Complexity

Behaviour Analysis, Complexity and Control of Networked Dynamical Systems

Lead Guest Editor: Hou-Sheng Su
Guest Editors: Zhiwei Gao, Xiao Ling Wang, Yue Song, and Miaomiao Wang

Behaviour Analysis, Complexity and Control of Networked Dynamical Systems

Complexity

Behaviour Analysis, Complexity and Control of Networked Dynamical Systems

Lead Guest Editor: Hou-Sheng Su
Guest Editors: Zhiwei Gao, Xiao Ling Wang, Yue Song, and Miaomiao Wang

Contents

Contents

Hindawi
Complexity
Volume 2022, Article ID 5714395, 24 pages
https://doi.org/10.1155/2022/5714395

Review Article

Self-Organization in Network Sociotechnical Systems

Svetlana Maltseva,[1] **Vasily Kornilov**,[1] **Vladimir Barakhnin**,[2,3] **and Alexander Gorbunov**[1]

[1]*National Research University Higher School of Economics, Moscow 101000, Russia*
[2]*Federal Research Center for Information and Computational Technologies, Novosibirsk 630090, Russia*
[3]*Novosibirsk State University, Novosibirsk 630090, Russia*

Correspondence should be addressed to Svetlana Maltseva; smaltseva@hse.ru

Received 6 September 2021; Revised 28 December 2021; Accepted 19 January 2022; Published 30 March 2022

Academic Editor: Yue Song

We can observe self-organization properties in various systems. However, modern networked dynamical sociotechnical systems have some features that allow for realizing the benefits of self-organization in a wide range of systems in economic and social areas. The review examines the general principles of self-organized systems, as well as the features of the implementation of self-organization in sociotechnical systems. We also delve into the production systems, in which the technical component is decisive, and social networks, in which the social component dominates; we analyze models used for modeling self-organizing networked dynamical systems. It is shown that discrete models prevail at the micro level. Furthermore, the review deals with the features of using continuous models for modeling at the macro level.

1. Introduction

The ability to self-organize is one of the key features of dynamic network systems. This quality is most evident on social media, in network communities, ecosystems, and various structures of the sharing economy.

There are several factors affecting the growth of sociotechnical systems, which provided room for interaction between entities of various nature: humans, intelligent devices or agents, and robots. Among them are the contemporary stage of interaction process automation; a surge in smart devices, intelligent equipment, and robots; and the development of artificial intelligence and machine learning. Even on traditional social media platforms, there are myriads of digital objects acting in the name of a human user.

Dynamic network systems, including those that can be attributed to sociotechnical systems, are usually focused not on rigid management systems but on the introduction of self-controlling, flexible, scalable, and distributed structures [1–3].

The idea of self-organization was formed a long time ago. Thus, R. Descartes put forward the hypothesis of self-organization as an ordering in the system due to its internal dynamics. W.R. Ashby formulated principles of the self-organizing dynamic system in 1947. The founder of synergetics, H. Haken, defines it as the science of self-organization.

The modern development and popularity of social networks and network communities has caused new, intensive research in the field of self-organizing systems, which is associated with the development of technologies that not only allow observing and analyzing the dynamics of social and economic systems but also developing applications for solving practical problems of influencing the audience of networks and the processes occurring in them.

Features of interaction in modern network systems, the dynamics of the processes occurring in them, and the principles of self-organization are not fully studied today.

Software agents became important participants of Internet communities in addition to humans. When the agents

have emerged as full participants in the processes, the term "Internet-of-agents" came into use [4, 5].

Multi-agent systems [6, 7] have shown that self-organization is effective when it comes to solving problems in various sectors such as transportation [8–11], logistics [12, 13], electric power [14, 15], e-commerce [16–18], health care [19], social services [20], etc.

The systems also revealed many problems that arise when self-organizing distributed systems are implemented, such as distributed problem solving, coordination and alignment, ethics and communication, as well as reliability and stability in critical states. For heterogeneous multi-agent systems, where the agents are designed by different developers, as well as for human-agent systems, the ethical rules for the agents [21–23] and their enforcement are crucial for interaction and problem-solving. The importance of these problems is shown in Ref. [23]. Evidently, ethical consequences of the present and future activities of the smart devices are vital, and we cannot be sure that intelligent autonomous devices would follow certain ethical rules. This prompts further research into machine ethics, as well as the ethics of the new technological stage as a whole.

Since the Internet of Things was introduced, network processes have featured smart things and tools. This calls for the development of new models, where smart things and tools are the end customers, signaling an imminent transition to a new stage, with self-organizing distributed systems being the foundation for the future organizational design.

Nowadays, "ecosystem" is one of the most frequently used terms in business. According to the complexity theory, an ecosystem is a complex of self-organizing, self-regulating, and self-developing systems. The models of biological ecosystems are admittedly essential in business applications because the principles of self-organization are crucial for the development of modern business and industrial structures. These artificial digital ecosystems are highly complex, indeterminate, and dynamically nontrivial. As of today, there are no theories that can give us an accurate understanding of their features, dynamics, and control of processes in them.

For self-organizing networked digital systems that control complex industrial processes, there are several issues to be addressed first: supporting the interactions between the participants following the process goals, observing the dynamics, and guaranteeing the stability of interactions. This, in turn, calls for improved methods to ensure process stability. Those methods should allow us to predict, reveal, and distinguish any critical phenomena in the system, which might lead to malfunction or even destruction of the systems, and subsequently to find ways to mitigate the adverse effect of those processes.

Critical phenomena in physics, thermodynamics, and other related fields include numerous anomalies observed at the points of second-order phase transition. Phase transitions describe the physical processes of transition from a medium state defined by a set of parameters to another state with another set of parameter values, and in thermodynamics, the critical point is the end point of the phase coexistence curve in which phases the thermodynamic equilibriums become identical in their properties. The classical theory of phase transitions was formulated by Landau and Lifshitz [24].

With the development of global networks and corporate networks, the study of dynamic processes and critical phenomena in disordered systems was also extended to network structures. Interest in the study of networks in this context is associated not only with a large number of their practical applications but also with the peculiarities of the inhomogeneity and fluctuations that arise in them.

In self-organizing digital network systems, the processes are led by machine-to-machine or human-to-machine interactions, with data entry and interpretation errors elimination being one of the major problems. The development of the robotic process automation (RPA) software designed to automate business activities formerly performed by humans, chat bots, and various recommendation systems helped to limit critical phenomena due to adverse effects of the aforementioned factors. However, there are still many problems where distributed resources and distributed problem-solving processes are not controlled efficiently enough, giving rise to critical phenomena.

Machine-to-machine interactions are dominated by signal and data transfer protocols. On the contrary, human-to-machine and human-to-human interactions are mostly performed through exchange of short messages, like in microblogging online communities. The data that need transferring in big volumes are picked from various existing data sources. The answer options could also be picked from a fixed set in a multiple-choice format. The critical phenomena that arise here could be caused by many factors, such as an untimely or incorrect response to a query, lack of requested information, mass-mailing of queries or opinions in the course of distributed problem solving, etc.

Modern networked dynamical systems are quite complex in technological and organizational aspects; therefore, controlling them creates multiple nontrivial challenges [25, 26].

This review examines two types of digital distributed systems: production systems and social networks.

In modern and promising concepts for the development of production systems, the technical component is the game changer; however, self-organization is considered a key characteristic of such systems, and decision-making and management in such systems are largely based on artificial intelligence methods and the study of the behavior of human communities.

Social networks are dominated primarily by the social component, although they have a complex and developed technical infrastructure and many digital objects.

2. Research Method

This systemic overview is an interdisciplinary study devoted to the development of network sociotechnical system self-organization models, with a focus on analyzing the use of these models in industrial systems and on social media. The main purpose of the review is to assess the current state and prospects of self-organization models in relation to the network sociotechnical system. The field under consideration is interdisciplinary, covering areas of application in computer science, physical sciences and engineering, and the

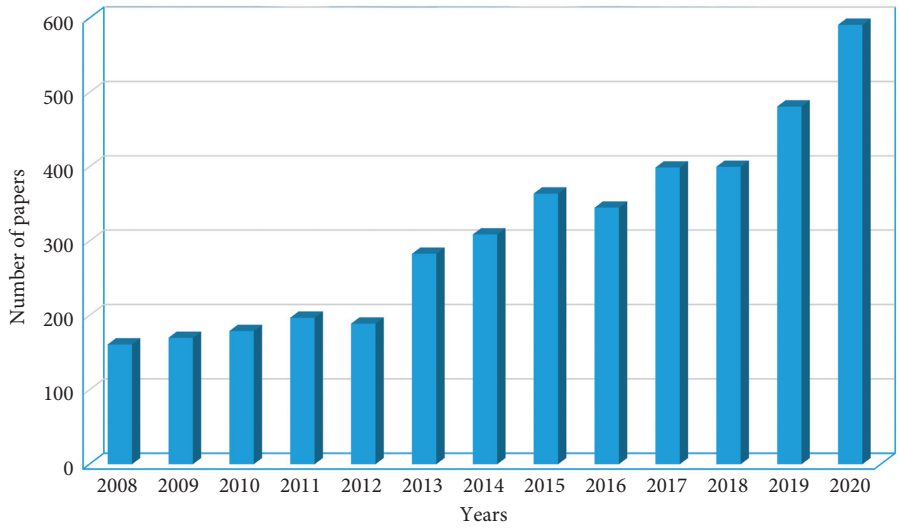


Figure 1: Papers published in the mentioned years.

social sciences. Although there has been a lot of self-organization research conducted over the years, the development of digital technologies has had a significant impact on the introduction of self-organization into modern systems. These aspects were the main reasons behind the choice of literature.

The overview rests on the literature accessible in Web of Science and ScienceDirect databases through the protocol PRISMA-P (Preferred Reporting Items for Systematic reviews and Meta-Analysis Protocols) [27].

English-language sources were mainly used to search for references, with the following keywords: (1) self-organizing sociotechnical system, (2) digital distributed manufacturing system, (3) social networks, (4) mathematical dynamics and modeling of self-organizing systems. We considered scientific articles, reviews, conference proceedings, and book chapters for the period 2008–2020. The subject fields were limited to Computer Science, Social Sciences, and Engineering. The initial search yielded around 4,000 scientific papers (see Figure 1). We can clearly see the global scientific community's growing interest in the subject.

To pick out the most academically valuable papers, we used additional criteria: publication level (a journal must fall into the first or second quartile of Web of Science) and citation index (a paper must be cited at least twice a year). To select the basic part of the list of publications, we compiled a thesaurus, including terms characterizing research areas associated with self-organizing systems, as well as with existing and promising digital technologies. When selecting basic resources, we factored in the bibliometric data related to authors (citations, number of cited publications), institutions, publisher, journal, and country. We then carried out an additional search based on the selected key topics; we sought to select the most significant publications, to show a retrospective of research, and to find interesting new studies that are often presented at conferences and workshops. Also, for the sake of completeness, we included a set of classic scientific papers and older monographs.

Four experts were involved in searching for and selecting the papers. One of them was responsible for the initial search, selection, and retrieval of bibliometric data and abstracts. Then, other researchers performed the expert selection of papers based on the aforementioned criteria. The full text of all picked papers was read and reviewed by each member of our research group. Outside experts were invited to resolve disputes.

More specific subjects–for which the papers were selected and organized–were defined and they formed the backbone of our review:

(i) Principles of operating self-organizing systems–Subsection 3.1–the principles of functioning of self-organizing systems.

(ii) Control in self-organizing systems–Subsection 3.2–general issues of control in self-organizing systems and features of control in digital distributed systems.

(iii) Self-organization in manufacturing systems–Subsection 3.3–the consideration of self-organization in distributed digital production systems.

(iv) Self-organized criticality–Subsection 3.4–the consideration of self-organized criticality and ensuring the stability of self-organizing systems, based on the study of social networks.

(v) Modeling of self-organizing systems–Subsection 3.5–mathematical models that are used to model self-organizing systems.

3. Results and Discussion

3.1. Principles of Operating Self-Organizing Systems. The self-organizing capabilities lead to a certain set of the system's functionalities, the list of which varies slightly according to different authors. The most commonly listed functionalities are as follows:

(i) self-monitoring;

(ii) self-adaptation;

(iii) self-healing;

(iv) self-configuration.

Some papers also mention self-tuning, self-construction, self-regulation, and self-reproduction.

As artificial intelligence and process automation evolve and provide additional possibilities, the list of the functionalities of the artificial systems will also expand. However, for artificial and heterogeneous systems and their artificial elements, this does not simply mean that they would approach human capabilities. Rather, this means a better adaptation to collaborative processing of objects of various natures, as well as enabling a safe interaction [28, 29].

Self-organizing systems that exhibit nonlinear dynamics are open, nondeterministic, exist far from equilibrium, and are based on the cooperative, collective behavior of their elements. There are many publications on the properties and behavior of such systems. The fundamental principles are considered in Ref. [30]. The following features of self-organizing systems are emphasized in Ref.[31]:

(i) *The endogenous global order:* a self-organizing system makes a transition from one stable state to another, and the final state is determined by internal processes in the system;

(ii) *Emergence*: due to the internal interactions between the elements of a self-organizing system, the properties (functionalities) of the system as a whole are not simply a sum of the properties (functionalities) of its elements;

(iii) *Simple local rules*: the overall complex system behavior can be based on simple (in terms of the amount of information) individual behavioral rules. Local information describes the mechanism for producing the global behavioral pattern and not the pattern itself;

(iv) *Instability*: a small variation in the system's parameters could lead to drastic changes in the system's behavior;

(v) *Several equilibrium states*: at the bifurcation point (when the system is unstable) there are many stable states for the system to pass to;

(vi) *Critical state*: this characteristic is related to the presence of threshold effects or phase changes. The exit from a critical state can be carried out as a result of the internal dynamics of the system. Also, the critical state can be maintained in a stable state due to internal dynamics (SOC phenomenon).

Nonlinear dynamics is seen in several important features of such systems.

Random small deviations of the system's parameters from their average values caused by either external conditions or internal interactions may grow and lead to drastic qualitative changes in the system's behavior. In some cases, one can observe a sensitivity threshold such that small changes below it do not result in significant consequences. It is difficult to predict the evolution and behavior of such systems based on the data for the prior intervals, because there is always a possibility of sudden changes in the pathway of processes in such systems, leading to a random and nonrepeating evolutionary path in the future. At the same time, a nonlinear system might have a limited number of possible evolutionary pathways determined by its spectrum of stable states.

The openness of the self-organizing systems such as social networks implies that the existence of such systems is ensured by information exchange with the environment. The input data (for example, a new current topic) increases the structural nonhomogeneity of the system, with new groups being created and the existing ones being regrouped. At the same time, the structural nonhomogeneities are getting smaller due to information dissipation: as time passes, the entry topic gets old, and the groups of network users that were formed around it cease to exist. However, the network could evolve; a stable group or a new community could emerge, involving some of the participants interested in the development of the topic [32].

In industrial systems and e-commerce systems, which are implemented as distributed digital human-to-machine systems, the entry feed is interpreted by devices and software as well as by humans. Such systems are not always created as open systems. External information or changes in the flows of external material resources can be the cause of noise in the system. Interaction with the external environment can also be used to get out of a critical state. For example, the systems created with Industry 4.0. are often considered as working under closed internal networks (intranet) only, to secure them from any external destabilizing information. Such systems are highly ordered and organized. However, they could also lose stability due to malfunction or lack of internal resources. For example, in manufacturing, if some parts are not supplied on time or stocked ahead at the warehouse, the planned manufacture processes could be seriously affected.

Open systems can better adapt to external perturbations due to self-organization, but at the same time they present a danger of an increase in entropy, loss of stability, and emergence of chaos.

As for hybrid human-machine networked dynamical systems, it is difficult to make them completely closed even with a high level of automation, and the increase in the AI potential is itself a source of poorly predictable disturbances due to external information influx. The complexity of human interactions with the machines while they perform tasks together is yet another source of instability [33].

Self-organizing systems are based on cooperative processes. These processes differ in the equilibrium and nonequilibrium states. In the nonequilibrium state, interaction is aimed not only at achieving the main goals but also at eliminating nonequilibrium states. Recent research has paid a great deal of attention to the features of interaction as a source of fluctuations which drive the system into a nonequilibrium state, as well as to the problems of maintaining the system's stability through the rules of interaction.

For example, in Ref. [29], the following methods of lowering the chances of the emergence of contradictions in a system are considered:

(1) *Tolerance* toward others and their goals contrary to own interests by way of internal adjustments or using additional resources to prevent conflicts and achieve common goals.

(2) *Politeness* as modifying one's own behavior and avoiding conflicts so as not to inconvenience other parties.

(3) *Compromise* as a combination of politeness and tolerance, when both parties have to modify their behavior to reduce conflicts.

(4) *Imposition,* when politeness is enforced by some limitations or internal changes.

(5) *Elimination* is a special case of the imposition strategy. An element of the system that does not contribute significantly to the system's development but causes conflicts by damaging the system's synergy can be eliminated.

(6) *Self-elimination* is a special case of politeness when an element eliminates itself for the sake of the system as a whole.

Obviously, these methods of maintaining the stability of the processes and the system's targeted development are imminent, in the first place, to biological systems, in which the participants are capable of evaluating their actions and their impact on the system as a whole. However, these methods become equally important for digital industrial systems in which heavy machine-to-machine and human-to-machine interactions are taking place. They also serve as a basis for self-organization, especially when the devices' AI reaches higher levels.

It is noteworthy that such functions are already implemented in modern devices. For example, many devices would not begin working unless the input data are entered correctly, or would stop operation if foreign objects appear in the work area, etc.

Whereas initial research on self-organization dealt exclusively with natural systems, these days a large part of research is aimed at artificial systems. This research is motivated to a great extent by the need to find new methods and strategies for controlling such systems in order to achieve the goals important for the development of economics and the society. Such innovation becomes essential because the traditional management schemes have mostly reached their limits of efficiency.

As shown in Ref. [34], for example, self-organizing production systems (SOPS) are aimed at the realization of Flexible Manufacturing Systems (FMS) and allow one to achieve high levels of automation and a wider range of manufactured products. This paper indicates the key features of SOPS that give them advantages in comparison with traditional systems. Among these features are the flexible automation of production systems and the possibility of automated planning, which allow them to manufacture a wider array of products. The production and control systems are distributed, meaning higher stability of the system, its increased capacity to respond to unforeseen and unpredictable events, and to quickly collect and present the information on the processes in the system.

These characteristics of modern production systems are listed in Ref. [35] as a basis for a set of metrics for the evaluation of modern production systems. The following metrics are listed:

(i) agility, which reflects the speed of response to unpredictable events;

(ii) adaptability in the course of processing products of various levels of complexity;

(iii) efficiency in the use of resources;

(iv) ability to reorganize, to change the system's topology in order to create new functionalities;

(v) self-organization, which reflects the system's ability for autonomous resolution of unforeseen internal issues.

Self-organizing networks (SON) are a separate class of self-organizing systems. Due to the automation and intellectualization of management, these networks have a dynamically changing structure; they can change and distribute functions between network nodes, for example, when a new device is connected or the traffic changes.

SON cases can be found in the sensor networks, in the mobile communications, and in transportation systems.

3.2. Control in Self-Organizing Systems. A large number of research papers on control problems in self-organizing sociotechnical distributed systems are available today. These papers deal with philosophical, social, ethical, mathematical, technological, economical, and many other aspects of such systems. In particular, there are many papers on self-organization and control in social networks. Principles and models of control in social networks are studied in detail in Ref. [36]. There are also papers that developed new concepts, for example, Refs. [37, 38].

Another direction of research deals with the applications of distributed and self-organizing systems to robotics systems and systems of intelligent agents [39–44].

Meanwhile, the functioning and stability of sociotechnical networked systems that contain elements of different origin remain some of the least researched problems.

Up-to-date results on stability and loss thereof in complex networks, including online social networks, are presented in the review [45].

It provides a classification of critical phenomena arising from loss of stability in networks via various mechanisms. It also presents the mechanisms and models that describe networks in equilibrium state and critical phenomena due to loss of stability. The loss of stability in networks leads to such critical phenomena as phase transitions, which are caused by a specific external impact (the fine-tuning of a control parameter), and to self-organization into a critical state (a self-organized critical state) without any specific external impact. Even though this review was published more than 10 years ago, it still reflects the state-of-the-art knowledge in this area. The main problem is the lack of a universal model (theory) that would allow us to describe, with good enough accuracy, the mechanisms and phenomenology of the phase transitions and network transitions into the self-organized critical

state. As of today, there are a number of specialized models that describe the loss of stability during phase transitions in social networks [45–47] and during transitions into a self-organized critical state [48–50].

Among the approaches to the modeling of self-organizing distributed systems, some of the most promising are based on the models that employ the order parameter, the conjugated field and a control parameter for the description of critical phenomena.

Such an approach allows one to design methods to predict, reveal, and identify the critical phenomena, as well as to develop recommendations on how to mitigate their negative impact and ensure the system's stability.

Many authors make a point of distinguishing between natural and artificial systems [28, 51].

Digital systems are artificial; their rules are determined by the developer according to the set goals, quality criteria, and the models of the system they are meant to control. They could be further improved in the course of their operation. A possibility for the improvement of the rules follows directly from the initial definition by W. R. Ashby, who defined self-organization as a process where some form of overall order of a complex system is evolved, reproduced, or perfected [52].

Today, the most complex systems are digital industrial systems, which combine various intelligent devices, use complex distributed machining processes, and which require machine-to-machine and human-to-machine interactions for their implementation.

The papers on self-organization in industrial systems [53–61] are aimed, in the first place, at the design of such systems and realization of their control systems.

In that connection, it seems interesting that the definition of self-organization in a technological system as a process of autonomous development of the optimal structure and the optimal algorithm of its operation according to the goal the system is supposed to meet, certain quality criteria, and external conditions [62].

A control problem for a group of robots with minimal human input is described in Ref. [63]. It is concluded that to design an artificial self-organizing system to control a process or object, one needs first of all to reveal or specify a set of local rules of self-organization which will be the basis for the specific structures and algorithms, which in turn would govern the control actions.

The target goal of performing a certain series of actions under the given conditions is considered as a problem on a directed, weighted graph, with its vertices corresponding to the actions needed to solve the problem and its edges determining the precedence relationships. At the same time, each robot in the group is characterized by a certain set of actions it is able to perform.

There are also additional requirements and limitations applied to the solution. The most complicated part is to develop the sets of actions, edges, and the model of the target problem. This can be done using two approaches:

(1) Creating the sets by the robots on their own in real time using AI methods.

(2) Creating the sets and the model by experts as ontological models and storing them in the knowledge bases of the group's robots.

The majority of researchers distinguish the following main mechanisms of control of a self-organizing system:

(i) Setting rules according to the goals of control so as to maintain an equilibrium in the system and making sure that those rules are followed.

(ii) Controlling the flow of resources and limiting their use.

(iii) Regulation of the noise level at the entrance and inside the system due to the development of the information environment, increasing the intensity of communications, and strengthening feedback. Analysis and regulation of external and internal noise allows, when adapting the system to new conditions, to make transient processes when it is necessary to change the state of the system and make its parameters smoother and more predictable.

Considering self-organization in sociotechnical systems, especially in production systems, researchers often investigate them within the framework of the concept of a System of Systems (SoS) [64–66]. SoS is composed of systems that are managerially and operationally independent.

The advantage of SoS is the ability to perform tasks that cannot be performed independently by any of its systems.

At the same time, the problem of such systems is that the potential of individual systems is not fully utilized when combining them; therefore, one of the important tasks is to ensure optimal performance through the use of various coordination mechanisms.

This approach makes it possible to increase the efficiency of distributed digital systems in many areas in which the level of autonomy of the combined systems is high, for example, in healthcare, transport, and others. This can often be observed in digital ecosystems as well, although not all authors regard them as sociotechnical systems.

3.3. Self-Organization in Manufacturing Systems. One of the groundbreaking works on the application of self-organization principles to manufacturing systems is by Kubota [67], where self-organizing manufacturing systems (SOMS) were defined.

SOMS are capable of reorganizing the hardware and software, and control in them is decentralized rather than centralized.

The main features of SOMS are generalized in Ref. [35]:

(i) Spontaneity: they adapt to external and internal changes without an intervention from outside by adjusting and reorganizing.

(ii) Distribution: controls are distributed among several autonomous nodes.

(iii) Emergence: SOMS as a whole may have properties not immanent to any of its elements.

(iv) Bottom-up method: the task arrangement is based on the solutions resulting from the communications between separate elements of the system, in contrast to making centralized solutions.

(v) Individual autonomy: simple tasks can be solved independently by individual components of the system, while complex problems are solved through component cooperation, taking into account their autonomy.

(vi) Self-learning: adapting to various media and self-optimization.

(vii) Reorganization: the ability to adjust and organize its own structure and create new functionalities in response to various demands.

The complexity of modern manufacturing problems has led to the evolution of manufacturing control systems. Self-organization fits well into modern distributed models for controlling large manufacturing systems.

These models are described as systems of interacting holons (the term was coined by Arthur Koestler in Ref. [68]) and include digital agents and digital twins [69–71].

A digital twin is a specific kind of model of a physical object. A feature of this model is the presence of a constant connection between two views of an object: the physical object that really exists and a virtual one–a model that contains the information about the physical object. The two views are related throughout the life cycle of an object. In the scientific literature, the idea was first described in Ref. [72]. Grieves [73, 74] defined a digital twin as a "set of virtual information constructs that fully describes a potential or actual physical manufactured product from the micro atomic level to the macro geometrical level." Two types of digital twins are distinguished: Digital Twin Prototype (DTP) and Digital Twin Instance (DTI). DTP contains the informational sets necessary to describe and produce a physical object. DTI describes a specific corresponding physical product that an individual Digital Twin remains linked to throughout the life of that physical product. Digital twins interact in the digital environment and can be combined as Digital Twin Aggregate (DTA) for different applications. The most important is the ability to analyze the current state of objects and the system as a whole and predict future behavior and performance.

As shown in many recent papers, the basic automation model such as the standard model IEC 62264 [72], which was meant to provide centralized control, is not always efficient as a means of controlling large systems with numerous intelligent devices. Such systems are able to quickly and correctly react to the changes in the environment and adapt to them without interference from outside.

This ability is a result of digitalization of the majority of processes. It led to the evolution in control systems and the development of new decentralized concepts of the system design.

According to the basic concepts that form the foundation of the system architecture and topology, we distinguish Fractal Manufacturing Systems, Bionic Manufacturing Systems, and Holonic Manufacturing Systems [75–77].

Fractal Manufacturing Systems (FMS) are developed to make traditional manufacturing models more flexible using the ideas of self-organization and the methods that help mitigate the critical phenomena, which could result in a transition to a chaotic state.

FMS are open; their main feature is the self-similarity of their fractal units.

FMS are focused on the use of analogy with the forms of organization of natural systems in the development of a framework for the planning, design, optimization, management, and assessment of processes to supply sustainable manufacturing under environment pressure through self-organization. In the FMS, the common features of fractals are: similarity, self-organization, self-optimization, and dynamics. The organizational model assumes that all departments and even individual employees are focused on achieving business goals and are capable of entrepreneurial innovative thinking and activity. In Ref. [78], four principles of fractal sustainable manufacturing are considered: the principle of manufacturing organization fractal; the principle of organization derived from the variety or fractalicity required; the principle of fractalization of sustainable product's life cycle; and the principle of fractal levels of sustainable manufacturing.

The concept of Bionic Manufacturing Systems (BMS) is based on the similarities between the manufacturing and biological processes, which ensure the proper functioning of all the individual parts and the entire organism by analyzing internal and external events and processes, adjusting the behavior of the organism and its components, and reproducing new elements. Ueda [79] was the first to introduce the concept of such systems based on the idea that dynamic adjustments are the result of self-growth, self-organization, adaptation, and evolution.

The basic element of BMS is called a modelon. A modelon could be composed of lower-level modelons, forming a hierarchical structure.

The information is exchanged both within and between the modelons. Similar to living organisms, the BMS possess components or mechanisms for adjustment and control which impose organizational and structural rules of interaction between modelons, so that they work toward the goals of individual processes and the common goals of the entire system [80, 81].

Holonic Manufacturing Systems (HMS) have a far higher level of self-organization than the bionic systems.

A holon's concept states that a holon has two main parts linked to the informational and physical processes. The informational part is related to the internal communications, decision-making processes, and human-to-machine interactions, whereas the physical part is related to the control over physical processes and equipment.

A holon is autonomous; it can draw up its own plans, strategies, and oversee their implementation. It can also cooperate with other holons while following certain rules, as well as be a part of another holon.

Holons interact with each other in the course of manufacturing processes, thus forming holarchic control structures. Initially, the interaction of holons was very

limited and did not have the character of developed negotiations, although they received useful information and coordinated their actions. The expansion of the possibilities of interaction between holons is largely associated with the development of multi-agent systems in the holistic paradigm. The concept of the Holonic Multi-Agent System is based on complete decentralization of management, when management occurs only through local interactions between agents. In this case, the agent can be the entity that focused not only on solving a separate part of the general problem. It can be an integral artificial active entity capable of solving various tasks to support its existence. In Ref. [82], it is proposed to create manufacturing or functional nodes that combine a holon and a digital agent.

The agents and holons can be based on physical resources or logical objects [82]:

(1) The functional approach: various units are used to represent functional modules in manufacturing systems; there are no direct relations between the abstract units and physical entities.

(2) The physical approach: various units are used to represent the physical entities; there is a direct relation between the abstract units and physical entities.

The properties of a digital agent in this case must ensure the interaction between all the nodes of an intelligent manufacturing platform. A digital agent within a holon must meet the following requirements [83]:

(i) it should be able to perceive and respond to the environment;

(ii) it should be goal oriented;

(iii) it should possess a large enough knowledge base to act independently;

(iv) it should be able to interact with other agents;

(v) it should be able to learn from prior behavior;

(vi) it should be mobile within the network.

When implementing this model, it is important to establish the correspondence and interaction between the production holon definition and the connected agents through the mechanisms of communications, negotiations, coordination, and collaboration between various elements of the system.

The known interaction mechanisms alone cannot solve the problem in its entirety. Particular solutions are aimed at the reduction of complexity, localization of information, reduction of communication volumes, and optimization of the global behavior [84].

The following coordination mechanisms are better known:

(i) "*Blackboard*." There is a common area of shared information where each participant posts the information on its demands and supports, and all the parties mutually coordinate their actions according to the information on the board.

(ii) *Market-based mechanisms* are based on auctions and contract networks: the participants list their tasks on the "market," then the relevant parties decide if they should bid to win contracts to perform those tasks.

(iii) *Mechanisms that model the processes in biological systems*, such as those used in BMS, are based on evolutionary processes, regulation of hormones, reinforced learning, and others.

Some other mechanisms of self-organization that could also be adapted to manufacturing control are presented in Refs. [85–87].

In human-machine systems, it is especially important to ensure error-free communication between the machine and the human, not only the communication between machines, because incorrect responses due to communication errors may often cause critical phenomena.

Digital twins play an important role in the improvement of control processes in self-organizing systems.

They model real objects or processes, and they model internal processes, technical features, and the behavior of a physical object under the action of perturbations and the environment. Digital twins not only contain current information on the state of the object, but historical data as well. This makes it possible to obtain accurate information about the system's productivity, predict the future states, and control the object or the process remotely in real time. Digital twins of the personnel responsible for certain functions can also be used. In fact, they are replaced by avatars, which are adapted to the information exchange formats and protocols in cyber-physical systems.

The application of digital twins [88–90] reduces uncertainty and risks in the system, because when a human worker is replaced by a software agent, the worker's functionality is better understood by digital devices interacting with them. This helps to eliminate most errors stemming from human-to-machine interactions. This makes the digital twins similar to the RPA-systems, which help reduce the number of errors by reassigning some processes from humans to automates.

Modern approaches to the SOMS implementation are reflected in the Reference Architectural Model Industry 4.0 (RAMI 4.0) [91, 92].

Technological developments ensure the improvement of the model's components in all three of its main aspects: the architectural levels of the ICT system, the Hierarchical Levels of interconnection of the elements in the manufacturing system, and control over the Life Cycle and Value Stream.

Technologies that help combine the physical world with cyberspace and enable the structural blocks of cyber-physical systems to communicate, make decisions, and perform control are crucial for development.

Smart manufacturing equipment becomes more intelligent and autonomous. This allows us to considerably broaden the array of functions performed by machines and robots, and modify the topology of the manufacturing systems, while enhancing its capability for self-organization.

To implement a Smart Enterprise that would properly interact with the environment, it is critical to develop the concepts of a Connected World and Smart Products.

The use of the Service-Oriented Architecture (SOA) and cloud computing provides the components of a manufacturing system with the abilities to adapt and reorganize quickly, makes the entire system more flexible, and makes distributed control and self-organization possible.

The development of big data technologies and the methods of predictive and prescriptive analytics greatly enhance the possibilities and functionalities of the system's architectural layer. This is especially true when it concerns predicting and automated decision-making, which help mitigate and reduce the risks and critical phenomena due to human-to-machine interactions [93].

Research is still far from being sufficient on the application of quantum computing to the manufacturing area. Some possible applications of quantum computing in manufacturing stemming from McKinsey's predictions are given in Ref. [94]. The main tools are quantum modeling, optimization, and artificial intelligence. The improvement of yields and suppressed by-product generation through better understanding of reactions, finding new catalysts, and the use of quantum algorithms to solve complex optimization problems of heat and mass transport are given as examples of future applications.

In Ref. [95], prospects for the use of quantum computing in various fields are considered. Materials science, advanced analytics for control processes, and risk modeling are highlighted in manufacturing and industrial design. In logistics, it is possible to solve many problems of supply chain optimization, risk modeling, increasing the speed of service, easier adaptation to changes, for example, in cases of canceled orders or rescheduled deliveries. Today, the financial sector is showing considerable interest in the prospects for the use of quantum computers [96]. The great interest in quantum computing in the social sciences as a whole should be noted [97–99]. Cybersecurity is also an area that quantum computing will significantly change.

In Ref. [100], business impact of the application of quantum technologies is considered for the following areas:

(i) material science: discovery of new candidates for drugs and materials with better properties; reduced time to market; and reduced number of real world trials;

(ii) engineering and design: improved model and simulation quality lead to better quality; faster time-to-solution provides process efficiencies;

(iii) production and logistics: faster and more efficient production and supply chain management; improved quality; reduced emission.

3.4. Self-Organized Criticality. The theory of self-organized criticality (SOC) has been widely used in natural sciences over the past decades.

The key notions of the general theory of phase transitions by Landau and Lifshitz [24] are the order parameter η, which characterizes the state of the system, and the so-called Landau functional, which determines the energy associated with the spatial variation of the order parameter. Landau interpreted a phase transition as the moment when the system's symmetry changes: above the transition point, the system has a higher symmetry ($\eta > 0$, "order") than it has below the transition point ($\eta = 0$, "chaos"). Such critical states occur in a system only when there is an external impact on it; formally stated, this happens when the control parameter S reaches a certain threshold value S_C.

In Refs. [101, 102], it is noted that a critical state can occur spontaneously as a result of self-organization in the system. A classical model developed in the late 1980s by Per Bak, Chao Tang, and Kurt Wiesenfeld within the framework of the concept of self-organized criticality (SOC) (the BTW model) is a sandpile model. As sand is added, the system approaches a critical state, where addition of one more grain of sand causes an avalanche: the components of the sandpile system transition from the chaotic state to that of a regular movement. The system's order parameter (flow of sand) is slightly greater than zero, $\eta = +0$, while the control parameter S, in this case the slope of the sandpile's surface, independently reaches a certain critical value S_C regardless of the initial state. It has been proved that the size and frequency of avalanches occurring in this critical state follow the power-law distribution [103, 104]. This result, in general, is typical of critical phenomena, when a system responds to some minor event by a catastrophic change in its state.

The SOC theory was used to interpret a wide variety of phenomena in nature and society, for instance, economics [101–107], biology [108–111], earthquakes [112], political science [113, 114], sociology [115–117], brain functions (neural networks) [118–123], agriculture [124], and other fields.

Critical state in dynamic systems is one of the fundamental concepts of physical phenomena. This is a special state such that the system loses stability of a less organized (disordered) state and makes a transition to a more ordered state at the so-called bifurcation points. When the system reaches this state, a small disturbance can completely change the system's behavior. This is a classic case when quantitative changes transition into qualitative ones.

Self-organized criticality emerges in systems that possess the above self-organization functions and consist of many elements connected directly or indirectly by cause-effect relationships. This allows us to link the system's dynamics at the macro level with the behavior of its elements at the micro level, thus ensuring its integrity.

Simple events at the micro level may have consequences that affect many other events in the system over time and trigger causal waves that can superpose, thus amplifying or attenuating each other. These processes can also be resulted in the change of parameters describing the dynamics of the system at the macro level.

Chains of micro-level events lead to the appearance of the *1/f*–pink noise and then of avalanches. At the same time, the system can stay in a critical state relatively long if stability is enforced one way or another.

The classic BTW model with its many versions as well as other models can be considered as cellular automata. Other

approaches used in the modeling of self-organization include multi-agent systems, evolutionary computing, neural networks, and hybrid intelligent systems.

The following models are used most frequently:

(i) Stick-slip model of fault failure [125];

(ii) Forest-fire model [126, 127];

(iii) Olami-Feder-Christensen's model, an earthquake model [112, 128].

(iv) Bak-Sneppen's model, BS-model, co-evolution model [129].

The examples of physical processes with SOC behavior are given in Ref. [130], where one can find those used as a metaphor for the developed models, as well as those used in the models of interaction in the digital environment, for example, traffic collisions, stock market crash, or a lottery win.

Very often SOC effects are observed in the agent-oriented models [131].

A large number of studies deal with the applications of the SOC theory for the simulation of the information propagation on the Internet, in particular, in social networks [48, 49, 132, 133].

In the community of users of a social network, a random message that is not interesting to other users and not retransmitted by them does not lead to a change in the state of the network. If, on the other hand, a message is retransmitted by a critical number of active users of the social network, it quickly becomes popular, and at some point this can lead to a system-wide transition to a critical state, which is characterized by a spontaneous increase in user activity and an avalanche of messages. The situation described above is very well in line with the concept of a self-organized critical system. Indeed, SOC-based models have been successfully used in studies of various phenomena observed not only in social networks but also in social and sociotechnical systems.

For example, in Ref. [133], empirical data and agent modeling are used to analyze the phenomenon of collective emotional behavior of users which is often observed on various web portals. The emphasis is on the quantitative evaluation of the collective emotions through fractal analysis of the underlying self-organizing dynamic processes, as well as the topology of the social networks which occur and coherently develop in these stochastic processes.

Initially, fractal market analysis was introduced by Mandelbrot and Hudson [134]. Fractal analysis is now widely used in all areas of science.

Fractal analysis consists of several methods to assign a fractal dimension and other fractal characteristics to a dataset. Fractal analysis is valuable in expanding our knowledge of the structure and function of various systems. Refs. [135, 136] discuss the use of fractal analysis of social networks.

Data from the Twitter social network are used in Ref. [132] to study fluctuations in the frequency of brand tweets. The frequency of tweets is an outcome of strongly correlated user behavior, which leads to turbulent collective dynamics with the characteristic $1/f$ noise. An integral parameter measuring the user's interest in a brand is used to simulate the collective human dynamics using a stochastic differential equation with multiplicative noise. The model is supported by a detailed analysis of the fluctuations in the rate of tweets; it reproduces both the exact dynamics of peaks and the $1/f$ noise.

From the point of view of the complexity paradigm, a microblogging social network can be described as a nonlinear, nonequilibrium dynamic system in a three-dimensional phase space. The flows of messages (posts) in microblogging networks are prone to the emergence of critical states, i.e., avalanches of messages. Analysis of the time series for these messages reveals several features of a classical SOC model: power-law statistics of the probability distribution, $1/f$ flicker noise of the power spectral density of the time series, and the presence of events that can be viewed as catastrophic [137–141].

The Twitter network is considered in Refs. [142, 143] using the SOC model, because Twitter is prone to avalanche-like message streams, from tens to thousands of posts per second. It is hypothesized that Twitter's self-organization in a critical condition is the result of a special persistent ("strategic") behavior of a relatively small number of users.

All Twitter users can be tentatively divided into two groups: a small number of strategically oriented users (SOUs), and the majority of randomly oriented users (ROUs). The basic state of the network, when messages are generated by ROUs, corresponds to a chaotic state in a classical dynamical system, since these messages are not connected to each other and quickly disappear from the user's horizon. The critical state occurs when the number of coherent SOUs messages aimed at the same outcome reaches a certain critical value, such that yet one more SOU causes the network to transition into an ordered state and creates an avalanche of messages.

Similar ideas were also voiced by other authors. For example, it is suggested in Ref. [144] that each social network has "influence agents," i.e., users holding certain key positions in the network, which gains them a strong structural ability to influence the entire population of users. The activity of influencers more often than the activity of other network users can lead to the emergence of critical states characterized by macroscopic effects.

Arguably, the social microblogging networks develop holistic properties immanent to complex systems. This conclusion is supported by a power-law dependence of the autocorrelation function for the messages time series. This means that the current number of messages depends mostly on the number of messages generated in the social network in the past. Here, we see a manifestation of the characteristic ability of SOC systems to possess a long time memory [145].

It is noted in Ref. [146] that, when a social network's dynamics is based on human effort, it is difficult for it to acquire new features such as collective social values. The vast amount of empirical data collected from various websites provides a unique opportunity to quantify social dynamics the same way this is done for complex physical systems. It is suggested that the dynamics of social knowledge exchange is

governed by the SOC mechanism; moreover, the emergence of hyperbolic geometry in such social systems was demonstrated.

It is noteworthy that self-organized systems do not always exhibit the properties of criticality. Research and differentiation between the SOC, SOC-like, and non-SOC processes was done in Ref. [130]. Based on the proposed metrics, the difference between various SOC or SOC-like processes was studied such as The Exponential-Growth model (EG-SOC), Fractal-Diffusive model (FD-SOC), Forced Self-Organized Criticality model (FSOC), Self-Organization Without Criticality (SO), Brownian Motion and Classical Diffusion, Hyper-Diffusion (thresholded), Levy Flight (thresholded), Nonextensive Tsallis Entropy, and Turbulence (laminar/turbulent). Power law of spatial scales, power law of time scales, power law of total energy, power law of energy dissipation rate, fractal geometry, intermittency in time evolution, statistical independence of events, critical threshold restoration, next-neighbor interaction, and nonlocal (long-range) coupling were used as metrics to quantify the system's properties.

An important result of Ref. [130] is the differences shown to exist between SO and SOC in terms of stability of the system states. Self-organization can be characterized by rather long intervals of stability of the emerging landscape of more or less stable areas of the entire system. This relative stability is regulated via long-term interactions through rules, pressure, and forces. SOC generates dynamic events (avalanches) when intermittent avalanches occur due to random-like disturbances.

One of the most important tasks for digital distributed systems, which have applications in various fields, is to ensure their stability.

Stability is the ability of a system to function without changing its own structure and to be in balance. A system is said to be stable if it does not exhibit large changes in its output for a small change in its input, initial conditions, or its system parameters. In a stable system, the output is predictable and finite for a given input.

An understanding of the minimally stable state is essential also in understanding the self-organized critical state [102].

In Ref. [147], it is noted that there is a class of complex systems with a large number of degrees of freedom that enter a critical mode by their very nature, as a result of the internal tendencies of evolution inherent in these systems. Critical states of such systems do not require precise adjustment of external control parameters and in fact have the property of self-support.

For real sociotechnical systems, it is important to detect and analyze critical phenomena and assess the state of the system in order to, if necessary, improve it in time, ensuring stable functioning.

A sociotechnical digital distributed system is a complex network structure. As shown in Ref. [147] "critical phenomena in network structures include a wide range (forms of behavior) of phenomena: structural changes in networks, the emergence of a critical state, various percolation phenomena, critical points in various optimization problems, and many others. Many of these critical phenomena are closely related, have a similar nature, and allow for a universal description."

Critical phenomena in such systems are primarily connected with the risk of loss of control and disintegration of the system. Since in such digital systems, especially manufacturing systems, the processes involve both equipment and people, critical phenomena may endanger not only the manufacturing processes but also people's health and lives, as well as the environment (sometimes, the criticality could instead allow the unlocking of situations otherwise blocked, or in deadlock). Hence, it is extremely important to study the critical phenomena and the methods of their prevention.

The problems of control of nondeterministic dynamics in self-organizing systems, an optimal balance between external impact that adversely affects the reliability, adaptability, and scalability of the system on the one hand, and the amount of uncertainty, which could make the verification and validation of the system an impossible task, on the other hand, are discussed in Refs. [148–152]. It is shown that finding such a balance helps solve various practical problems, for example, in traffic control [149], design of robotic teams [150], data visualization [151], development of services in Grid networks [152], and many others.

Many studies explore the methods and principles of making decentralized decisions in multi-agent systems [86, 153–158].

Cyber-physical and cyber-social systems are becoming more complex, and at the same time, the number of different approaches to modeling of those systems in the context of self-organization is growing.

Special issue [159] discusses an increase in the complexity of digital sociotechnical systems, including social media, the Internet of Things, RPA systems, digital business platforms, algorithmic decision-making, digital networks, and ecosystems.

Phase transitions in business processes supported by digital technologies are simulated in Ref. [160]. Gradual endogenous changes in the system may lead to a state of self-organizing criticality. As the process approaches this state, further incremental changes may result in nonlinear surges in the process complexity and significant modifications of its structure.

The review [161] examines the simulation of complex dynamic network structures where interactions occur within groups of three or more nodes. It is noted that such models correspond well to the architecture of real complex systems. Taking into account the higher-order structures in those systems could widen the scope of simulations and help with understanding and predicting the system's dynamic behavior. This could also help to provide an adequate description of the state of self-organizing criticality that is typical of many similar systems.

To characterize the stability of a self-organizing system, it is important to quantify the level of self-organization, since a tendency of gradual degradation of self-organization and transition to hierarchical control can occur in sociotechnical systems. On the other hand, technical systems may gradually

evolve toward self-organized criticality. A definition of quasi-stationary self-organization systems, when the system parameters vary much slower than typical dynamical motions in the system, is given in Ref. [162]. Self-organization in complex systems is regarded as a process of human decision-making, and vice versa; the decision-making is considered as self-organization in the nervous system of those who make decisions. A similarity between those processes is also demonstrated in a number of other publications, for example, in Ref. [163].

A number of studies present approaches that make it possible to quantify the state of self-organization in a system. The characteristics of structures, patterns, scenarios, or prospects are associated with specific states.

Several well-known metrics for quantification of self-organization are considered, such as the Shannon entropy [164], the von Foerster redundancy [165, 166] and some others, as well as the new approach proposed by the authors and based on the methods of the quantum theory of measurement [167, 168] and quantum information theory [169, 170].

The interest in self-organization as a process similar to the decision-making process reflects several important trends related to the development of digitization and digital systems in various fields. Those trends include progress in the intelligence and autonomy of equipment and robots, an increase in the rate of information exchange and the volume of information traffic, and an ever greater degree of process automation. Given all that, it is likely that a better understanding of theories and practices of decision-making in human communities, and subsequent application of them to digital systems, could help achieve a greater stability in self-organizing heterogeneous systems.

3.5. Modeling of Self-Organizing Systems. Mathematical modeling of sociotechnical networked self-organizing systems is a nontrivial task, because, as was shown above, they necessitate joint participation of and interactions between humans and miscellaneous digital equipment, including intelligent tools.

Modern digitalization capabilities allow for an agent-centric approach to these systems which combines the physical or functional elements of the systems with the agents that ensure the communications.

An overview of various methods of modeling self-organizing systems is presented in Ref. [171]. Modeling at the micro level can be used to describe the behavior of each element of the system as well as the communication between the elements. It has an advantage of providing a more detailed description of the real system, but the downside is a high dimensionality of the model's global state space. At the macro level, models do not deal with individual entities of the system and consider their equivalence classes instead, thus reducing the vast micro-level state space. At the micro level, the properties of the models of sociotechnical distributed systems discussed here (time, space, etc.) are conceptually discrete in nature, while, as a rule, synchronous time renewal is used, while at the macro level these properties are continuous.

Thus, two main approaches can be distinguished in mathematical modeling of digital distributed systems:

(a) discrete *ad hoc* models;

(b) discrete models based on the ideas borrowed from mathematical modeling in biology;

(c) continuous models using analogs from biology, mechanics, etc.

Among them, the most widely used approach relies on discrete modeling [172–180].

That is hardly surprising: a relatively small number of agents in the system allows us to formalize and numerically describe the interaction mechanisms between specific agents. On the other hand, this is exactly what makes a transition to "continuous models" difficult.

However, when modeling social networks with a huge number of actors (for example, the number of Twitter microblogging users exceeds 300 million), it becomes almost impossible to take into account the individual properties of each actor; therefore, in this situation, along with discrete modeling [181–185], an approach based on the use of systems of ordinary differential equations is used. The independent variable in such models is time; dependent are various quantitative characteristics that describe the number of tweets, retweets, etc.

Some discrete models are developed *ad hoc* [172–177], but more often, well-known models from biology based on the genetic algorithm or models of some aspects of the insect behavior (bees, ants) [179, 180] are used as a foundation for discrete models.

Continuous models [181–184] also often use analogs from biology and from hydrodynamics.

Considering particular models of digital distributed systems, we can evaluate the applicability of these approaches to specific systems.

3.5.1. Discrete Ad Hoc Models. A general method of analysis of multi-agent systems, which allows us to define strategies for local interactions by specifying a preferred global behavior of the system, and then to evaluate possible strategies using Markov analysis, is proposed in Ref. [172]. The choice of cooperation strategies is based on the iterative application of the "prisoner's dilemma." As shown, it is similar to the "tit for tat with forgiveness" strategy, which, under certain circumstances, surpasses the well-known "tit for tat" strategy. The latter is used, for example, in the peer-to-peer file-sharing network BitTorrent.

(1) A distributed, service-oriented, multi-agent system where the agents must communicate with each other to perform decentralized tasks of service discovery is considered in Ref. [173]. Since the system's structure may affect the efficiency of the service discovery, a structural self-organization mechanism should be used to facilitate the decentralized discovery.

During the discovery of decentralized services, different agents will present different intermediate utilities, and those utilities imply, to a certain degree, information about the distribution of the system's services. Hence, an intermediate utility can be used to improve the efficiency of both structural self-organization and the search for services. Moreover, the availability of global services is ensured by maintaining the global connectivity of the system's structure during parallel self-organizational processes.

A system is defined as a pair $\langle A, L\rangle$, where $A = \{i, \ldots, n\}$ is a set of agents, and $L \subseteq A \times A$ is a set of relationships between the agents; each relationship indicates a symmetric interaction between agents i and j.

Agent-to-agent relationships have a locked flag:

$L(i, j)$ indicates a locked relationship between i and j;

$L(i, j)$ indicates an unlocked relationship between i and j.

If a relationship becomes locked, it means that one of the agents intends to break it.

If a request for service is successful, the agent i is considering changing its relationships. If the following holds:

$$\max_{k \in A_i^K}\{\eta_{ik}\} > \min_{j \in A_i^N \wedge L(i,j)}\{SV_{ij}\} + \gamma, \tag{1}$$

where η_{ik} is an estimate of the potential weight (cost of creating a relationship) for the acquaintances $k \in A_i^K$ of agent i; SV_{ij} is the system value of the relationship between agents i and j, under the condition $SV_{ij} = \max\{\tau_{ij}, \tau_{ji}\}$; and τ_{ij} is an estimate of the weight for the neighbors j of agent i; $\gamma\,(\gamma > 0)$ is a constant.

In this case, it is possible to improve the quality of communication set of the agent by structural self-organization as follows:

(1) Agent i sends a message to agent j with the minimum SV_{ij} a message about blocking the communication (i, j); if agent j has already blocked the communication, it will respond with a denial; otherwise, it will send back a confirmation;

(2) If agent i receives a denial, it does nothing; if agent i receives a confirmation, a new type of request AQ is generated to find the appropriate agent to form the relationship $AQ = (i, j, DI, TTL', k, A_v' = \{i\})$, where i and j are the above agents' numbers; DI is the demand information; TTL' is the maximum number of requests forwards; k is the number of the acquaintance with the largest η_{ik}; and A_v is the set of agents that received the request in the course of forwarding.

(3) If $TTL' < 0$, the request is sent back to i. Otherwise, if $TTL' \geq 0$, then agent m decrements TTL' in the query by 1 and, if $m \notin A_v'$, includes itself in A_v';

(4) The request is forwarded to agent m^*:

$$m^* = \begin{cases} \underset{n \in L\left(A_m^N\right) \wedge n \notin A_v'}{\operatorname{argmax}} \{\eta_{in}\} \text{ if } L\left(A_m^N\right) - A_v' \neq \varnothing, \\ \text{precursor}\,(m) \text{ otherwise.} \end{cases} \tag{2}$$

Here, based on empirical estimates, it is assumed that $TTL' = 10$, $\gamma = 0.1$.

Compared to a number of known approaches (*simplicity-based search without structural adaptation*, *simplicity-based search with deterministic structural adaptation based on service property*, *simplicity-based search with probabilistic structural adaptation based on a decay function*), this model increases the efficiency and reduces the time needed for the system to come to a stable state, especially when it is adapting to varying content of the search queries.

(2) Two mechanisms of self-organization for a decentralized service discovery system that improve its performance are described in Ref. [174]. These mechanisms are based on local activities of agents, which only consider local information about the requests that they forward during the discovery process.

The following self-organization activities are noted: staying logged in, logging out, cloning, and changing the structural relationships with other agents. Each agent independently picks the actions it deems appropriate.

A system is defined as a Service Oriented Multi-Agent System (SOMAS): $\langle A, L\rangle$, where $A = \{i, \ldots, n\}$ is a set of agents, and $L \subseteq A \times A$ is a set of connections between them; each connection $(i, j) \in L$ indicates the presence of a symmetric interaction between agents i and j.

The function of the search for the most promising neighbor is defined as follows:

$$\mathcal{F}_N(t) = \underset{j \in N_i}{\operatorname{argmax}}\left[1 - \left(1 - \left(\frac{CH(j,t)}{\sum_{n \in N_i} CH(n,t)}\right)\right)^{\left|N_j\right|}\right]. \tag{3}$$

Here $t = (r_q, s_q, \varnothing, \varnothing)$ is a fake agent representing an unknown agent-provider of the requested service (or similar to it); CH is the Choice Homophily, which evaluates the proximity of agent-neighbor j to the fictitious agent i:

The function of selection of the local action of self-organization is defined as follows:

$$\mathcal{F}_A\colon st_i \longrightarrow \psi. \tag{4}$$

Here $\psi = (\text{clone}, \text{remain}, \text{leave}, \text{rewire})$ is a set of possible actions. Structural self-organization of agents follows CH.

The mechanism of structural self-organization promotes a decentralized discovery of the services in the system. The algorithm that supports the mechanism of self-organization consists of the following stages:

(i) the agents search for services;
(ii) the agents verify whether the local information is complete and reliable;
(iii) the agents decide whether to maintain, strengthen, or create new structural relationships;
(iv) the agents decide whether to stay logged in, log out, or get cloned.

Discrete models are sometimes used to analyze the behavior of the users of social networks (when the test data volume is relatively small). The model proposed in Ref. [175] can be used to integrate multiple social networks, import and process data by means of the modern big data technologies. This model allows one to record the activities of Internet users taking into account the combination of human and temporal factors, as well as to reveal both positive and negative trends in the evolution of opinions formed by network users.

Agents (Internet users) are represented by u_i, where $i = 1, \ldots, N_u$ is the number of users. The information exchange between users is expressed via posts, messages, or comments p_j, $j = 1, \ldots, N_w$.

Thus, a message is generated as an event of the following type:

$$g_{i,j} = \left(u_i, p_j, t^0_{i,j}\right), \tag{5}$$

where t is the time of the message generation.

The process of the message sending or processing is represented by event $e_{i,j}$. This is a logical variable that depends, among others, on the focus, i.e., the current scope of interest of the user; the focus is described by a cloud of weighted tags (keywords). Each user has his or her own ontology, which determines the basis of their perception (messages). It changes over time as a result of the process of remembering and forgetting the information received from messages. The ontology is represented by a sequence of contexts defined by tags. Changes in focus and context correlate with each other; their correlation is described by a logical variable. Therefore, the generation or processing of a message does not guarantee changes in focus and context.

The software that identifies the focus of social media is based on the knowledge search and big data analysis, which can also help to elucidate the impact of online bots.

In some cases, discrete models can be self-learning. In particular, there is a discussion in Ref. [176] of self-organization mechanisms that can build, in an automatic, flexible, decentralized manner, services for adapting the agents' local behavior and the structural relations between them.

The model defines the following roles of the agents: *consumer*, *supplier*, *work process*, *ontology*, and *reputation*.

The model employs an approach based on decentralization. To that end, dynamic communities of service-oriented agents are created, and requests are directed to those communities. By automatically creating new connections, a useful agent can seamlessly join a community. The community, on the other hand, is able to break connections with useless agents.

A novel self-learning mechanism, which allows distributed objects to change their structural relations and thus enable the system's evolution, is organized as follows. An agent makes decisions based both on its current state and on the quality of a potential reward for performing an action in that state. The reward is based on the agent's status, action, and feedback. The calculation of the reward is done in two stages: first, the reward is evaluated qualitatively, then it is converted into a numerical equivalent. If the system's evolution is stable, i.e., this process converges, then the agent is considering the possibility of creating new structural connections.

This approach improves the quality of service and reduces the network traffic.

3.5.2. Discrete Models Based on the Ideas of Mathematical Modeling of Biological Systems. The genetic algorithm is one of the traditional approaches. It is applied in Ref. [177] to mobile *ad hoc* networks (MANET), which are widely used in various situations, from military to business tasks, including disaster zone detection, mine field clearance, and transportation systems.

In real-world environments, it often does not make sense to deploy network nodes manually or using a centralized controller. To achieve self-organization of mobile nodes in an unknown area, the following nature-inspired approach is the most promising: each mobile node (agent) uses *the genetic algorithm* as a *self-propagation* mechanism. This allows it to determine its own speed and direction and ultimately to evenly distribute all network nodes.

Each agent (node) is characterized by a set of chromosomes of the following type: $\langle d_1, d_2, d_3, s_1, s_2 \rangle$, where d_1, d_2, d_3 are the direction bits, and s_1, s_2 are the speed bits.

Let us consider two nodes N_i and N_j. The force of attraction F_{ij} between the nodes N_i and N_j is expressed as follows:

$$F_{ij} = \begin{cases} F_{\max}, & \text{if } d_{ij} = 0, \\ \sigma_i\left(d_{th} - d_{ij}\right), & \text{if } 0 < d_{ij} < d_{th}, \\ 0, & \text{if } d_{th} \le d_{ij} \le R_{\text{com}}, \end{cases} \tag{6}$$

where d_{ij} is the Euclidean distance between nodes N_i and N_j; d_{th} is the radius within which the agents can be considered to be neighbors; σ_i is the expected degree of node N_i (a function of the average degree of nodes on the network and the number of N_i neighbors); R_{com} is the (maximum) communication radius of the node.

For the node N_i with k neighbors, the fitness value is defined as follows:

$$\text{minimize: } \sum_{j=1}^{k} F_{ij} = \sum_{j=1}^{k} \sigma_i\left(d_{th} - d_{ij}\right), \quad \text{for } 0 < d_{ij} \le d_{th},$$
$$\text{subject to: } d_{\text{mov}} \le d_{\max}, \tag{7}$$

where d_{mov} is the distance covered by the node based on the information in the chromosome; $d_{\max}$ is the maximum distance allowed.

The system's evolution is modeled by a nonhomogeneous Markov chain defined by its transition matrix (x_{kl}). Its elements are the probabilities of the node's transition from state k to state l at each step of the Markov chain; $k, l = 1, 2, \ldots, n$..

This approach demonstrates the evolution of a self-organizing system, which leads to an almost uniform distribution of nodes (agents) over the area.

To simulate dynamic interactions between Web services and solve some of the problems associated with their composition and adaptation, a stigma-based approach (a mechanism of spontaneous indirect interactions between individuals, when they leave tags in the operation domain to stimulate further activity of other individuals) is proposed in Ref. [178]. The proposed approach treats Web services and resources as sets of agents. The stigma-based self-organization of agents is used to develop and adapt compositions of Web services.

A service agent is defined as a pair $p_i = <id, F>$, where *id*–is the agent's identifier and F is the pheromone store, $F = \{f_1, f_2, \ldots, f_n\}$. Each fragrance contains a scalar value that represents the trajectory of a particular service agent. The quantity (volume) of a pheromone released and/or accepted by an agent is denoted by Q_{f_i}. Based on the user requests, service agent p_i requests an abstract work process, $r_j = \{r_{j1}, r_{j2}, \ldots, r_{jn}\}$, where r_{ji} is a specific service or resource that must be compiled to perform a single subtask. Thus, the requesting agent sends search queries through a local subnet of the associated agents to generate r_j. The operations of aggregation, evaporation, and storage of pheromones are introduced. Equations that determine the strength of smell f_i for agent p_i at time t and the sum of the amounts of pheromone f_i transferred to agent p_i at time t are given.

This approach enables self-organization even if one deals with incomplete (local) information and dynamic factors in a decentralized environment.

So-Grid, a set of bio-inspired algorithms adapted to the decentralized construction of the Grid information system, which is both adaptive and capable of self-organization, is presented in Ref. [179]. Such algorithms use the properties of swarm systems, where a number of entities (agents) perform simple operations on the local level, but together they produce an expanded form of swarm intelligence on the global level.

In particular, So-Grid provides two main functions: a logical rearrangement of resources, similar to the behavior of some ant and termite species, which move and collect objects from their environment, and a resource search similar to the mechanisms ants use to search for food.

In the Grid environment, a number of agents move autonomously like ants across the Grid through P2P connections and use skewed probability functions to: (i) replicate the resource descriptors, thereby facilitating resource discovery; (ii) collect similar resource descriptors in adjacent Grid hosts; (iii) promote distribution of the descriptors corresponding to fresh (recently updated) resources, as well as resources with high Quality of Service (QoS) characteristics.

The proposed approach is characterized by self-organization, scalability, and adaptability, making it useful for dynamic and partially unreliable distributed systems.

The So-Grid replication algorithm can reduce the system's entropy and efficiently distribute the content (information). Moreover, since the descriptors get gradually reorganized and replicated, the So-Grid discovery algorithm allows users to more quickly reach the Grid hosts that store information about more useful resources.

The model [180] based on the "honeybee" agents enables the discovery of cloud services on several levels.

The implementation of the proposed approach involves the following three phases:

(i) the development of a search model for a "hired collector bee" agent to discover private and dedicated cloud resources;

(ii) the deployment of a "scout bee" agent based on the mechanism of detecting unknown cloud sources;

(iii) the knowledge incurred through "waggle dancing mechanism." The onlooker bee agents will find unknown cloud sources, and the collected resources are clustered by resource cluster methods.

The "honeybee" agents are cooperative and can be efficiently used for search automation and grouping of cloud services.

The proposed solution greatly simplifies the search process: instead of spending a lot of time and effort discovering, evaluating, and exploring cloud clusters, users can easily discover, select, and use the services they need. In addition, cloud service providers (CSPs) can publish targeted information about their services.

3.5.3. Continuous Models Using Analogies from the Population Dynamics. The design of a synergistic system for managing self-organizing virtual communities, in particular, social networking services (SNS), enables the transition from chaos to a controlled process, thus achieving predictable results for the interaction between their agents. Several models that confirm this behavior of the systems are presented in Ref. [181].

A model of SNS can be expressed as follows:

$$\begin{cases} \dfrac{\mathrm{d}x_i(t)}{\mathrm{d}t} = f_i\left(x_1(t), \ldots, x_\lambda(t), y_{\lambda+1}(t), \ldots, y_\mu(t)\right), \\ \dfrac{\mathrm{d}y_j(t)}{\mathrm{d}t} = f_i\left(x_1(t), \ldots, x_\lambda(t), y_{\lambda+1}(t), \ldots, y_\mu(t), u_1(t), \ldots, u_\gamma(t)\right), \\ x_i(t_0) = x_i^0, \quad y_j(t_0) = y_j^0, \end{cases} \tag{8}$$

where $x_i(t)$, $y_j(t)$ are the indicators of agent interaction, $i = 1, 2, \ldots, \lambda$, $j = \lambda + 1, \lambda + 2, \ldots, \mu$;

$u_1(t)$ $l = 1,2,..\gamma$ stands for SNS control feedback; $x_i(t_0) = x_i^0$, $y_j(t_0) = y_j^0$, are the initial conditions.

The macro variable $\psi_v(x_i, y_j) = 0$, $v = 1, 2, \ldots$, is the parameter that determines the dynamics of interaction between the SNS agents. It ensures that the system is self-organized.

$$\psi_v = \psi_k\left(x_1, y_1, \ldots, x_\lambda, y_\mu\right) + \psi_d\left(x_1, y_1, \ldots, x_\lambda, y_\mu\right), \tag{9}$$

where $\psi_k(x_1, y_1, \ldots, x_\lambda, y_\mu)$ is the conservative component of the system, $k = 1, 2, \ldots$; $\psi_d(x_1, y_1, \ldots, x_\lambda, y_\mu)$ is the dissipative component in the system, $d = 1, 2, \ldots$.

The designed system state in the phase space shall be a point called a synergy splash point. However, its trajectory in the phase space must correspond to the following equation:

$$T_v \frac{d\psi_v(t)}{dt} + \psi_v(t) = 0, \tag{10}$$

where T_v is the period of all transitions initiated *by the SNS Actors'Interaction Synergetic Control*.

To simulate a specific situation, the model was defined as follows:

$$\begin{cases} \dfrac{dx(t)}{dt} = ax - xy - bx^2; \\ \dfrac{dy(t)}{dt} = -cy + xy, \end{cases} \tag{11}$$

where $x(t)$ stands for the process describing SNS actors' market for the information attractive to the virtual community investigated; $y(t)$ describes the attractive information supply; a indicates SNS actors' demand shift rate for attractive information, should $a > 0$ demand increase, should $a < 0$ demand decrease; b indicates SNS actors' rivalry in response to the substantially identical information posted; c indicates supply shift rate for information attractive to SNS actors.

The system may be in a certain stable condition called an attractor, which, however, is negatively affected by a potential threat to social media subjects.

A model of a social network of microblogs is given in Ref. [182], where the network is considered as a point-dissipative system describing the behavior of the following parameters, where

$$x_3(t) = N^*(t) - N_0(t), \tag{12}$$

is the difference between the number of network users who are in the excited and in the base states at time t;

$$x_1(t) = T(t) - T_0, \tag{13}$$

is the deviation from the number of tweets T_0 in the state of stable equilibrium;

$$x_2(t) = R(t) - R_0, \tag{14}$$

is a similar expression for retweets.

The rates of change of the variables can be written down in a parametric form, yielding a system of equations similar to the population dynamics equations:

$$\begin{aligned} x_1(t) &= -\alpha x_1(t) + \beta x_2(t), \\ x_2(t) &= -\gamma x_1(t) + c x_2(t) x_3(t), \\ x_3(t) &= \varepsilon\left(I_0 - N(t)\right) - k x_1(t) k x_2(t). \end{aligned} \tag{15}$$

It is shown that the above system is reduced to the well-known Lorentz dynamic system:

$$\begin{aligned} \dot{x} &= \sigma(y - x), \\ \dot{y} &= -y + xz, \\ \dot{z} &= b(\rho - z) - xy. \end{aligned} \tag{16}$$

This system has an attractor. The nonequilibrium dynamics of the system exhibits a singularity and a multiplexing character. A numerical experiment conducted on real-world data from the Twitter microblogging network showed that Twitter's dynamics is more often chaotic than regular.

Different levels of description of the dynamics for a large group of agents influenced by a small number of external agents are presented in Ref. [183]. The microscopic dynamics is described via classical flocking models supplemented by a metric on a set of agents and by a rule of their topological interaction. The mesoscopic description is obtained using the mean-field limit and is a hybrid model of a single kinetic equation for the system of agents and a system of ordinary differential equations that control the external agents. Finally, the macroscopic level is modeled using an appropriate system of equations of fluid dynamics. A similar approach to modeling of multi-agent social systems with the help of fluid dynamics analogies is also presented in Ref. [184].

An idea that self-organization in complex systems can be seen as a decision-making process similar to human decision-making and, vice versa, the decision-making process is nothing more than a kind of self-organization of the nervous system of those who make decisions, is presented in Ref. [162]. Similarities between these processes are also shown in a number of other works, for example, Ref. [163].

The mathematical formulation of the above approach is based on the probabilities of the system's states, i.e., the probabilities of its structures, models, evolutionary scenarios, etc. It is shown that the mathematical formalisms of self-organization and decision-making processes are identical. Thus, self-organization can be considered as an endogenous decision-making process, and, accordingly, the decision-making takes place through endogenous self-organization.

The proposed approach is illustrated by phase transitions in large statistical systems, evolutions and revolutions in social and biological systems, structural self-organization in dynamical systems, and probabilistic formulations in the classical and behavioral decision-making theories. In general, self-organization in these cases is described as a process of estimation of the probabilities of macroscopic states, or as a search for the most probable state. Application of the principle of minimal information is the standard way to obtain a probabilistic measure in classical systems. In other words, one needs to perform the conditional maximization of entropy under given constraints. Possible behavioral bias of decision-makers can be considered the same way that quantum fluctuations are treated in physical systems.

In applications to social systems, one uses a constraint called a systemic frustration or conflict, which is equivalent to the energy of a physical system. We can define an analogy of the free energy in social systems in the following way. The energy of system $E(\pi_j)$ in state j shall be called the state cost. The noise intensity T is an equivalent of the temperature of a social system, whereas the environment that generates the noise plays the role of a thermostat. The noise energy or noise cost for the system in state π_j is expressed by $TS(\pi_j)$, where $S(\pi_j)$ is the entropy of state π_j. Then, the equivalent of the free energy – the free cost – is the intrinsic state cost, i.e., the state cost less the noise cost:

$$F(\pi_j) = E(\pi_j) - TS(\pi_j). \tag{17}$$

Since the system is finite, the set of agents cannot exist in a single pure state, but rather it is characterized by probabilities of being in different states. Therefore, the probabilities obey the following distribution:

$$p(\pi_j) = \frac{1}{Z}\exp\{-\beta \cdot F(\pi_j)\}. \tag{18}$$

with the partition function

$$Z = \sum_j \exp\{-\beta \cdot F(\pi_j)\}, \tag{19}$$

where β is a Lagrange multiplier corresponding to the inverse temperature, $T = 1/\beta$.

Phase transitions occur between the dominant states, with continuous transitions corresponding to evolution, and discontinuous ones to revolutions or abrupt mode changes.

The above review highlights the following conclusions regarding the application of different model classes:

(i) The approach based on systems of ordinary differential equations is only meaningful when the system consists of a very large number of homogeneous actors (e.g., social networks). When the system consists of a relatively small number of actors with different roles and forms of interaction, discrete models are more appropriate.

(ii) In discrete models, the most difficult problem is to adequately describe the connections and interactions between actors. Sometimes such formalism is created ad hoc, requiring, as a rule, the development of original and sophisticated mathematical methods.

(iii) Another common approach involves the application of known biological models (the genetic algorithm or models for the behavior of colonies of insects such as bees or ants), where a well-developed mathematical formalism is available. However, it often requires the problem at hand to be adapted to fit the model, which can lead to the loss of important properties of the simulated system.

4. Conclusions

As shown in this review, self-organization is essential for the functioning of networked dynamical sociotechnical systems. The reasons are their high complexity and diversity, the complexity and increasing intelligence of their digital elements, and their distributed digital infrastructure. Moreover, self-organization becomes necessary when a system to be designed has to meet such criteria as high flexibility and process adaptability, active communication in the course of the processes of decision-making, operational tune-up, and system topology adjustments. Such tasks and objectives make centralized control inefficient and sometimes impossible under the imposed constraints.

In sociotechnical systems, a collaboration between humans and equipment is required even when the processes are digitized to a great extent. At the same time, two tendencies make those two groups of the systems' elements closer: the equipment becomes more intelligent, being developed toward modeling and reproduction of the human psyche. On the other hand, people are replaced by intelligent agents or digital twins that emulate human functionality and decision-making methods, and communicate on their behalf.

In this case, all the elements of the system become digital; however, as a rule, a human is still involved in monitoring, control, and decision-making. The development of big data technologies and communication networks makes it possible to bring social and technical entities closer together, thus enhancing the human ability to receive and analyze information, automate the decision-making processes, and interact with the equipment directly or remotely. As for the equipment, now it can learn to independently monitor, test, restore, and

configure the system processes. Thus, a conclusion made in Ref. [164] that self-organization can be considered as an endogenous decision-making process, and consequently, that decision-making takes place through endogenous self-organization, is clearly confirmed by practice. In the future, we can expect the scope of applications of the methods used to support decision-making in the models of digital distributed systems to expand even further.

The review focuses on two types of networked dynamical self-organizing sociotechnical systems, namely, manufacturing systems and social networks. In today's digital manufacturing systems, the hardware component dominates in most cases, while in social networks, despite their advanced digital infrastructure, the social component dominates, although a bot, an agent, or a digital twin might act on behalf of a particular person's account.

Despite a large number of available methods, the implementation of self-organizing systems, models, and cases does not have universal solutions for such common problems as distributed problem solving, coordination and approval, ethics and communications, and reliability and stability in critical states.

Human behavior brings uncertainty and unpredictability to the operation of digital systems, because people often do not act according to some preset rules or models. Hence, predictive modeling of human behavior cannot be reliable enough, although a fully developed digital infrastructure could limit any rule violations in the system to a considerable extent. Along with the advancement of research on human behavior, the main contributions toward a solution to those problems should come from further development of the models that replace humans in the processes, such as intelligent agents, digital twins, bots, etc. Such models can help reduce the risk of accidental errors as well as the risk of deliberate but unforeseen actions and decisions.

To provide communications, the system makes extensive use of multi-agent systems. Not only do they support human-to-machine and machine-to-machine interactions during cooperative activities, but they also ensure compliance with the system's rules, taking into account the interests of the system's elements they represent.

It is interesting to assess the level of self-organization in networked dynamical sociotechnical systems. Depending on the goals and objectives of the system, as well as its size, there may be different requirements for the level of its self-organization. Consequently, different implementations of self-organization could be chosen. One of the most difficult tasks is finding a balance between the development of rules and the maintenance of individual autonomy of the system's elements.

As shown in Ref. [130], self-organized systems do not always exhibit critical properties, hence, SOC or SOC-like and non-SOC processes are distinguished. For some self-organized systems, ensuring the system's stability of operation is one of the goals. The methods of prevention and mitigation of critical events are also of a great interest, as most digital industrial systems fall into this category.

Summarizing the review of models used in mathematical modeling of self-organizing digital distributed systems, it should be noted that the modeling of the micro level of systems based on the use of discrete models prevails.

This approach allows one to take into account more details of a real system, although it complicates the analysis of the model because of a very high dimensionality of the model's global state space [172]. In some cases, discrete models are created ad hoc [173–177]; in other cases, discrete models of digital distributed systems are based on analogies with appropriate biological models (the genetic algorithm or the behavioral models of social insects) [179, 180]. As for the macro-level continuous models of digital distributed systems, they significantly reduce the micro-level state space dimensionality by using equivalence classes. Normally, such models also use analogies, either from biology (population dynamics models) [182, 183] or fluid dynamics [184, 185].

Data Availability

All the data generated or analyzed during this study are included within this article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was supported by the Russian Foundation for Basic Research (Project 20-07-00651 A–Investigation into the stability of self-organizing distributed digital systems based on models of the social networks dynamics).

References

[1] Systems, S.-Organisation and Information, *An Interdisciplinary Perspective*, A. Pereira, A. W. Pickering, and R. Gudwin, Eds., Routledge, Taylor & Francis Group, Oxfordshire, UK, 2018.

[2] K. Wu and Q. Nan, "Information characteristics, processes, and mechanisms of self-organization evolution," *Complexity*, vol. 2019, Article ID 5603685, 2019.

[3] K. Han, G. Kokot, O. Tovkach, A. Glatz, I. S. Aranson, and A. Snezhko, "Emergence of self-organized multivortex states in flocks of active rollers," *Proceedings of the National Academy of Sciences*, vol. 117, no. 18, pp. 9706–9711, 2020.

[4] K.-H. N. Bui and J. J. Jung, "Internet of agents framework for connected vehicles: a case study on distributed traffic control system," *Journal of Parallel and Distributed Computing*, vol. 116, pp. 89–95, 2018.

[5] C. Savaglio, M. Ganzha, M. Bădică, M. Ivanović, and G. Fortino, "Agent-based Internet of things: state-of-the-art and research challenges," *Future Generation Computer Systems*, vol. 102, pp. 1038–1053, 2020.

[6] Springer Singapore, "Agents and multi-agent systems: technologies and applications," in *Proceedings of the 13th KES International Conference (KES-AMSTA-2019)*, G. Jezic, Y.-H. J. Chen-Burger, M. Kusek, R. Sperka, R. J. Howlett, and L. C. Jain, Eds., Vol. 148, Springer Singapore, Heidelberg, Germany, 2020.

[7] M. Wooldridge, *An Introduction to MultiAgent Systems*, John Wiley & Sons, Hoboken, NJ, USA, 2002.

[8] H. Moonen, *Multi-Agent Systems for Transportation Planning and Coordination, PhD Thesis*, Erasmus University Rotterdam, Rotterdam, Netherlands, 2009.

[9] B. Chen and H. H. Cheng, "A review of the applications of agent technology in traffic and transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 2, pp. 485–497, 2010.

[10] P. Skobelev, "Multi-agent systems for real-time adaptive resource management," in *Industrial Agents Industrial Agents: Emerging Applications of Software Agents in Industry*, P. Leitão and S. Karnouskos, Eds., pp. 207–229, Elsevier, Amsterdam, Netherlands, 2015.

[11] A. Bretas, A. Mendes, S. Chalup, M. Jackson, R. Clement, and C. Sanhueza, "Modelling railway traffic management through multi-agent systems and reinforcement learning," in *In Elsawah, S. (ed.) MODSIM2019, 23rd International Congress on Modelling and Simulation*, S. Elsawah, Ed., pp. 291–297, Modelling and Simulation Society of Australia and New Zealand, Canberra, Australia, December 2019, https://mssanz.org.au/modsim2019/B6/bretas.pdf.

[12] R. Dominguez and S. Cannella, "Insights on multi-agent systems applications for supply chain management," *Sustainability*, vol. 12, no. 5, p. 1935, 2020.

[13] Z. Lu, Z. Zhuang, Z. Huang, and W. Qin, "A framework of multi-agent based intelligent production logistics system," in *Proceedings of the 11th CIRP Conference on Industrial Product-Service Systems*, vol. 83, pp. 557–562, Amsterdam, Netherlands, January 2019.

[14] P. Oliveira, Z. Vale, H. Morais, T. Pinto, and I. Praça, "A multi-agent based approach for intelligent smart grid management," in *Proceedings of the 8th Power Plant and Power System Control Symposium (PPPSC'12 - IFAC)*, vol. 45, no. 21, pp. 109–114, Toulouse, France, September 2012.

[15] S. Howell, Y. Rezgui, J.-L. Hippolyte, B. Jayan, and H. Li, "Towards the next generation of smart grids: semantic and holonic multi-agent management of distributed energy resources," *Renewable and Sustainable Energy Reviews*, vol. 77, pp. 193–214, 2017.

[16] J. Zhong, W. Zhu, Y. Wu, and K. Wang, "Agent-based simulation of online trading," *Systems Engineering Procedia*, vol. 5, pp. 437–444, 2012.

[17] Springer, *Agent and Multi-Agent Systems in Distributed Systems - Digital Economy and E-Commerce*, A. Hakansson and R. Hartung, Eds., Springer, Berlin, Heidelberg, 2013.

[18] M. Amine Souissi, K. Bensaid, and R. Ellaia, "Multi-agent modeling and simulation of a stock market," *Investment Management and Financial Innovations*, vol. 15, no. 4, pp. 123–134, 2018.

[19] Springer International Publishing AG, "Agents and Multi-Agent Systems for Health Care," in *10th International Workshop A2HC 2017, and International Workshop A-HEALTH 2017 Revised and Extended Selected Papers*, S. Montagna, P. H. Abreu, S. Giroux, and M. I. Schumacher, Eds., Vol. 10685, Springer International Publishing AG, Berlin, Heidelberg, 2017.

[20] G. Aguado, V. Julian, A. Garcia-Fornes, and A. Espinosa, "A Multi-Agent System for guiding users in on-line social environments," *Engineering Applications of Artificial Intelligence*, vol. 94, Article ID 103740, 2020.

[21] V. Wiegel and J. van den Berg, "Combining moral theory, modal logic and MAS to create well-behaving artificial agents," *International Journal of Social Robotics*, vol. 1, no. 3, pp. 233–242, 2009.

[22] N. Cointe, G. Bonnet, and O. Boissier, "Ethics-based cooperation in multi-agent systems," in *Advances in Social Simulation*, H. Verhagen, M. Borit, G. Bravo, and N. Wijermans, Eds., Springer, Cham, Manhattan, NY, USA, 2020.

[23] M. Anderson and S. L. Anderson, "Machine ethics: creating an ethical intelligent agent," *AI Magazine*, vol. 28, no. 4, pp. 15–26, 2007.

[24] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Butterworth–Heinemann, Oxford, UK, Third Edition, 1980.

[25] P. Savaget, M. Geissdoerfer, A. Kharrazi, and S. Evans, "The theoretical foundations of sociotechnical systems change for sustainability: a systematic literature review," *Journal of Cleaner Production*, vol. 206, pp. 878–892, 2019.

[26] M. Ottens, M. Franssen, P. Kroes, and I. V. D. Poel, "Modelling infrastructures as socio-technical systems," *International Journal of Critical Infrastructures*, vol. 2, no. 2/3, pp. 133–145, 2006.

[27] D. Moher, L. Shamseer, M. Clarke et al., "Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement," *Systematic Reviews*, vol. 4, no. 1, p. 1, 2015.

[28] G. Zambrano, M. Carvalho, and D. Trentesaux, "Cooperation models between humans and artificial self-organizing systems: motivations, issues and perspectives," in *Proceedings of the 6th International Symposium on Resilient Control Systems (ISRCS)*, pp. 156–161, San Fransisco, CA, USA, August 2013.

[29] C. Gershenson, *Design and Control of Self-Organizing Systems, PhD Dissertation*, Vrijei Universiteit Brussele, Faculteit Wetenschappen, 2007.

[30] H. Haken, *Information and Self-Organization: A Macroscopic Approach to Complex Systems*, Springer-Verlag, Berlin Heidelberg, 2006.

[31] G. Di Marzo Serugendo, M.-P. Gleizes, and A. Karageorgos, "Self-organization in multi-agent systems," *The Knowledge Engineering Review*, vol. 20, no. 2, pp. 165–189, 2005.

[32] A. Dmitriev, V. Dmitriev, O. A. Tsukanova, and S. V. Maltseva, "A nonlinear dynamical approach to the interpretation of microblogging network complexity," in *Complex Networks & Their Applications VI (COMPLEX NETWORKS 2017)*, C. Cherifi, H. Cherifi, M. Karsai, and M. Musolesi, Eds., vol. 689, Manhattan, NY, USA, Springer Cham, 2018.

[33] N. Naikar, "Human-automation interaction in self-organizing sociotechnical systems," *Journal of Cognitive Engineering and Decision Making*, vol. 12, no. 1, pp. 62–66, 2018.

[34] I. Graessler, J. Hentze, and A. Poehler, "Self-organizing production systems: implications for product design," *Procedia CIRP*, vol. 79, pp. 546–550, 2019.

[35] J. Zhang, X. Yao, J. Zhou, J. Jiang, and X. Chen, "Self-organizing manufacturing: current status and prospect for Industry 4.0," in *Proceedings of the 5th International Conference on Enterprise Systems (ES)*, pp. 319–326, Beijing, China, September 2017.

[36] A. G. Chkhartishvili, D. A. Gubanov, and D. A. Novikov, "Social Networks: Models of Information Influence, Control and Confrontation," *Studies in Systems, Decision and Control*, Springer International Publishing AG, vol. 189, Berlin, Heidelberg, 2019.

[37] C. Kadushin, *Understanding Social Networks: Theories, Concepts, and Findings*, Oxford University Press, Oxford, UK, 2012.

[38] D. J. Watts, *Six Degrees: The Science of a Connected Age*, W. W. Norton & Company, Fifth Avenue, NY, USA, 2004.

[39] B. Brendan, *Designing Distributed Systems: Patterns and Paradigms for Scalable*, Reliable Services, O'Reilly Media, Sebastopol, CA, USA, 2018.

[40] Wiley, "Self-Organizing Networks: Self-Planning," in *Self-Optimization and Self-Healing for GSM, UMTS and LTE*, J. Ramiro and K. Hamied, Eds., Wiley, Hoboken, NJ, USA, 2011.

[41] F. Ducatelle, G. A. Di Caro, C. Pinciroli, and L. M. Gambardella, "Self-organized cooperation between robotic swarms," *Swarm Intelligence*, vol. 5, 2011.

[42] S. Murata and H. Kurokawa, *Self-Organizing Robots*, Springer, Tokyo, Japan, 2012.

[43] C. Hrabia, M. Lützenberger, and S. Albayrak, "Towards adaptive multi-robot systems: self-organization and self-adaptation," *The Knowledge Engineering Review*, vol. 33, p. E16, 2018.

[44] J. K. Gupta, M. Egorov, and M. Kochenderfer, "Cooperative multi-agent control using deep reinforcement learning," in *Autonomous Agents and Multiagent Systems, Lecture Notes in Computer Science*, G. Sukthankar and J. Rodriguez-Aguilar, Eds., vol. 10642, Manhattan, NY, USA, Springer Cham, 2017.

[45] V. Avetisov, A. Gorsky, S. Maslov, S. Nechaev, and O. Valba, "Phase transitions in social networks inspired by the Schelling model," *Physical Review E - Statistical, Nonlinear and Soft Matter Physics*, vol. 98, no. 3, pp. 032308-1–032308-11, 2018.

[46] W. Lin, M. Li, S. Zhou, J. Liu, G. Chen, and Z. Gu, "Phase transitions in normalized cut of social networks," *Physics Letters A*, vol. 383, no. 25, pp. 3037–3042, 2019.

[47] L. M. Floría, C. Gracia-Lázaro, J. Gómez-Gardeñes, and Y. Moreno, "Social network reciprocity as a phase transition in evolutionary cooperation," *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 79, Article ID 026106, 2009.

[48] P. A. Noël, C. D. Brummitt, and R. M. D'Souza, "Bottom-up model of self-organized criticality on networks," *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 89, Article ID 012807, 2014.

[49] B. Tadic, M. M. Dankulov, and R. Melnik, "Mechanisms of self-organized criticality in social processes of knowledge creation," *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, vol. 96, Article ID 032307, 2017.

[50] Y. Wang, H. Fan, W. Lin, Y.-C. Lai, and X. Wang, "Growth, collapse and self-organized criticality in complex networks," *Scientific Reports*, vol. 6, no. 1, Article ID 24445, 2016.

[51] Self-organising Software, *From Natural to Artificial Adaptation. Natural Computing Series*, M. Serugendo, G. Di, M.-P. Gleizes, and A. Karageorgos, Eds., Springer-Verlag, Berlin Heidelberg, 2011.

[52] W. R. Ashby, *Introduction to Cybernetics*Chapman & Hall, London, UK, 2015.

[53] Z. Zhou, H. Wang, and P. Lou, *Manufacturing Intelligence for Industrial Engineering: Methods for System Self-Organization, Learning, and Adaptation*, Engineering Science Reference, Hershey, PA, 2010.

[54] IGI Global, *Engineering Science Reference*, IGI Global, Hershey, PA, 2010.

[55] R. Dobrescu and G. Florea, "Unified framework for self-organizing manufacturing systems design," in *Service Orientation in Holonic and Multi Agent Manufacturing and Robotics*, T. Borangiu, A. Thomas, and D. Trentesaux, Eds., vol. 472, Berlin, Heidelberg, Springer, 2013.

[56] J. Leng, P. Jiang, C. Liu, and C. Wang, "Contextual self-organizing of manufacturing process for mass individualization: a cyber-physical-social system approach," *Enterprise Information Systems*, vol. 14, no. 8, pp. 1124–1149, 2020.

[57] H. V. D. Parunak and S. A. Brueckner, "Software engineering for self-organizing systems," *The Knowledge Engineering Review*, vol. 30, no. 4, pp. 419–434, 2015.

[58] K. Bellman, C. Landauer, N. Dutt et al., "Self-aware cyber-physical systems," *ACM Trans. Cyber-Phys. Syst*, vol. 4, 2020.

[59] S. Azaiez, M. Boc, L. Cudennec et al., "Towards flexibility in future industrial manufacturing: a global framework for self-organization of production cells," *Procedia Computer Science*, vol. 83, pp. 1268–1273, 2016.

[60] M. Chertow and J. Ehrenfeld, "Organizing self-organizing systems," *Journal of Industrial Ecology*, vol. 16, no. 1, pp. 13–27, 2012.

[61] A. Darabseh and N. M. Freris, "A software-defined architecture for control of IoT cyberphysical systems," *Cluster Computing*, vol. 22, no. 4, pp. 1107–1122, 2019.

[62] T. K. Kaiser and H. Hamann, "Self-Organization for Resilient Robot Self-Assembly with Minimal Surprise," *Robotics and Autonomous Systems*, vol. 122, Article ID 103293, 2019.

[63] I. A. Kalyaev, S. G. Kapustjan, and A. R. Gajduk, "Self-organizing distributed control systems of intellectual robot groups constructed on the basis of network model," *Upravlenie bol'shimi sistemami [Large-Scale Systems Control*, vol. 30, no. 1, pp. 605–639, 2010, in Russ.

[64] R. Lock and I. Sommerville, "Modelling and analysis of socio-technical system of systems," in *Proceedings of the 2010 15th IEEE International Conference on Engineering of Complex Computer Systems*, pp. 224–232, Oxford, UK, March 2010.

[65] B. Sauser and J. Boardman, "Taking hold of system of systems management," *Engineering Management Journal*, vol. 20, no. 4, pp. 3–8, 2008.

[66] "ISO/IEC/IEEE International Standard - systems and software engineering - system of systems (SoS) considerations in life cycle stages of a system," *ISO/IEC/IEEE 21839:2019*, pp. 1–40, 2019.

[67] N. Kubota, T. Fukuda, F. Arai, and K. Shimojima, "Genetic algorithm with age structure and its application to self-organizing manufacturing system," in *Proceedings of the IEEE Symposium on Emerging Technologies and Factory Automation*, pp. 472–477, Tokyo, Japan, November 1994.

[68] A. Koestler, *The Ghost in the Machine*, Macmillan, Macmillan, NY, USA, reprint ed. edition, 2016.

[69] M. Neuer, "Cognitive perception and self-organization for digital twins in cyber-physical steel production systems," in *Proceedings of the Industry 4.0 and Steelmaking Webinar of Steel Times International*, Future Steel Forum, Prague, Czech Republic, June 2020.

[70] N. Herakovič, H. Zupan, M. Pipan, J. Protner, and M. Šimic, "Distributed manufacturing systems with digital agents," *Journal of Mechanical Engineering*, vol. 65, pp. 650–657, 2019.

[71] A. Haghighi, "Combined strength of holons, agents and function blocks in cyber-physical systems," *Journal of Manufacturing Systems*, vol. 40, pp. 25–34, 2016.

[72] IEC 62264-1, "Enterprise-control system integration — Part 1: models and terminology," April 2013, https://www.iso.org/ru/standard/57308.html.

[73] M. Grieves, *Virtually Perfect: Driving Innovative and Lean Products through Product Lifecycle Management*, Space Coast Press, Brevard, FL, USA, 2011.

[74] M. Grieves, *Origins of the Digital Twin Concept*Working Paper, Florida Institute of Technology/NASA, FL, USA, 2016

[75] A. Tharumarajah, "Comparison of the bionic, fractal and holonic manufacturing system concepts," *International Journal of Computer Integrated Manufacturing*, vol. 9, no. 3, pp. 217–226, 1996.

[76] I. Gräßler and A. Pöhler, "Implementation of an adapted holonic production architecture," *Procedia CIRP*, vol. 63, pp. 138–143, 2017.

[77] S. Raileanu, "Production scheduling in a holonic manufacturing system using the open-control concept," *UPB Scientific Bulletin, Series C: Electrical Engineering*, vol. 72, no. 3, pp. 39–52, 2010.

[78] M. E. Peralta, M. Marcos, F. Aguayo, J. R. Lama, and A. Córdoba, "Sustainable fractal manufacturing: a new approach to sustainability in machining processes," *Procedia Engineering*, vol. 132, pp. 926–933, 2015.

[79] K. Ueda, "A concept for bionic manufacturing systems based on DNA-type information," *Human Aspects in Computer Integrated Manufacturing*, in *Proceedings of the IFIP 8th International PROLAMAT Conference*, pp. 853–863, Tokyo, Japan, June 1992.

[80] A. Tharumarajah, "From fractals and bionics to holonics," in *Agent-Based Manufacturing. Advanced Information Processing*, S. M. Deen, Ed., Springer, Berlin, Heidelberg, pp. 11–30, 2003.

[81] M. Segrest, *Modern Manufacturing (Volume 1): Best Practices from Industry Champions (Modern Manufacturing Case Studies)*Navigate Content Inc., Gulf Shores, AL, USA, 2019.

[82] L. Wang and A. Haghighi, "Combined strength of holons, agents and function blocks in cyber-physical systems," *Journal of Manufacturing Systems*, vol. 40, pp. 25–34, 2016.

[83] P. Leitão, "Agent-based distributed manufacturing control: a state-of-the-art survey," *Engineering Applications of Artificial Intelligence*, vol. 22, no. 7, pp. 979–991, 2009.

[84] E. Adam, T. Berger, Y. Sallez, and D. Trentesaux, "An open-control concept for a holonic multiagent system," in *Holonic and Multi-Agent Systems for Manufacturing (HoloMAS 2009)*, V. Mařík, T. Strasser, and A. Zoitl, Eds., vol. vol 5696, Berlin, Heidelberg, Springer, 2009.

[85] G. Di Marzo Serugendo, M.-P. Gleizes, and A. Karageorgos, "Self-organisation and emergence in MAS: an overview," *Informatica*, vol. 30, no. 1, pp. 45–54, 2006.

[86] D. Ye, M. Zhang, and A. V. Vasilakos, "A survey of self-organization mechanisms in multiagent systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 3, pp. 441–461, 2017.

[87] M. Shin and M. Jung, "MANPro: mobile agent-based negotiation process for distributed intelligent manufacturing," *International Journal of Production Research*, vol. 42, no. 2, pp. 303–320, 2004.

[88] E. Lee, "The past, present and future of cyber-physical systems: a focus on models," *Sensors*, vol. 15, no. 3, pp. 4837–4869, 2015.

[89] H. Ishii and B. Ullmer, "Tangible bits: towards seamless interfaces between people, bits, and atoms," in *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, pp. 234–241, Atlanta, Georgia, USA, March 1997.

[90] S. P. A. Datta, "Emergence of digital twins - is this the march of reason?" *Journal of Innovation Management*, vol. 5, no. 3, pp. 14–33, 2017.

[91] M. Hankel and B. Rexroth, "Industrie 4.0: The Reference Model Industrie 4.0 (RAMI 4.0)," ZVEI," April 2015, https://www.zvei.org/fileadmin/user_upload/Themen/Industrie_4.0/Das_Referenzarchitekturmodell_RAMI_4.0_und_die_Industrie_4.0-Komponente/pdf/ZVEI-Industrie-40-RAMI-40-English.pdf.

[92] Y. Wang, T. Towara, and R. Anderl, "Topological approach for mapping technologies in reference architectural model industrie 4.0 (RAMI 4.0)," in *Proceedings of the World Congress on Engineering and Computer Science (WCECS)*, vol. II, San Francisco, USA, October 2017.

[93] J. Lee, E. Lapira, B. Bagheri, and H.-A. Kao, "Recent advances and trends in predictive manufacturing systems in big data environment," *Manufacturing Letters*, vol. 1, no. 1, pp. 38–41, 2013.

[94] C. Dilmegani, "Top 20+ Quantum Computing Applications/Use Cases in 2021, AIMultiple," February 2021, https://research.aimultiple.com/quantum-computing-applications.

[95] *How Quantum Computing Will Transform These 9 Industries, CB Insights*, https://www.cbinsights.com/research/quantum-computing-industries-disrupted, 2021.

[96] R. Orus, S. Mugel, and E. Lizaso, "Quantum computing for finance: overview and prospects," *Reviews in Physics*, vol. 4, Article ID 100028, 2019.

[97] E. Haven and A. Khrennikov, *Quantum Social Science*, Cambridge University Press, Cambridge, NY, USA, 2013.

[98] E. Haven, A. Khrennikov, and T. R. Robinson, *Quantum Methods in Social Science: A First Course*, World Scientific Publishing Europe Ltd, London, UK, 2017.

[99] H. Wu, "Quantum mechanism helps agents combat "bad" social choice rules," *International Journal of Quantum Information*, vol. 9, no. 1, pp. 615–623, 2011.

[100] A. Bayerstadler, G. Becquin, B. Julia et al., "Industry quantum computing applications," *EPJ Quantum Technol*, vol. 8, no. 25, 2021.

[101] P. Bak, C. Tang, and K. Wiesenfeld, "Self-organized criticality: an explanation of the 1/fnoise," *Physical Review Letters*, vol. 59, no. 4, pp. 381–384, 1987.

[102] P. Bak, C. Tang, and K. Wiesenfeld, "Self-organized criticality," *Physical Review A*, vol. 38, no. 1, pp. 364–374, 1988.

[103] S. Corry and D. Perkinson, *Divisors and Sandpiles*, AMS, Providence, Rhode Island, 2018.

[104] D. Marković and C. Gros, "Power laws and self-organized criticality in theory and nature," *Physics Reports*, vol. 536, pp. 41–74, 2014.

[105] P. Bak, *How Nature Works: The Science of Self-Organized Criticality*, Copernicus, New York, NY, 1996.

[106] D. Stauffer and D. Sornette, "Self-organized percolation model for stock market fluctuations," *Physica A: Statistical Mechanics and Its Applications*, vol. 271, no. 3-4, pp. 496–506, 1999.

[107] K. Katahira, Y. Chen, and E. Akiyama, "Self-organized Speculation Game for the spontaneous emergence of financial stylized facts," *Physica A: Statistical Mechanics and Its Applications*, vol. 582, Article ID 126227, 2021.

[108] P. Bak and K. Sneppen, "Punctuated equilibrium and criticality in a simple model of evolution," *Physical Review Letters*, vol. 71, no. 24, pp. 4083–4086, 1993.

[109] G. F. R. Ellis and J. Kopel, "The dynamical emergence of biology from physics: branching causation via biomolecules," *Frontiers in Physiology*, vol. 9, 2019.

[110] M. A. Muñoz, "Colloquium: criticality and dynamical scaling in living systems," *Reviews of Modern Physics*, vol. 9, no. 90, Article ID 031001, 2018.

[111] A. Stožer, R. Markovič, J. Dolenšek et al., "Heterogeneity and delayed activation as hallmarks of self-organization and criticality in excitable tissue," *Frontiers in Physiology*, vol. 10, p. 869, 2019.

[112] Z. Olami, H. J. S. Feder, and K. Christensen, "Self-organized criticality in a continuous, nonconservative cellular automaton modeling earthquakes," *Physical Review Letters*, vol. 68, no. 8, pp. 1244–1247, 1992.

[113] G. G. Brunk, "Self-organized criticality: a new theory of political behaviour and some of its implications," *British Journal of Political Science*, vol. 31, no. 2, pp. 427–445, 2001.

[114] L. Peng, Y. Hou, L. Mengdi, and Z. Zhuo, "The sandpile model and empire dynamics," *Chaos, Solitons & Fractals*, vol. 143, Article ID 110615, 2021.

[115] D. L. Turcotte and J. B. Rundle, "Self-organized complexity in the physical, biological, and social sciences," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 1, pp. 2463–2465, 2002.

[116] L. I. Borodkin, "Methods of complexity science in political history studies," *International Trends*, vol. 1, pp. 4–16, 2005, In Russian.

[117] S. W. Fleming, "Scale-free networks, 1/f dynamics, and nonlinear conflict size scaling from an agent-based simulation model of societal-scale bilateral conflict and cooperation," *Physica A: Statistical Mechanics and Its Applications*, vol. 567, Article ID 125678, 2021.

[118] O. Kinouchi, R. Pazzini, and M. Copelli, "Mechanisms of self-organized quasicriticality in neuronal network models," *Frontiers in Physiology*, vol. 8, Article ID 583213, 2020.

[119] K. Mikkelsen, A. Imparato, and A. Torcini, "Emergence of slow collective oscillations in neural networks with spike-timing dependent plasticity," *Physical Review Letters*, vol. 110, no. 20, Article ID 208101, 2013.

[120] F. S. Matias, P. V. Carelli, C. R. Mirasso, and M. Copelli, "Self-organized near-zero-lag synchronization induced by spike-timing dependent plasticity in cortical populations," *PLoS One*, vol. 10, Article ID e0140504, 2015.

[121] M. Khoshkhou and A. Montakhab, "Spike-timing-dependent plasticity with axonal delay tunes networks of izhikevich neurons to the edge of synchronization transition with scale-free avalanches," *Frontiers in Systems Neuroscience*, vol. 13, p. 73, 2019.

[122] I. Y. Tyukin, D. Iudin, F. Iudin et al., "Simple model of complex dynamics of activity patterns in developing networks of neuronal cultures," *PLoS ONE*, vol. 14, no. 6, Article ID e0218304, 2019.

[123] M. Aguilera, X. E. Barandiaran, M. G. Bedia, and F. Seron, "Self-organized criticality, plasticity and sensorimotor coupling. Explorations with a neurorobotic model in a behavioural preference task," *PLoS One*, vol. 10, no. 2, Article ID e0117465, 2015.

[124] J. M. Torres-Rojo and R. Bahena-González, "Scale invariant behavior of cropping area losses," *Agricultural Systems*, vol. 165, pp. 33–43, 2018.

[125] H. J. S. Feder and J. Feder, "Self-organized criticality in a stick-slip process," *Physical Review Letters*, vol. 66, no. 20, pp. 2669–2672, 1991.

[126] B. Drossel and F. Schwabl, "Self-organized critical forest-fire model," *Physical Review Letters*, vol. 69, no. 11, pp. 1629–1632, 1992.

[127] C. L. Henley, "Self-organized percolation: a simpler model," *Bulletin of the American Physical Society*, vol. 34, p. 838, 1989.

[128] K. Christensen and Z. Olami, "Variation of the Gutenberg-Richterbvalues and nontrivial temporal correlations in a Spring-Block Model for earthquakes," *Journal of Geophysical Research*, vol. 97, no. B6, pp. 8729–8735, 1992.

[129] K. Sneppen, P. Bak, H. Flyvbjerg, and M. H. Jensen, "Evolution as a self-organized critical phenomenon," *Proceedings of the National Academy of Sciences*, vol. 92, no. 11, pp. 5209–5213, 1995.

[130] M. Aschwanden, "Theoretical Models of Self-Organized Criticality (SOC) Systems," 2012, https://arxiv.org/abs/1204.5119.

[131] D. L. Turcotte, "Self-organized criticality," *Reports on Progress in Physics*, vol. 62, no. 10, pp. 1377–1429, 1999.

[132] A. Mollgaard and J. Mathiesen, "Emergent user behavior on twitter modelled by a stochastic differential equation," *PLoS One*, vol. 10, no. 5, Article ID e0123876, 2015.

[133] B. Tadic, V. Gligorijevic, M. Mitrovic, and M. Suvakov, "Co-evolutionary mechanisms of emotional bursts in online social dynamics and networks," *Entropy*, vol. 15, pp. 5084–5120, 2013.

[134] B. Mandelbrot and R. L. Hudson, *The Misbehavior of Markets: A Fractal View of Financial Turbulence*, Basic Books, New York, NY, USA, 2004.

[135] L. Kirichenko, V. Bulakh, and T. Radivilova, "Fractal time series analysis of social network activities," in *Proceedings of the 2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T)*, pp. 456–459, Kharkov, Ukraine, October 2017.

[136] S. Tsugawa and H. Ohsaki, "Emergence of fractals in social networks: analysis of community structure and interaction locality," in *Proceedings of the 2014 IEEE 38th Annual Computer Software and Applications Conference*, pp. 568–575, Vasteras, Sweden, July 2014.

[137] M. Aguilera, I. Morer, X. Barandiaran, and M. Bedia, "Quantifying political self-organization in social media. Fractal patterns in the Spanish 15M movement on twitter," in *Proceedings of the ECAL 2013: The Twelfth European Conference on Artificial Life*, pp. 395–402, Sicily, Italy, September 2013.

[138] T. De Bie, J. Lijffijt, C. Mesnage, and R. Santos-Rodríguez, "Detecting trends in twitter time series," in *Proceedings of the 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*, pp. 1–6, Vietri sul Mare, Italy, September 2016.

[139] D. R. Bild, Y. Liu, R. P. Dick, Z. M. Mao, and D. S. Wallach, "Aggregate characterization of user behavior in twitter and analysis of the retweet graph," *Acm. T. Internet Techn.*vol. 15, no. 4, 2015.

[140] C. Remy, N. Pervin, F. Toriumi, and H. Takeda, "Information diffusion on twitter: everyone has its chance, but all chances are not equal," in *Proceedings of the 2013 International Conference on Signal-Image Technology & Internet-Based Systems*, pp. 483–490, Kyoto, Japan, December 2013.

[141] J. P. Gleeson and R. Durrett, "Temporal profiles of avalanches on networks," *Nature Communications*, vol. 8, p. 1227, 2017.

[142] S. Pramanik, Q. Wang, M. Danisch, J.-L. Guillaume, and B. Mitra, "Modeling cascade formation in Twitter amidst mentions and retweets," *Soc. Netw. Anal. Min.*vol. 7, 2017.

[143] M. M. Uddin, M. Imran, and H. Sajjad, "Understanding Types of Users on Twitter," June 2014, https://arxiv.org/pdf/1406.1335.pdf.

[144] A. Steyer and J.-B. Zimmermann, "Self organised criticality in economic and social networks," in *Economics with Heterogeneous Interacting Agents*, A. Kirman and J. B. Zimmermann, Eds., vol. 503, Berlin, Heidelberg, Springer, 2001.

[145] J. W. Kantelhardt, "Fractal and multifractal time series," in *Mathematics of Complexity and Dynamical Systems*, R. Meyers, Ed., Springer, Berlin, Heidelberg, 2012.

[146] B. Tadic, "Self-organised criticality and emergent hyperbolic networks: blueprint for complexity in social dynamics," *European Journal of Physics*, vol. 40, no. 2, Article ID 024002, 2019.

[147] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, "Critical phenomena in complex networks," *Reviews of Modern Physics*, vol. 80, no. 4, pp. 1275–1335, 2008.

[148] M. Prokopenko, "Design vs. Self-organization," in *Advances in Applied Self-Organizing Systems, Advanced Information and Knowledge Processing*, M. Prokopenko, Ed., Springer, London, UK, 2008.

[149] S. B. Cools, C. Gershenson, and B. D'Hooghe, "Self-organizing traffic lights: a realistic simulation," in *Advances in Applied Self-Organizing Systems, Advanced Information and Knowledge Processing*, M. Prokopenko, Ed., Springer, London, UK, 2008.

[150] G. Baldassarre, "Self-organization as phase transition in decentralized groups of robots: a study based on Boltzmann entropy," in *Advances in Applied Self-Organizing Systems, Advanced Information and Knowledge Processing*, M. Prokopenko, Ed., Springer, London, UK, 2008.

[151] A. V. Moere, "A model for self-organizing data visualization using decentralized multiagent systems," in *Advances in Applied Self-Organizing Systems, Advanced Information and Knowledge Processing*, M. Prokopenko, Ed., Springer, London, UK, 2008.

[152] T. Schlegel and R. Kowalczyk, "Self-organizing nomadic services in grids," in *Advances in Applied Self-Organizing Systems, Advanced Information and Knowledge Processing*, M. Prokopenko, Ed., Springer, London, UK, 2008.

[153] G. Mathews and H. Durrant-Whyte, "Decentralized decision making for ad-hoc multi-agent systems," in *Advances in Applied Self-Organizing Systems, Advanced Information and Knowledge Processing*, M. Prokopenko, Ed., Springer, London, UK, 2013.

[154] M. Rubenstein, A. Cornejo, and R. Nagpal, "Programmable self-assembly in a thousand-robot swarm," *Science*, vol. 345, no. 6198, pp. 795–799, 2014.

[155] U. Wilensky and W. Rand, *An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo*, The MIT Press, Cambridge, MA, USA, 2015.

[156] Z. Zuo, Q.-L. Han, B. Ning, X. Ge, and X.-M. Zhang, "An overview of recent advances in fixed-time cooperative control of multiagent systems," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 6, pp. 2322–2334, 2018.

[157] G. Vásárhelyi, C. Virágh, G. Somorjai, T. Nepusz, A. Eiben, and T. Vicsek, "Optimized flocking of autonomous drones in confined environments," *Science Robot*, vol. 3, no. 20, 2018.

[158] J. Florez-Lozano, F. Caraffini, C. Parra, and M. Gongora, "Training data set assessment for decision-making in a multiagent landmine detection platform," in *Proceedings of the 2020 IEEE Congress on Evolutionary Computation (CEC)*, Glasgow, UK, April 2020.

[159] H. Benbya, N. Nan, H. Tanriverdi, Y. Yoo, and W. McKelvey, "Complexity and IS research. Special issue," *MIS Quarterly: Management Information Systems*, vol. 44, no. 1, 2020.

[160] B. T. Pentland, P. Liu, W. Kremser, and T. Haerem, "The dynamics of drift in digitized processes," *MIS Quarterly*, vol. 44, no. 1, pp. 19–47, 2020.

[161] F. Battiston, G. Cencetti, I. Iacopini et al., "Networks beyond pairwise interactions: structure and dynamics," *Physics Reports*, vol. 874, pp. 1–92, 2020.

[162] V. I. Yukalov and D. Sornette, "Self-organization in complex systems as decision making," *Advances in Complex Systems*, vol. 17, Article ID 1450016, 2014.

[163] K. C. Laycraft, "Decision-making as a self-organizing process," *Ann. Cogn. Sci.*vol. 3, no. 1, pp. 86–99, 2019.

[164] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, The University of Illinois Press, Urbana, IL, USA, 1949.

[165] G. A. Pask and H. von Foerster, "A predictive model for self-organizing systems, Part I," *Cybernetica*, vol. 3, no. 4, pp. 258–300, 1961.

[166] H. von Foerster, "Action without Utility. An Immodest Proposal for the Cognitive Foundations of Behavior," *Cybernetics and Human Knowing*, vol. 10, pp. 27–50, 2003

[167] P. Busch, P. J. Lahti, and P. Mittelstaedt, "The quantum theory of measurement," *Lect. Notes in Physics Monographs*, Springer, vol. 2, Berlin, Heidelberg, 1996.

[168] M. Steiner and R. Rendell, "The quantum measurement problem," *Prog. on the Phys. of Quantum Measurement*, vol. 1, 2018.

[169] M. M. Wilde, *Quantum Information Theory*, Cambridge University Press, Cambridge, NY, USA, 2013.

[170] M. Hayashi, *A Group Theoretic Approach to Quantum Information*, Springer Cham, Manhattan, NY, USA, 2017.

[171] R. Holzer, P. Wüchner, and H. de Meer, "Modeling of self-organizing systems: an overview," *Electron. Commun. of the EASST*, vol. 27, 2010.

[172] C. Auer, P. Wüchner, and H. de Meer, "A method to derive local interaction strategies for improving cooperation in self-organizing systems," *Lecture Notes in Computer Science*, vol. 5343, pp. 171–180, 2008.

[173] Y. Li and Y. Jiang, "Self-organization based service discovery approach considering intermediary utility," in *Proceedings of the 2016 IEEE International Conference on Web Services (ICWS)*, pp. 308–315, IEEE, San Francisco, CA, USA, June 2016.

[174] E. del Val, M. Rebollo, and V. Botti, "Combination of self-organization mechanisms to enhance service discovery in open systems," *Information Sciences*, vol. 279, pp. 138–162, 2014.

[175] A. Ivaschenko, A. Khorina, V. Isayko, D. Krupin, V. Bolotsky, and P. Sitnikov, "Modeling of user behavior for social media analysis," in *Proceedings of the 2018 Moscow Workshop on Electronic and Networking Technologies (MWENT)*, pp. 1–4, IEEE, Moscow, Russia, March 2018.

[176] N. Rodrigues, P. Leitão, and E. Oliveira, "Dynamic composition of service oriented multi-agent system in self-organized environments," in *Proceedings of the 2014 Workshop on Intelligent Agents and Technologies for Socially Interconnected Systems (IAT4SIS '14)*, pp. 1–6, Association for Computing Machinery, New York, NY, USA, August 2014.

[177] C. Ş. Şahin, M. Ü. Uyar, S. Gundry, and E. Urrea, "Self organization for area coverage maximization and energy

conservation in mobile ad hoc networks," *Lecture Notes in Computer Science*, vol. 7050, pp. 49–73, 2012.

[178] A. Moustafa, M. Zhang, and Q. Bai, "Stigmergic modeling for web service composition and adaptation," *Lecture Notes in Computer Science*, vol. 7458, pp. 324–334, 2012.

[179] A. Forestiero, C. Mastroianni, and G. Spezzano, "So-Grid," *ACM Transactions on Autonomous and Adaptive Systems*, vol. 3, no. 2, pp. 1–37, 2008.

[180] M. N. Faruk, G. L. V. Prasad, and C. Nalini, "An Artificial Bee Colony (ABC) model for cloud service discovery and composition," *International Journal of Pure and Applied Mathematics*, vol. 117, no. 7, pp. 113–125, 2017.

[181] R. Hryshchuk and K. Molodetska, "Synergetic control of social networking services actors' interactions," in *Recent Advances in Systems, Control and Information Technology (SCIT 2016)*, R. Szewczyk and M. Kaliczyńska, Eds., vol. vol 543, Manhattan, NY, USA, Springer Cham, 2017.

[182] A. V. Dmitriev, O. A. Tsukanova, and S. V. Maltseva, "Investigation into the regular and chaotic states of microblogging networks as applied to social media monitoring," in *Proceedings of the 2016 IEEE 13th International Conference on E-Business Engineering (ICEBE)*, pp. 293–298, Macau, China, November 2016.

[183] G. Albi and L. Pareschi, "Modeling of self-organized systems interacting with a few individuals: from microscopic to macroscopic dynamics," *Applied Mathematics Letters*, vol. 26, no. 4, pp. 397–401, 2013.

[184] E. Tadmor, "Mathematical aspects of self-organized dynamics: consensus, emergence of leaders, and social hydrodynamics," *SIAM News*, vol. 48, no. 9, 2015.

[185] V. Dmitriev, A. Dmitriev, S. V. Maltseva, and S. Balybin, "Three-parameter kinetics of self-organized criticality on twitter," in *Complex Networks and Their Applications VIII (COMPLEX NETWORKS 2019)*, H. Cherifi, S. Gaito, J. Mendes, E. Moro, and L. Rocha, Eds., vol. vol 881, Manhattan, NY, USA, Springer Cham, 2020.

Hindawi
Complexity
Volume 2022, Article ID 6271232, 14 pages
https://doi.org/10.1155/2022/6271232

Research Article

Distributed Sliding Mode Fault-Tolerant LFC for Multiarea Interconnected Power Systems under Sensor Fault

Xichao Zhou,[1,2] Zhenlan Dou,[3] Weiming Zhang,[4] Yiwei Zhang,[5] Deyi Wang,[1] Chunyan Zhang,[3] and Dezhi Xu[4]

[1]*School of Electrical Engineering, Xi'an University of Technology, Xi'an 710043, China*
[2]*Science and Technology Research and Development Center, State Grid Integrated Energy Service Group Co., Ltd., Beijing 100053, China*
[3]*Technical Engineering Department, State Grid Shanghai Integrated Energy Service Co., Ltd., Shanghai 200023, China*
[4]*School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China*
[5]*State Grid Zhejiang Yuyao Power Supply Co., Ltd., Ningbo 315400, China*

Correspondence should be addressed to Weiming Zhang; wmzhang21@163.com

Received 10 September 2021; Revised 13 December 2021; Accepted 12 January 2022; Published 10 February 2022

Academic Editor: Yue Song

This paper focuses on the problem of load frequency control and sensor fault-tolerant control in the multiarea power grid. To solve these problems, a sliding mode control strategy based on an interval observer is designed. First, an interval observer is designed to obtain the boundary estimation information of the system state online, which is used to reconstruct the load disturbance online. Based on the reconstructed values, an integral sliding mode controller is designed to keep the system frequency stable when the load is disturbed. Then, sensor fault is diagnosed by interval residual. An augmented observer is designed to estimate sensor fault online by equivalent transformation, and the sensor fault is compensated by the fault estimation to reduce the influence of fault on system performance and ensure the reliable and stable operation of the power system. The simulation results show the superiority of the proposed control strategy.

1. Introduction

1.1. Literature Review. Because of the increasing size, interconnection, and complexity of power systems, today's power grids are composed of multiple highly interconnected areas, and various communication devices are used to protect, monitor, and control the power grid [1, 2]. The areas are connected via tie lines to eliminate any mismatching between generating and demand. Frequency, as an important index of power quality, should be stable within a certain range, so the active power balance should be guaranteed in the operation of power system [3]. However, because of the continuous growth of the installed capacity of power systems, the continuous access of new energy and distributed energy storage system have brought new challenges to maintaining system frequency stable [4, 5]. For example, the high permeability of renewable energy, such as wind power, can lead to power quality and stability problems in the power grid, which adversely affect the operation frequency of the grid [6, 7]. Thus, frequency regulation is becoming more prominent in the reconstituted power systems [8]. Besides, the change of the load will cause the corresponding deviation of the system frequency and the power of the link line, resulting in the corresponding change of the power generated [9]. At the same time, network security [10, 11] and sensor fault [12–14] cannot be ignored. When the sensor is faulty due to malicious attack signals, the fault will lead to a large frequency deviation. Even in the normal operation of communication network, sensors will inevitably be damaged or affected by interference signals, resulting in fault. The sensor faults can make a harmful effect on the power grid or even cause damage to the power grid during operating [15, 16]. Therefore, it is particularly important to design an effective LFC strategy with fault diagnosis and FTC scheme to reliably control the frequency of power system within an acceptable range.

1.2. Research Gap and Motivation. At present, the main method for improving the frequency stability of each area in power system is the load frequency control, which continuously adjusts the active output of the frequency-modulated generator set in the system to maintain the real-time active power balance between the generating side and the demand side [17]. However, with the increasing power demand and the increasing complexity of the system structure, the uncertainty of the system model is bound to increase, which reduces the accuracy and speed of load frequency control. In order to design a load frequency controller with good control performance, scholars at home and abroad have made great achievements in the physics-based dynamic model and data-driven methods, such as fuzzy logic control [18, 19], optimal control [20], model-free adaptive control [21], model predictive control [22, 23], robust control [24, 25], and other advanced control theories [26, 27]. Among them, a distributed optimal load frequency control scheme that can realize frequency regulation and economic scheduling was proposed in [20]. Because of the characteristics of random fluctuations in load disturbances, an LFC strategy based on a disturbance observer is proposed in [27]. However, the complex structure of the disturbance observer limits the practical application of these methods. The interval observer can determine the bound values of the state by assuming that the uncertainty in the system is unknown but bounded [28–30]. This advantage has been applied to state estimation [29] and fault detection [31]. However, unknown input reconstruction based on the interval estimation is still a problem to be solved.

In addition, in recent decades, research on system faults has been focused on fault detection and FTC and lots of significant fruits have been achieved [32–36]. For a class of nonlinear Markovian jump systems, an observer-based fault-tolerant control scheme is proposed to make the closed-loop system randomly stable [32]. The distributed fault-tolerant control for interconnected systems is proposed to ensure the overall stability and constraints of the system when faults occur [33]. In order to solve the problem of sensor fault in nonlinear systems with uncertain parameters, a fault estimation technique based on an adaptive sliding mode observer was investigated in [37]. In addition to the large and complicated power network of a power system, the power system is operated, protected, and controlled through an intricate communication system [38]. Therefore, reducing the impact of sensor failures on the system is a serious challenge.

Based on the above research, we propose an interval disturbance reconstruction technique to rapidly and accurately decouple and reconstruct the load disturbances in each region, so that the sliding mode load frequency control strategy based on interval disturbance reconstruction can effectively and quickly respond to the load changes and keep the system frequency deviation within a small range. Aiming at the sensor fault problem, we design a controller based on fault compensation to reduce the influence of the fault on the system and improve the system performance.

1.3. Contribution and Paper Organization. In this paper, load disturbance and sensor fault are considered to stabilize the system frequency. Considering load fluctuation reconstruction, parameter uncertainty, sensor fault diagnosis, fault estimation, and fault transition, this paper proposes a fault-tolerant load frequency control strategy based on interval disturbance reconstruction. The main contributions are summarized as follows:

(1) Interval observers are designed in each area of the LFC model of the multiarea interconnected power system, and the boundary value of system state quantity in each area is estimated rapidly and stably by using the designed observer.

(2) Using the estimation values of the bound values of the system state combined with the disturbance reconstruction technology, the system uncertainty items including load disturbance and uncertain parameters are reconstructed online.

(3) The interval residual errors based on the interval observer can be used for online monitoring and diagnosis of system sensor faults.

(4) The system sensor fault is regarded as a system state. An augmented observer is designed to estimate the sensor fault online, and the fault estimate is used for fault compensation. The sliding mode fault-tolerant load frequency control strategy is proposed to maintain the reliability and stability of the system under the problems of load disturbance and sensor fault.

The remainder of this research is organized as follows. The LFC model in the multiarea interconnected power grid and some preliminaries are introduced in Section 2. The interval observer-based disturbance reconstruction and the new-type load frequency controller are presented in Section 3. Sensor fault diagnosis, fault estimation, and the corresponding FTC strategy are given in Section 4. The simulation results are demonstrated in Section 5, and the conclusions are shown in Section 6.

2. Problem Formulation and Preliminaries

2.1. LFC System. The LFC problem for the multiarea power grid is considered in this paper. A three-area interconnected power grid is shown in Figure 1. Although the power grid has nonlinear and time varying characteristics, under stable operating conditions, the load variation range is small. It is possible to consider a linearized model near the steady-state operating point of system. A multiarea power grid usually consists of N LFC areas. For the sake of simplicity, use one equivalent generator to replace other generators in an area.

In each LFC subsystem, the frequency dynamic equation can be written as

$$\Delta \dot{f}_i = -\frac{D_i}{H_i}\Delta f_i + \frac{1}{H_i}\Delta P_{mi} - \frac{1}{H_i}\Delta P_{tie,i} - \frac{1}{H_i}\Delta P_{di}, \quad (1)$$

where Δf_i, ΔP_{di}, D_i, and H_i are the deviation of frequency, load disturbances, equivalent inertia constant, and equivalent damping coefficient; ΔP_{mi} and $\Delta P_{tie,i}$ are generator

output power and tie-line power, respectively, which can be described as

$$\begin{aligned}\Delta\dot{P}_{mi} &= -\frac{1}{T_{ti}}\Delta P_{mi} + \frac{1}{T_{ti}}\Delta P_{gi},\\ \Delta\dot{P}_{tie,i} &= 2\pi\sum_{j\neq i}^{N} T_{ij}\left(\Delta f_i - \Delta f_j\right),\end{aligned} \tag{2}$$

where T_{ti} is the time constants of turbine, T_{ij} is the synchronizing power coefficient in connection areas, N means there are N control areas, and ΔP_{gi} is governor valve position, which is determined by Δf_i and power generation set points ΔP_{ci} and defined as

$$\Delta\dot{P}_{gi} = -\frac{1}{R_i T_{gi}}\Delta f - \frac{1}{T_{gi}}\Delta P_{gi} - \frac{1}{T_{gi}}u_i, \tag{3}$$

where R_i and T_{gi} are the droop coefficient and time constants of governor and u_i is a power command signal generated by the automatic generation control (AGC) system to achieve the balance between the generating side and the demand side.

Combining (1)–(4), the state-space equation of i-th area can be described as

$$\begin{aligned}\dot{x}_i(t) &= A_{ii}x_i(t) + B_i u_i(t) + \sum_{j\neq i}^{N} A_{ij}x_j(t) + F_i\Delta P_{di}(t),\\ y_i(t) &= C_i x_i(t) + f_{si}(t),\end{aligned} \tag{4}$$

where $x_i(t) \triangleq \left[\begin{array}{cccc}\Delta f_i & \Delta P_{mi} & \Delta P_{gi} & \Delta P_{\text{tie},i}\end{array}\right]^T$,

$$A_{ii} \triangleq \begin{bmatrix} -\frac{D_i}{H_i} & \frac{1}{H_i} & 0 & -\frac{1}{H_i} \\ 0 & -\frac{1}{T_{ti}} & \frac{1}{T_{ti}} & 0 \\ -\frac{1}{R_i T_{gi}} & 0 & -\frac{1}{T_{gi}} & 0 \\ 2\pi\sum_{j\neq i}^{N} T_{ij} & 0 & 0 & 0 \end{bmatrix},$$

$$A_{ij} \triangleq \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -2\pi T_{ij} & 0 & 0 & 0 \end{bmatrix},$$

$$C_i \triangleq \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$B_i \triangleq \left[\begin{array}{cccc} 0 & 0 & -\frac{1}{T_{gi}} & 0 \end{array}\right]^T,$$

$$F_i \triangleq \left[\begin{array}{cccc} -\frac{1}{H_i} & 0 & 0 & 0 \end{array}\right]^T. \tag{5}$$

Also, $x_i(t) \in \mathbb{R}^n$, $u_i(t) \in \mathbb{R}^m$, $y_i(t) \in \mathbb{R}^p$, and $f_{si}(t) \in \mathbb{R}^p$.

The actual power system is large and complex. Considering the system parameter uncertainties, the system model can be expressed as

$$\begin{aligned}\dot{x}_i(t) &= \left(A_{ii} + \Delta A_{ii}\right)x_i(t) + \left(B_i + \Delta B\right)_i u_i(t)\\ &\quad + \sum_{j\neq i}^{N}\left(\left(A_{ij} + \Delta A_{ij}\right)x_j(t) + \left(F_i + \Delta F_i\right)\Delta P_{di}(t),\right.\end{aligned} \tag{6}$$

where $\Delta A_{ii}, \Delta B_i, \Delta A_{ij}, \Delta F_{1i}$, and ΔF_{2i} denote the parameters of uncertainties.

Define

$$\begin{aligned}\mathrm{d}_i(t) &= \Delta A_{ii}x_i(t) + \Delta B_i u_i(t) + \sum_{j\neq i}^{N}\Delta A_{ij}x_j(t)\\ &\quad + \left(F_i + \Delta F\right)_i\Delta P_{di}(t).\end{aligned} \tag{7}$$

Then, the model in (6) can be rewritten as

$$\begin{aligned}\dot{x}_i(t) &= A_{ii}x_i(t) + B_i u_i(t) + \sum_{j\neq i}^{N} A_{ij}x_j(t) + d_i(t),\\ y_i(t) &= C_i x_i(t) + f_{si}(t).\end{aligned} \tag{8}$$

Definition 1 (see [29]). A matrix is called Metzler matrix with nonnegative off-diagonal elements. A matrix is defined as a Hurwitz matrix if all eigenvalues have strictly negative real parts.

Lemma 1 (see [39]). *For a system of the form $\dot{x}(t) = Ax(t) + \omega(t)$, where matrix A is Metzler, it has $x(t) \geq 0$ for all $t > 0$ with $x(0) > 0, \omega(t) > 0$.*

Lemma 2 (see [40]). *For the transfer function of the form $\mathscr{G}(s) = \mathscr{C}(sI - \mathscr{A})^{-1}\mathscr{B} + \mathscr{D}$, we can get the following two equivalent conditions:*

(1) $\|\mathscr{C}(sI - \mathscr{A})^{-1}\mathscr{B} + \mathscr{D}\|_\infty < \rho$ with $\mathscr{A}$ stable.

(2) There exists a positive matrix P satisfying

$$\begin{bmatrix} P\mathscr{A} + \mathscr{A}^T P & P\mathscr{B} & \mathscr{C}^T \\ * & -\rho I & \mathscr{D}^T \\ * & * & -\rho I \end{bmatrix} < 0. \tag{9}$$

Assumption 1. The system state vector $x_i(t)$ is available for measurement.

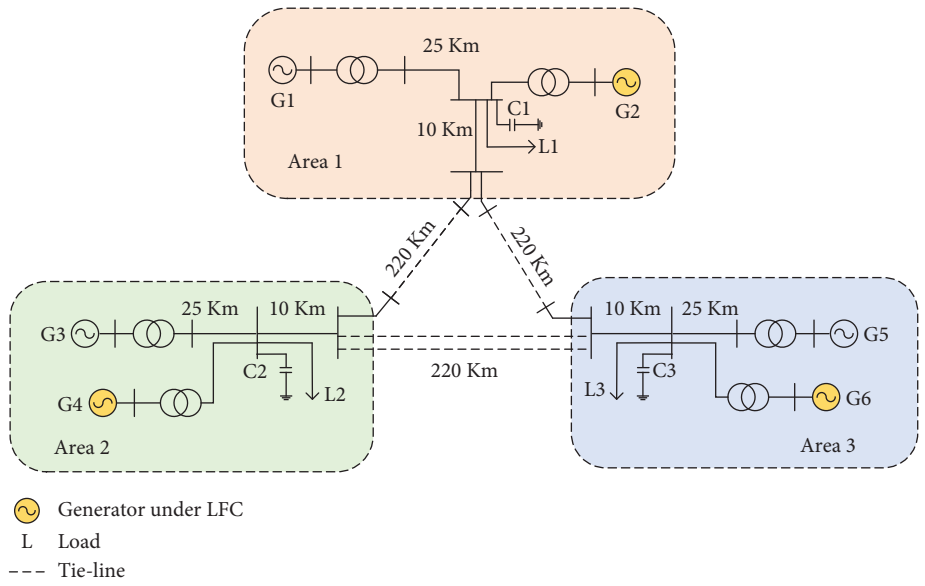


Figure 1: Three-area power system.

Assumption 2. There exist two known vectors $d_i^+ > 0$ and $d_i^- < 0$ such that $d_i^- \le d_i(t) \le d_i^+$.

3. Interval Observer-Based Disturbance Reconstruction and Sliding Mode LFC (DR-SMC-LFC)

In this section, the load disturbance in the power system will be reconstructed with the designed interval observer. To make the perfect transient and steady-state performance under the load disturbance, an integral-type sliding mode controller will be developed in this section by using the reconstructed value.

3.1. Interval Observer Design. Design an interval observer of the following form:

$$\begin{aligned}\dot{\hat{x}}_i^+(t) &= A_{ii}y_i(t) + B_iu_i(t) + \sum_{j\neq i}^{N} A_{ij}y_j(t) \\ &\quad + d_i^+ - d_i^- + L_i\left(y_i(t) - C_i\hat{x}_i^+(t)\right), \\ \dot{\hat{x}}_i^-(t) &= A_{ii}y_i(t) + B_iu_i(t) + \sum_{j\neq i}^{N} A_{ij}y_j(t) \\ &\quad + d_i^- - d_i^+ + L_i\left(y_i(t) - C_i\hat{x}_i^-(t)\right),\end{aligned} \tag{10}$$

where $\hat{x}_i^+(t)$ and $\hat{x}_i^-(t)$ are the boundary estimations of $x_i(t)$, respectively, and L_i is the gain of interval observer which needs to be defined in advance.

Theorem 1. *Supposing that Assumptions 1 and 2 are held and a system without sensor faults, if the initial states of the proposed interval observer (10) meet the following inequalities,*

$$\hat{x}_i^-(0) \le x_i(0) \le \hat{x}_i^+(0), \tag{11}$$

then (10) is an interval observer for system (8) with $-L_i$ designed as Metzler and Hurwitz.

Proof. Denote estimation errors of the interval observer as $e_{1i}(t) = \hat{x}_i^+(t) - x_i(t)$ and $e_{2i}(t) = x_i(t) - \hat{x}_i^-(t)$. If the system runs normally without any sensor fault, from (8) and (10), the error dynamics system can be obtained as follows:

$$\begin{aligned}\dot{e}_{1i}(t) &= A_{ii}(C_i - I)x_i(t) + \sum_{j\neq i}^{N} A_{ij}(C_i - I)x_j(t) \\ &\quad + d_i^+ - d_i^- - d_i(t) - L_iC_ie_{1i}(t), \\ \dot{e}_{2i}(t) &= A_{ii}(C_i - I)x_i(t) + \sum_{j\neq i}^{N} A_{ij}(C_i - I)x_j(t) \\ &\quad - d_i^- + d_i^+ + d_i(t) - L_iC_ie_{2i}(t).\end{aligned} \tag{12}$$

Recalling $C_i = I_n$,

$$\begin{aligned}\dot{e}_1(t) &= d_i^+ - d_i^- - d_i(t) - L_ie_1(t), \\ \dot{e}_2(t) &= -d_i^- + d_i^+ + d_i(t) - L_ie_2(t).\end{aligned} \tag{13}$$

According to Assumption 2, it can be obtained that

$$\begin{aligned}d_i^+ - d_i^- - d_i(t) &\ge 0, \\ -d_i^- + d_i^+ + d_i(t) &\ge 0.\end{aligned} \tag{14}$$

Choosing $-L_i$ as a Metzler and Hurwitz matrix, we can obtain that $\hat{x}_i^-(t) \le x_i(t) \le \hat{x}_i^+(t)$, for $t \ge 0$ from Lemma 1, and the condition of Hurwitz ensures the system is stable when operating. □

3.2. Disturbance Reconstruction. From Theorem 1, we can get $\hat{x}_i^-(t) \le x_i(t) \le \hat{x}_i^+(t)$, for all $t \ge 0$, so it is reasonable to obtain the following equation:

$$x_i(t) = \Phi_i(t)\hat{x}_i^+(t) + \left(I - \Phi_i(t)\right)\hat{x}_i^-(t), \tag{15}$$

where $\Phi_i(t) = \mathrm{diag}\{a_1(t), a_2(t), \ldots, a_n(t)\}$ and $0 \le a_i(t) \le 1, i = 1, 2, \ldots, n$.

After obtaining the interval observer according to Theorem 1, the following theorem is proposed in combination with (15), which is used to reconstruct the load disturbance.

Theorem 2. *For an LFC model (8) satisfying Assumptions 1 and 2 and an interval observer satisfying Theorem 1, the external load disturbance can be reconstructed by the following algebraic computation:*

$$\hat{d}_i(t) = \hat{M}_1\hat{x}_i^+ - \hat{M}_1\hat{x}_i^- + \hat{M}_2\left(d_i^+ - d_i^-\right), \tag{16}$$

where $\hat{M}_1$ and $\hat{M}_2$ are defined as

$$\begin{aligned}\hat{M}_1 &= \hat{\Phi}_i(t)\hat{N}_1 + \left(I - \hat{\Phi}_i(t)\right)\hat{N}_2 + \dot{\hat{\Phi}}_i(t) - A_{ii}\hat{\Phi}_i(t),\\ \hat{M}_2 &= 2\hat{\Phi}_i(t) - I,\end{aligned} \tag{17}$$

with $\hat{N}_1 = A_{ii}\hat{\Phi}_i(t) + L_i\hat{\Phi}_i(t) - L_i$ *and* $\hat{N}_2 = A_{ii}\hat{\Phi}_i(t) + L_i\hat{\Phi}_i(t)$.

Proof. According to (15) and (10), we can obtain

$$\begin{aligned}\dot{\hat{x}}_i^+(t) &= N_1\hat{x}_i^+ + \left(A_{ii} - N_1\right)\hat{x}_i^- + B_iu_i(t) + d_i^+ - d_i^-\\ &\quad + \sum_{j\neq i}^{N} A_{ij}\left(\Phi_j(t)\hat{x}_j^+(t) + I - \Phi_j(t)\right)\hat{x}_j^-(t),\\ \dot{\hat{x}}_i^-(t) &= N_2\hat{x}_i^+ + \left(A_{ii} - N_2\right)\hat{x}_i^- + B_iu_i(t) + d_i^- - d_i^+\\ &\quad + \sum_{j\neq i}^{N} A_{ij}\left(\Phi_j(t)\hat{x}_j^+(t) + I - \Phi_j(t)\right)\hat{x}_j^-(t),\end{aligned} \tag{18}$$

where $N_1 = A_{ii}\Phi_i(t) + L_i\Phi_i(t) - L_i$ and $N_2 = A_{ii}\Phi_i(t) + L_i\Phi_i(t)$.

According to (15), the first derivative of $x_i(t)$ is

$$\begin{aligned}\dot{x}_i(t) &= \left(\Phi_i(t)N_1 + \left(I - \Phi_i(t)\right)N_2 + \dot{\Phi}_i(t)\right)\hat{x}_i^+ + B_iu_i(t)\\ &\quad + \left(A_{ii} - \Phi_i(t)N_1 - \left(I - \Phi_i(t)\right)N_2 - \dot{\Phi}_i(t)\right)\hat{x}_i^-\\ &\quad + \sum_{j\neq i}^{N} A_{ij}\left(\Phi_j(t)\hat{x}_j^+(t) + \left(I - \Phi_j(t)\right)\hat{x}_j^-(t)\right)\\ &\quad + \left(2\Phi_i(t) - I\right)\left(d_i^+ - d_i^-\right).\end{aligned} \tag{19}$$

Substituting (8) and (15) into (19), we can obtain the reconstruction of $d_i(t)$:

$$d_i(t) = M_1\hat{x}_i^+ - M_1\hat{x}_i^- + M_2\left(d_i^+ - d_i^-\right), \tag{20}$$

where $M_1 = \Phi_i(t)N_1 + (I - \Phi_i(t))_2 + \dot{\Phi}_i(t) - A_{ii}\Phi_i(t)$ and $M_2 = 2\Phi_i(t) - I$.

Then, the external disturbance can be reconstructed by using (20). As for the computation of $\Phi_i(t)$, we can convert (15) to the following equality to calculate $\Phi_i(t)$:

$$\begin{aligned}x_i(t) &= \begin{bmatrix}\hat{x}_{i1}^+ & & & \\ & \hat{x}_{i2}^+ & & \\ & & \ddots & \\ & & & \hat{x}_{in}^+\end{bmatrix}\begin{bmatrix}a_1(t)\\ a_2(t)\\ \vdots\\ a_n(t)\end{bmatrix}\\ &\quad + \begin{bmatrix}\hat{x}_{i1}^- & & & \\ & \hat{x}_{i2}^- & & \\ & & \ddots & \\ & & & \hat{x}_{in}^-\end{bmatrix}\begin{bmatrix}1 - a_1(t)\\ 1 - a_2(t)\\ \vdots\\ 1 - a_n(t)\end{bmatrix}.\end{aligned} \tag{21}$$

Define

$$\begin{aligned}Q_1(t) &= \mathrm{diag}\{\hat{x}_{i1}^+, \hat{x}_{i2}^+, \ldots, \hat{x}_{in}^+\},\\ Q_2(t) &= \mathrm{diag}\{\hat{x}_{i1}^-, \hat{x}_{i2}^-, \ldots, \hat{x}_{in}^-\},\\ a_v(t) &= \left[a_1(t), a_2(t), \ldots, a_n(t)\right]^T.\end{aligned} \tag{22}$$

Then, we can obtain

$$\begin{aligned}\left(Q_1(t) - Q_2(t)\right)a_v(t) &= x_i(t) - \hat{x}_i^-(t),\\ a_v(t) &= \left(Q_1(t) - Q_2(t)\right)^{-1}\left(x_i(t) - \hat{x}_i^-(t)\right).\end{aligned} \tag{23}$$

Consequently, the disturbance can be reconstructed by using (16). □

3.3. Sliding Mode Load Frequency Controller Design. In order to maintain the system frequency stable when load disturbances occur, a new-type load frequency controller is designed and applied in this subsection which is used in the integral-type sliding mode technique. The integral-type switching surfaces can be designed as

$$s_i(t) = G_iy_i(t) - \int_0^t G_i\left(A_{ii} + B_iK_i\right)y_i(\theta)d\theta, \tag{24}$$

where G_i and K_i are constant matrices with appropriate dimension.

Traditional SMC utilizes the boundary value of load disturbance to design a control law. However, the uncertain term $d_i(t)$ always changes in actual engineering. The controller based on the boundary value ξ_i may not preform very well, so we redesigned the controller by using the reconstructed value $\hat{d}_i(t)$ and the adaptive law as follows:

$$\begin{cases} u_i(t) = K_i y_i(t) - \left((G_i B)_i \right)^{-1} \left[G_i \widehat{d}_i(t) + \kappa_i s_i(t) \right. \\ \qquad \left. + G_i \sum_{j \neq i}^{N} A_{ij} y_j(t) + \widehat{\tau}_i \| G_i \| \mathrm{sat}\left(s_i(t)\right) \right] \\ \mathrm{sat}\left(s_i(t)\right) = \begin{cases} \dfrac{s_i(t)}{\varepsilon_i}, \| s_i(t) \| \leq \varepsilon_i \\ \mathrm{sign}\left(s_i(t)\right), \| s_i(t) \| > \varepsilon_i \end{cases} \end{cases}, \tag{25}$$

where $\widehat{d}_i(t)$ is the reconstruction of $d_i(t)$ from (16) and $\| d_i(t) - \widehat{d}_i(t) \| < \tau_i$, in which τ_i is the positive constant, and we define the following adaptive law of $\widehat{\tau}_i$:

$$\dot{\widehat{\tau}}_i = n_i \| G_i \| \| s_i(t) \|. \tag{26}$$

Theorem 3. *The closed-loop control system will remain asymptotically stable under the load frequency controller (25) based on load disturbance reconstruction with adaptive law (26).*

Proof. Define a Lyapunov function as

$$V_i(t) \triangleq 0.5 s_i^2(t) + 0.5 \frac{1}{n_i} \widetilde{\tau}_i^2, \tag{27}$$

where $\widetilde{\tau}_i = \tau_i - \widehat{\tau}_i$ and $\dot{\widetilde{\tau}}_i = \dot{\tau}_i - \dot{\widehat{\tau}}_i = -\dot{\widehat{\tau}}_i$.

According to the switching surface (24), we can obtain

$$\begin{aligned} \dot{V}_i(t) &= s_i(t) \dot{s}_i(t) - \frac{1}{n_i} \widetilde{\tau}_i \dot{\widehat{\tau}}_i \\ &= s_i(t) \left[G_i C_i A_{ii} x_i(t) + G_i C_i B_i u_i(t) + G_i C_i d_i(t) \right. \\ &\quad \left. + G_i C_i \sum_{j \neq i}^{N} A_{ij} x_j(t) - G_i A_{ii} C_i x_i(t) \right. \\ &\quad \left. - G_i B_i K_i C_i x_i(t) \right] - \frac{1}{n_i} \widetilde{\tau}_i \dot{\widehat{\tau}}_i. \end{aligned} \tag{28}$$

Recalling $C_i = I_n$,

$$\begin{aligned} \dot{V}_i(t) &= s_i(t) \left[G_i B_i u_i(t) + G_i \sum_{j \neq i}^{N} A_{ij} x_j(t) + G_i d_i(t) \right. \\ &\quad \left. - G_i B_i K_i x_i(t) \right] - \frac{1}{n_i} \widetilde{\tau}_i \dot{\widehat{\tau}}_i \\ &= s_i(t) \left[G_i \left(d_i(t) - \widehat{d}_i(t) \right) - \kappa_i s_i(t) \right. \\ &\quad \left. - \widehat{\tau}_i \| G_i \| \mathrm{sat}\left(s_i(t)\right) \right] - \frac{1}{n_i} \widetilde{\tau}_i \dot{\widehat{\tau}}_i \\ &\leq \| G_i \| \| s_i(t) \| \tau_i - \kappa_i \| s_i(t) \|^2 - \| G_i \| \| s_i(t) \| \widehat{\tau}_i \\ &\quad - \frac{1}{n_i} \widetilde{\tau}_i \dot{\widehat{\tau}}_i \\ &= \| G_i \| \| s_i(t) \| \left(\tau_i - \widehat{\tau}_i - \widetilde{\tau}_i \right) - \kappa_i \| s_i(t) \|^2 \\ &= -\kappa_i \| s_i(t) \|^2 < 0. \end{aligned} \tag{29}$$

Consequently, it can be concluded that the system state will eventually remain stable under the proposed controller, which implies the new-type LFC controller can ensure the stability of each control area in power system. □

4. Sensors Fault Diagnosis and Estimation and FTC Strategy

The closed-loop control of the system requires precise measurements. However, any fault that may affect the sensor

will inevitably lead to performance degradation and cause system damage. Therefore, to ensure the stability and reliability when the system is operating, it is necessary to perform fault detection and fault tolerance control when faults occur. In this section, the fault diagnosis based on interval residual errors, sensor fault estimation based on the augmented observer, and FTC strategy are defined, and the whole control strategy is redesigned as shown in Figure 2.

4.1. Fault Diagnosis. According to the interval observer designed in Section 3, the following interval residual errors are designed for sensor fault diagnosis:

$$\begin{aligned} r_i^+(t) &= y_i(t) - \widehat{y}_i^-(t), \\ r_i^-(t) &= y_i(t) - \widehat{y}_i^+(t), \end{aligned} \tag{30}$$

where

$$\begin{aligned} \widehat{y}_i^+(t) &= C_i^+ \widehat{x}_i^+(t) - C_i^- \widehat{x}_i^-(t), \\ \widehat{y}_i^-(t) &= C_i^+ \widehat{x}_i^-(t) - C_i^- \widehat{x}_i^+(t). \end{aligned} \tag{31}$$

According to Theorem 1, if $-L_i$ is designed as a Metzler matrix, then under the initial condition $\widehat{x}_i^-(0) \le x_i(0) \le \widehat{x}_i^+(0)$ and without sensor faults, we can get

$$\begin{aligned} &\widehat{x}_i^-(t) \le x_i(t) \le \widehat{x}_i^+(t), \\ &\widehat{y}_i^-(t) \le y_i(t) \le \widehat{y}_i^+(t), \\ &0 \in [r_{is}^-(t), r_{is}^+(t)], s = 1, 2 \ldots, 4. \end{aligned} \tag{32}$$

In case of faults, even if $-L_i$ is designed as Metzler and Hurwitz, the nonnegativity of the error dynamics cannot be guaranteed. Therefore, $0 \in [r_{is}^-(t), r_{is}^+(t)]$ will not be guaranteed at this time, which can be used for detecting faults.

4.2. Fault Estimation. The following augmented system (33) is designed which is equivalent to the original dynamic system (8). This equivalent transformation treats the sensor failure $f_{si}(t)$ in the original system (8) as a state one, so that an effective state observer strategy can be designed, while estimating the states and sensor faults.

$$\begin{cases} \dot{\overline{x}}_i(t) = \overline{A}_{ii}\overline{x}_i(t) + \overline{B}_i u_i(t) + w_i(t) + \sum_{j \ne i}^{N} \overline{A}_{ij}\overline{x}_j(t) \\ y_i(t) = \overline{C}_i \overline{x}_i(t) \end{cases}, \tag{33}$$

where

$$\begin{aligned} &\overline{x}_i \triangleq \begin{bmatrix} x_i \\ f_{si} \end{bmatrix}, \overline{A}_{ii} \triangleq \begin{bmatrix} A_{ii} & 0 \\ 0 & 0_{p\times p} \end{bmatrix}, \overline{B}_i \triangleq \begin{bmatrix} B_i \\ 0_{p\times m} \end{bmatrix} \\ &\overline{C}_i \triangleq \begin{bmatrix} C_i & I_p \end{bmatrix}, \overline{A}_{ij} \triangleq \begin{bmatrix} A_{ij} & 0 \\ 0 & 0_{p\times p} \end{bmatrix}, w_i = \begin{bmatrix} d_i \\ \dot{f}_{si} \end{bmatrix}. \end{aligned} \tag{34}$$

For the augmented system (33), a state observer for i th subsystem is constructed as follows:

$$\begin{cases} \dot{\widehat{\overline{x}}}_i(t) = \overline{A}_{ii}\widehat{\overline{x}}_i(t) + \overline{B}_i u_i(t) + L_{pi}\left(y_i(t) - \widehat{y}_i(t)\right) \\ \quad + \sum_{j \ne i}^{N} \overline{A}_{ij}\widehat{\overline{x}}_j, \\ \widehat{y}_i(t) = \overline{C}_i \widehat{\overline{x}}_i(t), \end{cases} \tag{35}$$

where $\widehat{\overline{x}}_i(t)$ is the estimation of $\overline{x}_i(t)$ by the augmented observer and L_{pi} is a constant matrix with appropriate dimension.

Define the state estimate error and sensor fault estimate error as $e_x(t) = \widehat{x}_i(t) - x_i(t)$, $e_{fsi}(t) = \widehat{f}_{si}(t) - f_{si}(t)$, and $\overline{e}_i(t) = \left[e_x^T(t) \;\; e_{fsi}^T(t) \right]^T$. From (33) and (35), the error dynamic system is obtained as

$$\begin{cases} \dot{\overline{e}}_i(t) = \left(\overline{A}_i - L_{pi}\overline{C}_i\right)\overline{e}_i(t) + \sum_{j \ne i}^{N} \overline{A}_{ij}\overline{e}_j - w_i(t) \\ e_{fsi}(t) = \overline{I}_p^T \overline{e}_i(t) \quad i = 1, 2, \ldots, N \end{cases}, \tag{36}$$

where $\overline{I}_p = \left[0 \;\; I_p \right]^T$.

Theorem 4. *Given a H_∞ performance level ρ, if there exist matrix $W_i \in \mathbb{R}^{(n+p)N \times pN} > 0$ and symmetric matrices $P_i \in \mathbb{R}^{(n+p)N \times (n+p)N} > 0$ such that condition (37) is satisfied,*

$$\begin{bmatrix} \Gamma_1 & -P & I_N \otimes \overline{I}_p \\ * & -\rho I_{(n+p)N} & 0 \\ * & * & -\rho I_{pN} \end{bmatrix} < 0, \tag{37}$$

where

$$\begin{aligned} \Gamma_1 &= P\overline{A} - W\overline{C} + (P\overline{A} - W\overline{C})^T, \\ P &= \mathrm{diag}\{P_1, P_2, \ldots, P_N\}, \\ W &= \mathrm{diag}\{W_1, W_2, \ldots, W_N\}, \\ \overline{C} &= \mathrm{diag}\{\overline{C}_1, \overline{C}_2, \ldots, \overline{C}_N\}, \\ \overline{A} &= \begin{bmatrix} \overline{A}_{11} & \overline{A}_{12} & \cdots & \overline{A}_{1N} \\ \overline{A}_{21} & \overline{A}_{21} & \cdots & \overline{A}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{A}_{N1} & \overline{A}_{N2} & \cdots & \overline{A}_{NN} \end{bmatrix}, \end{aligned} \tag{38}$$

then the error dynamics of the whole interconnected power system satisfies the H_∞ performance. Also, we can calculate the gain matrices L_p by $L_{pi} = P_i^{-1} W_i$.

Proof. Define augmented vectors and matrices as follows:

$$\begin{aligned} L_p &= \mathrm{diag}\left\{L_{p1}, L_{p2}, \ldots, L_{pN}\right\}, \\ \overline{e}(t) &= \left[\overline{e}_1^T(t) \;\; \overline{e}_2^T(t) \;\; \cdots \;\; \overline{e}_N^T(t) \right]^T, \\ e_{fs}(t) &= \left[e_{fs1}^T(t) \;\; e_{fs2}^T(t) \;\; \cdots \;\; e_{fsN}^T(t) \right]^T, \\ w(t) &= \left[w_1^T(t) \;\; w_2^T(t) \;\; \cdots \;\; w_N^T(t) \right]^T. \end{aligned} \tag{39}$$

So, we can get the global error dynamics of power grid as follows:

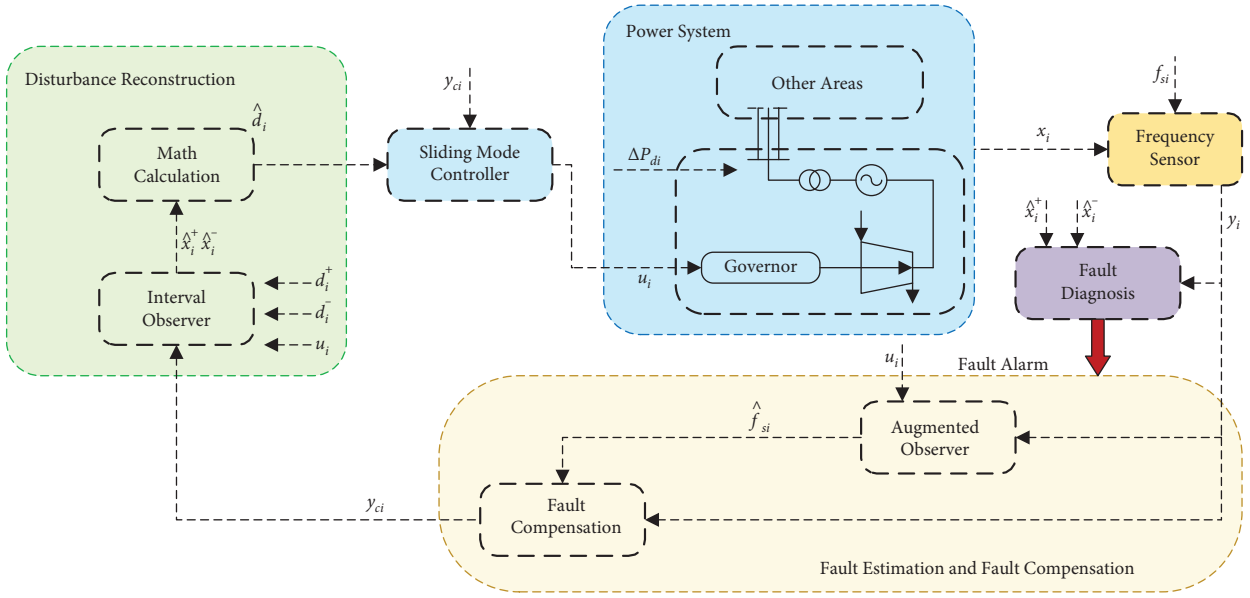


Figure 2: The structure of the proposed control approach.

$$\begin{cases} \dot{\bar{e}}(t) = \left(\overline{A} - L_p\overline{C}\right)\bar{e}(t) - w(t) \\ e_{fs}(t) = I_N \otimes \overline{I}_p^T \bar{e}(t) \end{cases} \tag{40}$$

From the global view, $W = PL_p$, so (37) can be rewritten as

$$\begin{bmatrix} \Gamma_2 & -P & I_N \otimes \overline{I}_p \\ * & -\rho I_{(n+p)N} & 0 \\ * & * & -\rho I_{pN} \end{bmatrix} < 0, \tag{41}$$

where

$$\Gamma_2 = P\overline{A} - PL_p\overline{C} + \left(P\overline{A} - PL_p\overline{C}\right)^T. \tag{42}$$

According to Lemma 2, if (37) is satisfied, the global error dynamics system (40) will satisfy the H_∞ performance $\| - (I_N \otimes \overline{I}_p^T)(sI - (\overline{A} - L_p\overline{C}))^{-1}\|_\infty < \rho$. □

Remark 1. The H_∞ performance in Theorem 4 can inhibit the influence of $w(t)$ on sensor failure estimation error $e_{fs}(t)$. The condition in (37) is in the form of LMI which can make the design of the gain matrix L_{pi} more convenient.

4.3. Fault Compensation. According to the augmented observer designed in the previous section, the online sensor failure estimations can be obtained by

$$\hat{f}_{si}(t) = \overline{I}_p^T \hat{\bar{x}}_i(t). \tag{43}$$

To keep the system stable when operating, the signals of system output are compensated accordingly by using the sensor fault estimation information, when faults occur.

$$\begin{aligned} y_{ci}(t) &= y_i(t) - \hat{f}_{si}(t) \\ &= C_i x_i(t) + f_{si}(t) - \hat{f}_{si}(t) \\ &= C_i x_i(t) + e_{fsi}(t), \end{aligned} \tag{44}$$

where $y_{ci}(t)$ is the compensated output signal.

Remark 2. The H_∞ performance used in Theorem 4 can inhibit the influence of $w(t)$ on sensor failure estimation error $e_{fs}(t)$. We can obtain accurate sensor fault estimations as long as we choose appropriate H_∞ performance level γ. Therefore, the fault compensation part can use the sensor fault estimations, which greatly eliminate the interference of the sensor faults on the system performance and guarantee the reliability and stability when the system is running.

4.4. Fault-Tolerant Control. Once the sensor faults occur, the compensation signal described in (44) is used to replace the output signal of the system measurement. Therefore, the

DR-SMC-LFC controller proposed in this paper can be redefined as

$$s_i(t) = G_i y_{ci}(t) - \int_0^t G_i\left(A_{ii} + B_i K_i\right) y_{ci}(\theta)\mathrm{d}\theta,$$

$$\begin{cases} u_i(t) = K_i y_{ci}(t) - \left(G_i B\right)_i^{-1}\left[G_i \widehat{d}_i(t) + \kappa_i s_i(t)\right. \\ \left. + G_i \sum_{j\neq i}^{N} A_{ij} y_{cj}(t) + \widehat{\tau}_i \left\|G_i\right\| \mathrm{sat}\left(s_i(t)\right)\right] \\ \mathrm{sat}\left(s_i(t)\right) = \begin{cases} \dfrac{s_i(t)}{\varepsilon_i}, \left\|s_i(t)\right\| \le \varepsilon_i \\ \mathrm{sign}\left(s_i(t)\right), \left\|s_i(t)\right\| > \varepsilon_i \end{cases} \end{cases} \tag{45}$$

In this way, it can not only keep the system frequency stable when the load changes in the power system but also timely compensate the measurement signal when the faults occur and realize FTC.

5. Simulation Results

To investigate the performance of the proposed method in a multiarea LFC, a three-area interconnected power grid model was built on the MATLAB/Simulink environment. Two cases with or without sensor faults under load disturbance are simulated, and the simulation results are demonstrated in Figures 3–5. The system model parameters are standardized with 100 MW as the reference value, and the established system parameters [15] are shown in Table 1. The purpose of LFC is to maintain the frequency within the nominal range and minimize the connection power between different control areas. Therefore, $\Delta f_i(t)$ and $\Delta P_{tie,i}$ can be used to compare the quality of different control strategies.

5.1. Case 1. In this case, fully considering the interconnectivity of the multiarea power grid, different load disturbances are applied in the three areas, respectively.

In order to realize the estimation of the bound values of the state in each control area by the designed interval observer, take area 1 as an example and design the parameters and initial values of the interval observer (10) as follows:

$$\begin{aligned} &\widehat{x}_1^+(0) = [0.005; 0.005; 0.005; 0.005], \widehat{x}_1^-(0) = -\widehat{x}_1^+(0), \\ &d_1^+(0) = [0.01; 0.01; 0.01; 0.01], d_1^-(0) = -d_1^+(0), \\ &L_1 = \begin{bmatrix} 10 & -2 & -2 & -2 \\ -2 & 10 & -2 & -2 \\ -2 & -2 & 10 & -2 \\ -2 & -2 & -2 & 10 \end{bmatrix}. \end{aligned} \tag{46}$$

Taking the deviation of system frequency and tie-line power as an example, the real-time boundary estimations' effect of frequency deviation is shown in Figure 3(a) and the real-time boundary estimations' effect of tie-line power deviation is shown in Figure 3(b). From the simulation results, it can be seen that the estimations of the bound values of the state can respond rapidly from its initial state, quickly follow the change of the state in each control area, and ensure the state condition of $\widehat{x}_i^-(t) \le x_i(t) \le \widehat{x}_i^+(t)$.

By using the real-time boundary estimations obtained by using the interval observer combined with the interval disturbance reconstruction technology proposed in this paper, the load disturbances in the three areas are reconstructed in real time. Figure 3(c) shows the reconstruction effect of interval disturbances which is compared with actual load disturbances. The advantage of interval disturbance reconstruction is that it does not need to obtain the precise boundary value of disturbance. The key of disturbance reconstruction is to ensure $d_i^- \le d_i(t) \le d_i^+$, which leaves some leeway for the designer to select the appropriate disturbance boundary. The simulation results show that although the three areas are interconnected, the interval disturbance reconstruction technology can still quickly and accurately decouple and reconstruct the load disturbances in each area, which provides an important basis for the design of the subsequent sliding mode load frequency controller.

The real-time disturbance reconstruction value obtained through interval disturbance reconstruction is used to design the new load frequency controller. To achieve control performance, take area 1 as an example; the parameters of controller (25) are designed as follows:

$$K_1 = [6, 1, 0, 3], G_1 = B_1^{\dagger}, \kappa_1 = 0.02, n_1 = 0.03. \tag{47}$$

To further study the effectiveness of the designed interval disturbance reconstruction based sliding mode load frequency control strategy, the control effect is compared with the PI-type load frequency controller (PI-LFC) and the traditional sliding mode controller (SMC-LFC). The design parameters of the PI load frequency controller are obtained through repeated experiments, and the traditional sliding mode controller is selected as follows with disturbance boundary value $\xi_i = 0.01$:

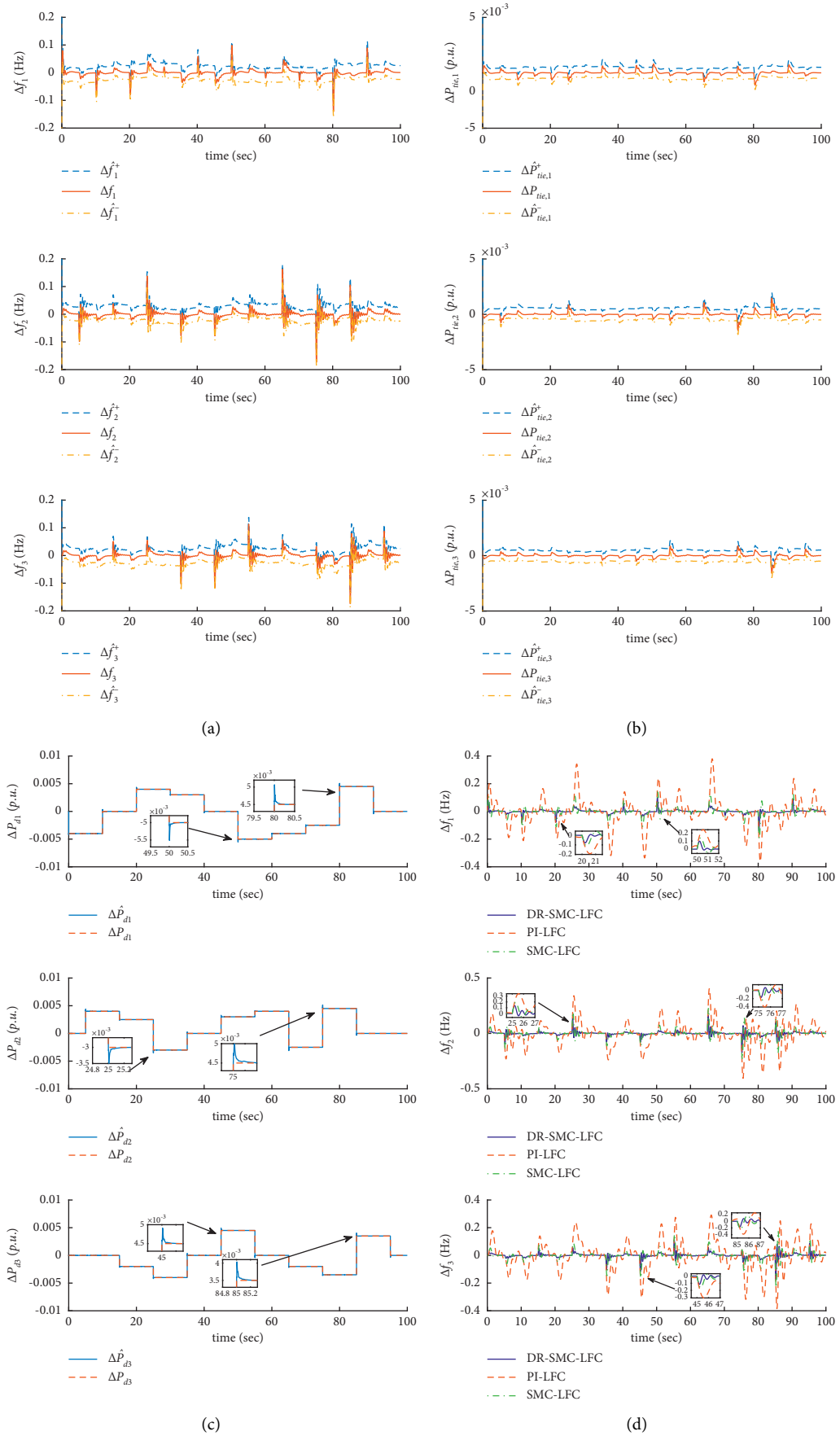


Figure 3: Simulation performance. (a) The interval estimation for $\Delta f_i(t)$ of each power grid. (b) The interval estimation for $\Delta P_{tie,i}(t)$ of each power grid. (c) The reconstruction of the load disturbance $\Delta P_{di}(t)$ of each power grid. (d)$\Delta f_i(t)$ of each power grid under load disturbance.

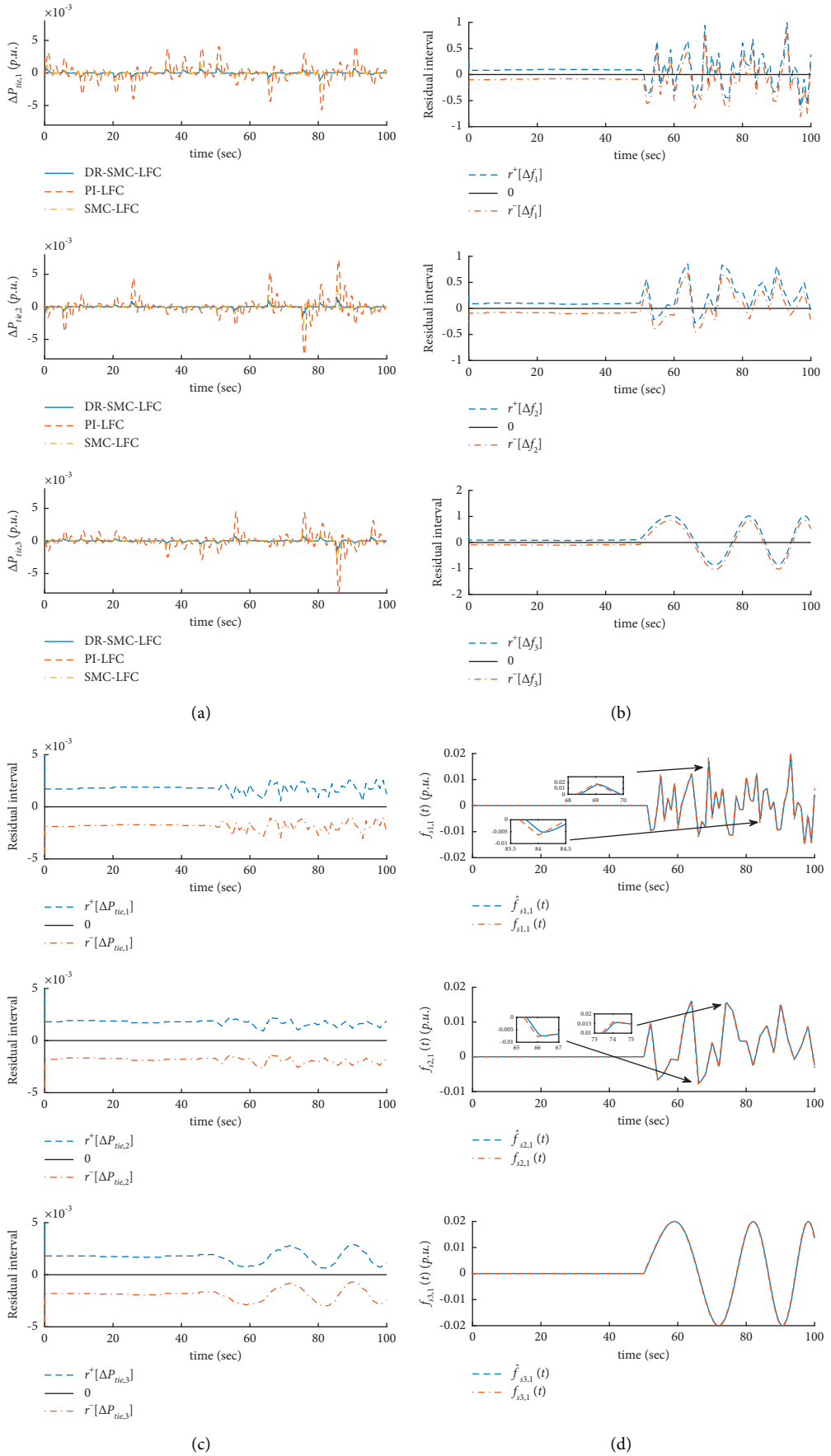


Figure 4: Simulation performance. (a) $\Delta P_{tie,i}(t)$ of each power grid under load disturbance. (b) Interval residual errors of $\Delta f_i(t)$ of each power grid. (c) Interval residual errors of $\Delta P_{tie,i}(t)$ of each power grid. (d) The estimation of $f_{si}(t)$ of each power grid.

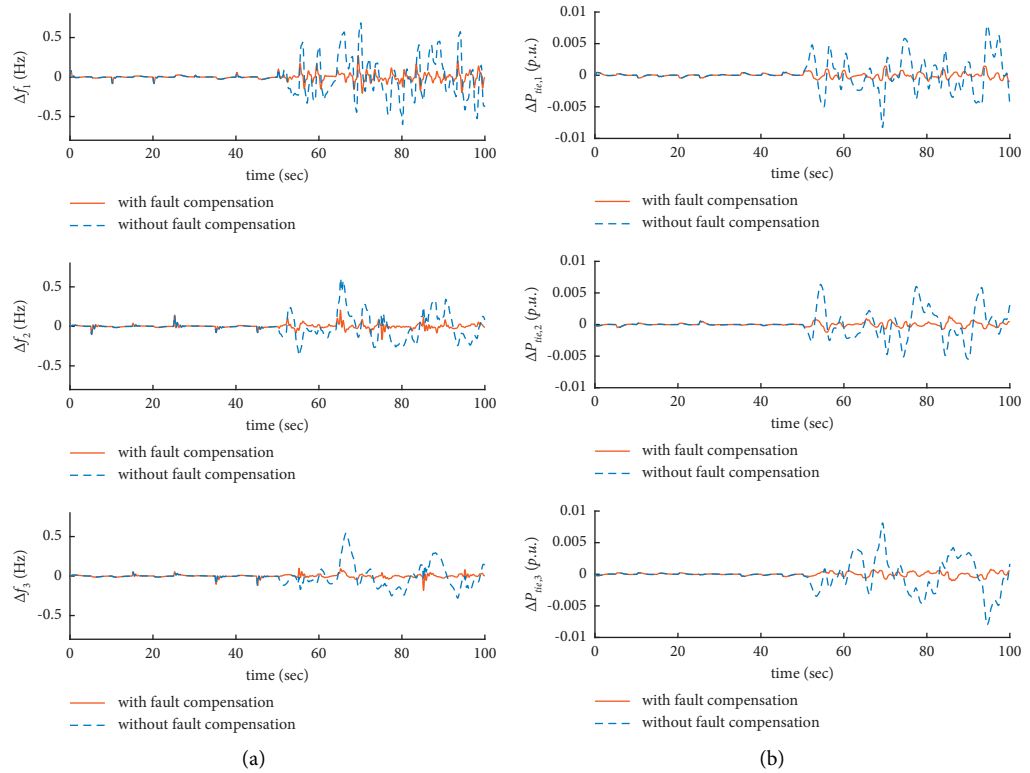


Figure 5: Simulation performance. (a)$\Delta f_i(t)$ and (b)$\Delta P_{tie,i}(t)$ of each power grid under load disturbance and sensor fault.

Table 1: The parameters of the three-area power grid.

	Area 1	Area 2	Area 3
D_i (p.u./Hz)	0.025	0.025	0.025
H_i (p.u. s)	0.3708	0.2708	0.2708
T_{ti} (s)	0.19	0.23	0.23
T_{gi} (s)	0.05	0.05	0.05
R_i (Hz/p.u.)	2.15	2.45	2.45
	$T_{12} = 0.13$, $T_{13} = 0.12$, $T_{23} = 0.10$ (p.u./Hz)		

$$u_i(t) = K_i x_i(t) - (G_i B_i)^{-1}\left[G_i \xi_i + \kappa_i s_i(t) + G_i \sum_{j \neq i}^{N} A_{ij} x_j(t) + m_i \operatorname{sign}(s_i(t))\right]. \tag{48}$$

The comparison simulation results of the three control strategies are shown in Figures 3(d)–4(a). Figure 3(d) shows the frequency deviation of each area under different control strategies; Figure 4(a) shows the power deviation on the tie line between each area and the other areas under different control strategies.

The simulation results show that, under the influence of the same load disturbances, the PI-type load frequency controller cannot quickly respond to the change of the load disturbances, resulting in a large system frequency deviation and tie-line power deviation, and it is difficult to maintain the stability of the grid. The traditional sliding mode controller has a better control effect by virtue of its own sliding mode variable structure, but its control performance is slightly lacking due to the difficulty of setting the appropriate disturbance boundary value. The interval disturbance reconstruction based sliding mode load frequency control strategy can effectively and quickly respond to load changes and maintain the system frequency deviation within a small range of changes.

5.2. Case 2. In actual engineering, the power grid is not only affected by load disturbances in different areas but also faces the problem of sensor fault due to the different characteristics of the sensor itself and the interference of the operating environment. Therefore, in this case, in addition to the different load disturbances in the three areas in Case 1, the problem of sensor fault in the system is also considered, and simulation experiments are carried out. For the sake of simplicity, suppose the form of sensor fault signal is as follows and sensor fault occurs after 50 s:

$$f_{s1}(t) = \left[f_{s1,1}(t) \;\; 0 \;\; 0 \;\; 0 \right]^T,$$
$$f_{s2}(t) = \left[f_{s2,1}(t) \;\; 0 \;\; 0 \;\; 0 \right]^T, \quad (49)$$
$$f_{s3}(t) = \left[f_{s3,1}(t) \;\; 0 \;\; 0 \;\; 0 \right]^T,$$

where $f_{s1,1}(t)$, $f_{s2,1}(t)$, and $f_{s3,1}(t)$ are shown in Figure 4(d).

Interval residual errors based on interval observers have considered system disturbances, so there will be no false alarms due to external disturbances, which can be used for fault diagnosis. Figures 4(b) and 4(c) show the interval residual errors of $\Delta f_i(t)$ and $\Delta P_{tie,i}(t)$, respectively. When the system is operating normally, the interval residual errors satisfy $0 \in [r_{is}^-(t), r_{is}^+(t)]$, but when faults occur, there are $0 \notin [r^-[\Delta f_i], r^+[\Delta f_i]]$ and $0 \in [r^-[\Delta P_{tie,i}], r^+[\Delta P_{tie,i}]]$, from which the fault locations can be diagnosed.

The sensor fault signals will be estimated by the augmented observer online, once the fault is detected. The estimated value of the sensor faults is shown in Figure 4(d), from which we can find that the augmented observer can well track and estimate the sensor fault signals quickly.

Figures 5(a) and 5(b) show $\Delta f_i(t)$ and $\Delta P_{tie,i}(t)$ response, respectively, in each area after sensor faults occur. It can be seen from the simulation experiment that, in the case of the same fault, the controller without considering the fault compensation is difficult to avoid the interference of sensor fault, so that the system frequency and the tie-line power have a large deviation, and it is difficult to maintain the stability of the grid. However, the controller based on fault compensation can well reduce the influence of faults on the system, so that the system frequency and the deviation of tie-line power can be kept within a small range and the stability of the grid can be maintained.

6. Conclusion

A new LFC scheme based on load disturbance reconstruction is proposed for the multiarea power grid, and the corresponding FTC strategy is also developed in this paper. The conclusions of this paper are as follows:

(1) The disturbance is reconstructed based on interval estimation, which solves the problem of decoupling and reconstruction of load disturbance.

(2) An integral sliding mode controller based on disturbance reconstruction is designed to ensure that the frequency deviation of the system is small. The simulation results show that the proposed controller can respond to the load changes effectively and quickly and the system frequency deviation can be kept in a small range.

(3) Sensor faults are diagnosed by interval residual. An augmented observer is designed to estimate sensor faults. The results show that the influence of sensor faults on system performance is greatly eliminated by using sensor fault estimation to compensate the system, and the system runs reliably and stably.

In the future research work, we will also study how to reduce the influence of malicious attacks on the power grid and ensure the stable operation of the power grid.

Data Availability

The data used to support the findings of this study cannot be made freely available. Access to the data is restricted.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (61973140) and Beijing Natural Science Foundation (21JC0026).

References

[1] A. Ameli, A. Hooshyar, E. F. El-Saadany, and A. M. Youssef, "Attack detection and identification for automatic generation control systems," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4760–4774, 2018.

[2] Y. Arya, "ICA assisted (FTIDN)-D-lambda controller for AGC performance enrichment of interconnected reheat thermal power systems," *Journal of Ambient Intelligence and Humanized Computing*, vol. 15, no. 3, 2021.

[3] C. Li, Y. Wu, Y. Sun et al., "Continuous under-frequency load shedding scheme for power system Adaptive frequency control," *IEEE Transactions on Power Systems*, vol. 35, no. 2, pp. 950–961, 2020.

[4] C. Wang, Y. Mi, Y. Fu, and P. Wang, "Frequency control of an isolated micro-grid using double sliding mode controllers and disturbance observer," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 923–930, 2018.

[5] D. Xu, J. Liu, X.-G. Yan, and W. Yan, "A novel adaptive neural network constrained control for a multi-area interconnected power system with hybrid energy storage," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 8, pp. 6625–6634, 2018.

[6] M. Sharma, S. Dhundhara, Y. Arya, and S. Prakash, "Frequency excursion mitigation strategy using a novel COA optimised fuzzy controller in wind integrated power systems," *IET Renewable Power Generation*, vol. 14, no. 19, pp. 4071–4085, 2020.

[7] F. Zhu, X. Zhou, Y. Zhang, D. Xu, and J. Fu, "A load frequency control strategy based on disturbance reconstruction for multi-area interconnected power system with hybrid energy storage system," *Energy Reports*, vol. 7, pp. 8849–8857, 2021.

[8] M. Sharma, S. Dhundhara, Y. Arya, and S. Prakash, "Frequency stabilization in deregulated energy system using coordinated operation of fuzzy controller and redox flow battery," *International Journal of Energy Research*, vol. 45, no. 5, pp. 7457–7475, 2021.

[9] M. Elsisi, M. Soliman, M. A. S. Aboelela, and W. Mansour, "Improving the grid frequency by optimal design of model predictive control with energy storage devices," *Optimal Control Applications and Methods*, vol. 39, no. 1, pp. 263–280, 2018.

[10] M. Q. Tran, M. Elsisi, K. Mahmoud, M. K. Liu, M. Lehtonen, and M. M. F. Darwish, "Experimental setup for online fault diagnosis of induction machines via promising IoT and

machine learning: towards industry 4.0 empowerment," *IEEE access*, vol. 9, pp. 115429–115441, 2021.

[11] M. Elsisi, M. Q. Tran, K. Mahmoud, D. E. A. Mansour, M. Lehtonen, and M. M. F. Darwish, "Towards secured online monitoring for digitalized GIS against cyber-attacks based on IoT and machine learning," *IEEE Access*, vol. 9, 2021.

[12] S. Yin, H. Gao, J. Qiu, and O. Kaynak, "Descriptor reduced-order sliding mode observers design for switched systems with sensor and actuator faults," *Automatica*, vol. 76, pp. 282–292, 2017.

[13] M. Liu and P. Shi, "Sensor fault estimation and tolerant control for Itô stochastic systems with a descriptor sliding mode approach," *Automatica*, vol. 49, no. 5, pp. 1242–1250, 2013.

[14] K. Wang, Y. Shao, L. Xie, J. Wu, and S. Guo, "Adaptive and fault-tolerant data processing in healthcare IoT based on fog computing," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 1, pp. 263–273, 2020.

[15] X. Su, X. Liu, and Y.-D. Song, "fault-tolerant control of multiarea power systems via a sliding-mode observer technique," *IEEE*, vol. 23, no. 1, pp. 38–47, 2018.

[16] J. Wu, Z. Chen, Y. Zhang, Y. Xia, and X. Chen, "Sequential recovery of complex networks suffering from cascading failure blackouts," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 2997–3007, 2020.

[17] N. Jaleeli, L. S. VanSlyck, D. N. Ewart, L. H. Fink, and A. G. Hoffmann, "Understanding automatic generation control," *IEEE Transactions on Power Systems*, vol. 7, no. 3, pp. 1106–1122, 1992.

[18] H. A. Yousef, K. Al-Kharusi, M. H. Albadi, and N. Hosseinzadeh, "Load frequency control of a multi-area power system: an adaptive fuzzy logic approach," *IEEE Transactions on Power Systems*, vol. 29, no. 4, pp. 1822–1830, 2014.

[19] S. Kayalvizhi and D. M. Vinod Kumar, "Load frequency control of an isolated micro grid using fuzzy adaptive model predictive control," *IEEE Access*, vol. 5, pp. 16241–16251, 2017.

[20] S. Trip and C. De Persis, "Distributed optimal load frequency control with non-passive dynamics," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 1232–1244, 2018.

[21] C. Chen, M. Cui, X. Fang, B. Ren, and Y. Chen, "Load altering attack-tolerant defense strategy for load frequency control system," *Applied Energy*, vol. 280, Article ID 116015, 2020.

[22] A. M. Ersdal, L. Imsland, K. Uhlen, D. Fabozzi, and N. F. Thornhill, "Model predictive load-frequency control taking into account imbalance uncertainty," *Control Engineering Practice*, vol. 53, pp. 139–150, 2016.

[23] A. M. Ersdal, L. Imsland, and K. Uhlen, "Model predictive load-frequency control," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 777–785, 2016.

[24] H. Bevrani, *Robust Power System Frequency Control*, Springer, New York, NY, USA, 2009.

[25] P. Ojaghi and M. Rahmani, "LMI-based robust predictive load frequency control for power systems with communication delays," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 4091–4100, 2017.

[26] X. Shang-Guan, Y. He, C. Zhang, L. Jiang, J. W. Spencer, and M. Wu, "Sampled-data based discrete and fast load frequency control for power systems with wind power," *Applied Energy*, vol. 259, Article ID 114202, 2020.

[27] Y. Mi, Y. Fu, D. D. Li, C. S. Wang, P. C. Loh, and P. Wang, "The sliding mode load frequency control for hybrid power system based on disturbance observer," *International Journal of Electrical Power & Energy Systems*, vol. 74, pp. 446–452, 2016.

[28] Z. Wang, C.-C. Lim, and Y. Shen, "Interval observer design for uncertain discrete-time linear systems," *Systems & Control Letters*, vol. 116, pp. 41–46, 2018.

[29] J. Li, Z. Wang, W. Zhang, T. Raïssi, and Y. Shen, "Interval observer design for continuous-time linear parameter-varying systems," *Systems & Control Letters*, vol. 134, Article ID 104541, 2019.

[30] S. Guo, B. Jiang, F. Zhu, and Z. Wang, "Luenberger-like interval observer design for discrete-time descriptor linear system," *Systems & Control Letters*, vol. 126, pp. 21–27, 2019.

[31] F. Xu, J. Tan, X. Wang, and B. Liang, "Conservatism comparison of set-based robust fault detection methods: set-theoretic UIO and interval observer cases," *Automatica*, vol. 105, pp. 307–313, 2019.

[32] M. Liu, P. Shi, L. Zhang, and X. Zhao, "fault-tolerant control for nonlinear markovian jump systems via proportional and derivative sliding mode observer technique," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 58, no. 11, pp. 2755–2764, 2011.

[33] F. Boem, A. J. Gallo, D. M. Raimondo, and T. Parisini, "distributed fault-tolerant control of large-scale systems: an active fault diagnosis approach," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 1, pp. 288–301, 2020.

[34] A. M. H. Teixeira, J. Araújo, H. Sandberg, and K. H. Johansson, "Distributed sensor and actuator reconfiguration for fault-tolerant networked control systems," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 4, pp. 1517–1528, 2018.

[35] X. Su, P. Shi, L. Wu, and Y.-D. Song, "fault detection filtering for nonlinear switched stochastic systems," *IEEE Transactions on Automatic Control*, vol. 61, no. 5, pp. 1310–1315, 2016.

[36] C. Ke, C. Li, and Q. Zhang, "Distributed adaptive fault-tolerant consensus of nonlinear multi-agent systems via state-constraint impulsive protocols with time-delay," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 3112–3121, 2020.

[37] X. G Yan and C. Edwards, "Adaptive sliding-mode-observer-based fault reconstruction for nonlinear systems with parametric uncertainties," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 11, pp. 4029–4036, 2008.

[38] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*, Wiley, New York, NY, USA, 2012.

[39] L. Farina and S. Rinaldi, *Positive Linear Systems: Theory and Applications*, Wiley, New York, NY, USA, 2000.

[40] P. Gahinet and P. Apkarian, "A linear matrix inequality approach toH∞ control," *International Journal of Robust and Nonlinear Control*, vol. 4, no. 4, pp. 421–448, 1994.

Hindawi
Complexity
Volume 2022, Article ID 1456570, 24 pages
https://doi.org/10.1155/2022/1456570

Research Article

Covert Communication for Wireless Networks with Full-Duplex Multiantenna Relay

Ling Yang,[1,2] **Weiwei Yang**,[1] **Liang Tang**,[1,3] **Liwei Tao**,[1] **Xingbo Lu**,[1] **and Zhengyun He**[2]

[1]*College of Communications Engineering, PLA Army Engineering University, Nanjing 210 007, China*
[2]*College of Railway Transportation, Hunan University of Technology, Zhuzhou 412 007, China*
[3]*College of Science, Hunan University of Technology, Zhuzhou 412 007, China*

Correspondence should be addressed to Weiwei Yang; weiweiyang1981@126.com

Received 19 August 2021; Accepted 22 November 2021; Published 7 January 2022

Academic Editor: Hou-Sheng Su

In this work, we investigated a covert communication method in wireless networks, which is realized by multiantenna full-duplex single relay. In the first stage, the source node sends covert messages to the relay, and the relay uses a single antenna to send interference signals to the adversary node to protect the covert information being transmitted. In the second stage, the relay decodes and forwards the covert information received in the first stage; at the same time, the relay uses zero-forcing beamforming to send interference signals to the warden to ensure covert transmission. The detection error rate, transmission outage probability, maximum effective covert rate, and other performance indicators are derived in two stages, and the total performance of the system is derived and analyzed. Then, the performance indicators are verified and analyzed by simulation. Our analysis shows that the maximum effective covert rate of using the characteristics of multiantenna to interfere with Willie in the second stage is taken as the total covert performance of the system, and the transmission interruption probability is significantly less than that of the first stage, so the corresponding maximum effective concealment efficiency will be greater.

1. Introduction

1.1. Background. Networks are ubiquitous in nature and human society. Modern communication has been highly concerned by academic and industrial circles, resulting in many emerging sciences, such as the internet of things [1], complex networks [2–5], cognitive networks [6, 7], virtual reality, and augmented reality technology. Due to the broadcast characteristics of wireless channels, whether users are legal or not, researchers pay more and more attention to the security and privacy of information transmission. Taking full advantage of the uncertainty and unpredictability of wireless channels, physical layer security technology [8, 9] has become a mature technology, which has been applied to achieve secure transmission, focusing on information theory technology [10, 11]. However, in many communication cases, not only the privacy and integrity of information need to be considered, but also the security of communication behavior needs to be protected, such as the existence of hidden communication [12]. In this context, covert communication [13, 14] emerged as a new security technology.

In 2013, B. A. Bash proved the reliable communication with low detection probability under additive white Gaussian noise (AWGN) channel theoretically for the first time [15] and proposed the square root rule of covert communication under AWGN channel. Then, the basic research of covert communication has attracted the interest of relevant scholars and has been studied in AWGN channels [16], discrete memoryless channels (DMC) [17, 18], and binary symmetric channel (BSC) [19]. At the same time, the basic limitations of covert communication under different channel conditions have been explored [20]. With the deepening of theoretical research, the concept of covert communication has gradually formed and developed. So far, researchers represent by Dr. B. A. Bash and Dr. Shihao Yan team of Macquarie University in Australia have made contributions to the basic theoretical research and performance analysis of covert communication technology;

covert communication technology has attracted more and more researchers' attention, and its related research has gradually developed and enriched [21–23].

1.2. Previous Works and Motivations. Recently, more and more scenes and technologies about covert communication have been studied. In this paper, we focus on full duplex (FD), relay, and multiantenna covert technology.

Reference [24] explores the possibility and condition of covert communication in quasi-static wireless fading channel by using an FD receiver to transmit artificial noise (AN) to enhance covert performance. Reference [25] not only studies the influence of channel estimation inaccuracy and randomness on covert wireless transmission performance but also explores the effect of FD relay on covert performance. Furthermore, in ad hoc networks, the internet of things, and other random wireless networks, reference [26] also uses FD technology to effectively enhance the covert of communication.

Reference [25] also studies the influence of channel uncertainty of detecting attackers on covert wireless transmission performance in two-hop relay communication system. Reference [27] studies the covert and transmission reliability of covert wireless transmission schemes with the assistance of a single relay and shows that cooperative relay can enhance the performance of covert wireless transmission. In [28], a covert wireless transmission scheme is designed with the assistance of a wireless energy acquisition relay, and the expression of minimum error detection probability is derived. In [29], the problem of the covert wireless transmission assisted by the untrusted relay is studied. References [30, 31] study the covert wireless communication scheme with multihop relay transmission. In particular, reference [31] considers the problem of multihop covert relay transmission in unmanned aerial vehicle communication. Most of these researches focus on reducing transmission power by relay transmission to enhance covertness or improving the communication performance of legal links by using cooperative diversity gain. However, the probability of signal exposure caused by relay transmission increases, and the performance degradation is not clear.

From [32], an active eavesdropping scheme assisted by the covert pilot attack is designed for wireless monitoring scenarios, and covert wireless communication is realized by using the uncertainty of the detection channel. The communication process is divided into the channel estimation stage and the data transmission stage. Through the use of a malicious detection node (source node) equipped with multiple antennas to realize hidden pilot attacks on channel status information in the channel estimation stage, the information beam of the communication process is forced to point to the wireless monitoring node in the data transmission stage, and the probability of wireless monitoring success is improved. Reference [33] considers centralized and distributed antenna systems and discusses the coverage and reliability of transmission. In addition, the covert communication model of multiantenna detector is also studied in reference [34], and the influence of the increase of the number of detectors' antennas on the covert performance is discussed. In recent years, researchers consider using multiantenna jammers to improve the performance of covert communication systems [35]. In [36], the enhancement effect of multiantenna AN nodes on covert performance is studied, and it is pointed out that directional beamforming is the optimal AN strategy. In [14], an FD multiantenna receiver is used to achieve covert communication based on the uncertainty of interference power. The receiver first selects the best antenna to receive the covert information and then randomly selects one of the remaining antennas to generate AN, resulting in the uncertainty of the detector so as to achieve covert communication.

In the present work, the FD, multiantenna and cooperative transmission method under covert communication needs to be further studied. On the one hand, the existing research of covert communication often focuses on the covert transmission of its own important information based on the relay forwarding the source's message and seldom considers that the information sent by the source are covert messages at the beginning, so the established scene has certain limitations. On the other hand, the reliability of relay nodes in covert wireless communications or covert wireless networks in the presence of eavesdropping is rarely involved. How to combine multiantenna, FD, cooperative technology, and covert communication is worthy of further study. Based on this, the research point of covert communication in multiantenna FD relay system is proposed.

1.3. Our Approach and Contributions. The main contributions of this paper are summarized as follows:

(i) We prove that the use of multiantenna FD relay is an effective way to achieve covert wireless communication in fading wireless channels. Multiantenna FD relay uses the advantages of multiantenna to design different antenna selection schemes in two stages to transmit signals with different power to cause Willie's confusion.

(ii) Based on the assumption of the Willie radiometer, we analytically derive Willie's optimal detection thresholds for the two stages of the system. When we define the optimal decision rules for the minimum detection error rate, the predetermined thresholds of the two stages are consistent, and the optimal detection performance is obtained with the minimum detection error probability.

(iii) For the given covert constraints, we give the design criteria of the optimal interference power of the first stage relay and the optimal transmission power of the second stage relay to forward the covert messages so as to maximize the optimal effective covert rate of the system.

(iv) Our analysis shows that for the same parameters, the transmission outage probabilities of the first stage, the second stage, and the total results have the same trend. Since the total maximum effective covert rate is smaller of the two stages, the total maximum effective covert rate has the same conclusion as that of the first stage.

1.4. Organization. The rest of this paper is organized as follows. The system model is introduced in Section 2. Sections 3 and 4, respectively, introduce the two stages of the system. The first stage is that the source node sends covert messages to the relay, and the second stage is that the relay forwards covert messages to the destination node. The detection and covert performance of the system are deduced and analyzed in the two stages. In Section 5, the total performance of the system is studied. Section 6 provides the theoretical analyses are verified by numerical results. Finally, Section 7 describes some concluding remarks. A list of the fundamental variables is provided in Table 1.

2. System Model

As shown in Figure 1, we consider a covert wireless communication model in a single relay multiantenna network, which includes a transmitter (Alice), a relay (Relay), a receiver (Bob), and an adversary (Willie). Among them, Relay is equipped with N_R antennas, and other nodes are equipped with a single antenna. Since there is no direct communication link between Alice and Bob, it is necessary to help Alice send covert messages to Bob through Relay, while Willie tries to detect such covert transmissions, and each role knows each other's existence and location. Information transmission is divided into two stages. In the first stage, Alice sends covert messages to Relay, and Relay sends artificial noise (AN) signals to Willie while receiving the messages. In order to improve the effective covert rate as much as possible, Relay selects the best antenna between it and Alice for receiving, one of the remaining $N_R - 1$ antennas was selected at random to send AN to Willie. In the second stage, Relay selects an optimal antenna to forward Alice's signal to Bob by decode-and-forward and uses the remaining $N_R - 1$ antennas to transmit AN signal to Willie by using zero-forcing beamforming (ZFB).

Assuming that the transmission of the first stage is completed before the transmission of the second stage, it is necessary to ensure that neither stage can be detected by Willie in order to realize the covert transmission from Alice to Bob. Since Willie needs to detect whether Alice and Relay send covert messages in the first and second stages, respectively, considering the worse case, it is assumed that Willie knows h_{AW} and h_{RW} and that Relay knows h_{AR}, and since Relay knows Willie's existence, it is assumed that Relay also knows h_{RW} [1]. The wireless channels are subject to independent quasi-static Rayleigh fading; the channels change independently between time slots and remain unchanged within the same communication time slot.

3. The First Stage: Alice Transmits Covert Messages to Relay

3.1. Relay Receives Covert Messages. The instantaneous signal-to-interference-plus-noise ratio ($SINR$) at Relay is given by

$$\gamma_{Relay} = \frac{P_A \max_{1\le i\le N_R}\left(\left|h_{AR_i}\right|^2\right)}{\varphi P_{R_j}\left|h_{R_jR_j}\right|^2 + \sigma_R^2}. \tag{1}$$

Relay selects the best i-th antenna to receive the signal according to the channel conditions between Alice and Relay, and its channel coefficient is expressed as h_{AR_i}, where $i = 1, 2, \ldots, n$, P_A is the power of Alice sending covert messages, P_{R_j} is the power of AN sent by Relay, $P_{R_jR_j}$ is the channel coefficient of Relay itself, and j is the j-th antenna randomly selected by Relay in $N_R - 1$ antennas. Since Relay itself knows AN signal, the residual noise can be reconstructed and eliminated through self-interference elimination. ϕ is used to represent the self-interference elimination coefficient. $\phi = 0$ is the ideal situation, while $0 \le \phi \le 1$ refers to different self-interference elimination levels [37]. σ_R^2 is the channel variance of Relay. The maximum power of Relay is $P_R^{\max}$. Assuming that the power is evenly distributed among the antennas of Relay, let $P_{R_j}^{\max} = P_R^{\max}/N_R$. We assume that P_A is fixed [38], and both Relay and Willie know it. The P_{R_j} changes from slot to slot, following a continuous uniform distribution over the interval $[0, P_{R_j}^{\max}]$, having a probability density function ($P\ DF$) given by

$$f_{P_{R_j}}(x) = \begin{cases} \dfrac{1}{P_{R_j}^{\max}}, & 0 \le x \le P_{R_j}^{\max}, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Reasons for setting P_{R_j} to change between slots: the purpose of Relay to send AN power is to make the power received by Willie uncertain. Willie knows h_{AW} in a time slot. If the AN power is constant, Willie can directly detect the covert transmission when Alice sends fixed covert messages, so the P_{R_j} is not fixed.

In the first stage, when Alice sends covert messages, the signal received by Relay is as follows:

$$\begin{aligned} Y_{\text{Relay}}[i] = {} & \sqrt{P_A}\max\left(h_{AR_i}\right)x_A[i] \\ & + \sqrt{\varphi P_{R_j}}h_{R_jR_j}V_R[i] + n_R[i], \end{aligned} \tag{3}$$

where $x_A[i]$ represents the signals transmitted by Alice, satisfying $E[x_A[i]x_A^{\dagger}[i]] = 1$, $i = 1, 2, \ldots, n$ represents the symbol index, n is the total number of channels used in each slot, and $n_R[i]$ is the complex additive white Gaussian noise ($AWGN$) at Relay with σ_R^2 as its variance, that is, $n_R[i] \sim CN(0, \sigma_R^2)$.

3.2. Detection Metrics at Willie. In the first stage, the SINR at Willie, in case Alice transmits, is given by

$$\gamma_{Willie1} = \frac{P_A\left|h_{AW}\right|^2}{P_{R_j}\left|h_{R_jW}\right|^2 + \sigma_W^2}, \tag{4}$$

where h_{AW} is the channel coefficient between Alice and Willie, h_{R_jW} represents the channel coefficient between the j-th antenna that Relay randomly selects to send interference and Willie, σ_W^2 is the noise variance of Willie. In a communication time slot, Willie has to decide whether Alice has transmitted covert messages to Relay in the first stage. Therefore, Willie is faced with a binary hypothesis testing

Table 1: List of fundamental variables.

Symbol	Description
$SINR$	Interference plus noise ratio
$AWGN$	The complex additive Gaussian noise
$f(\cdot)$	The cumulative distribution function
P_X	The power send by X, $X \in \{Alice, Relay, Bob, Willie\}$
y_A	The received signal at A, $A \in \{Bob, Relay, Willie\}$
$h_{a,b}$	The channel between a and b that both consist a single antenna
$n_A[i]$	The AWGN at A, $A \in \{Bob, Relay, Willie\}$
γ_A	The SINR on A, $A \in \{Bob, Willie\}$
n	The number of each channel use
σ_A^2	The variance of AWGN at node A, $A \in \{Bob, Willie\}$
T_W	The average power received at Willie
τ	The predefined detection threshold
ξ	The detection error rate
τ^*	The optimal predefined detection threshold
ξ^*	The minimum detection error rate
$\overline{R_C}$	The maximized effective covert rate
$F(\cdot)$	The probability density function
ε	The predetermined covert communication condition value
R	The transmission rate
δ	The transmission outage probability
N_A	Antenna number of A, $A \in \{Bob, Relay, Willie\}$
$h_{a,b}$	The channel vector between a and b consisting of one and multiple antennas, respectively
$w_{a,b}$	The received weighting vector on link $a-b$
$\dagger$	The conjugate transpose operator
$\|\bullet\|_F$	The Frobenius norm
$T^{\perp}$	The projection idempotent matrix
α	False alarm rate
β	Miss detection rate

problem, where the null hypothesis H_0 means that Alice has not sent covert messages, and the alternative hypothesis H_1 means that Alice has sent covert messages to Relay.

The signal received by Willie in the first stage can be expressed as follows:

$$y_{willie}[i] = \begin{cases} \sqrt{P_{R_j}}\, h_{R_jW} v_R[i] + n_W[i], & H_0, \\ \sqrt{P_A}\, h_{AW} x_A[i] + \sqrt{P_{R_j}}\, h_{R_jW} v_R[i] + n_W[i], & H_1, \end{cases} \tag{5}$$

where $n_W[k]$ is the AWGN at Willie with σ_W^2 as its variance, that is, $n_W[k] \sim CN(0, \sigma_W^2)$. Willie does not know the value of P_{R_j} in the time slot, but the value of P_A is fixed and known. In the first stage, Willie attempts to detect whether y_{willie} is H_0 or H_1. Through the application of Neyman–Perason criterion [39] and likelihood ratio test, the optimal decision rule for Willie to minimize his detection error is as follows:

$$T_W \underset{D_0}{\overset{D_1}{\gtrless}} \tau, \tag{6}$$

where $T_W = 1/n \sum_{i=1}^{n} y_{Willie1}[i]$ is the average power received at Willie in the slot and τ is a predefined Willie's detection threshold. D_0 and D_1 are binary decisions that infer whether Alice sends covert messages. In this paper, we consider an infinite number of channel uses, that is, $n \longrightarrow \infty$. Therefore, we have

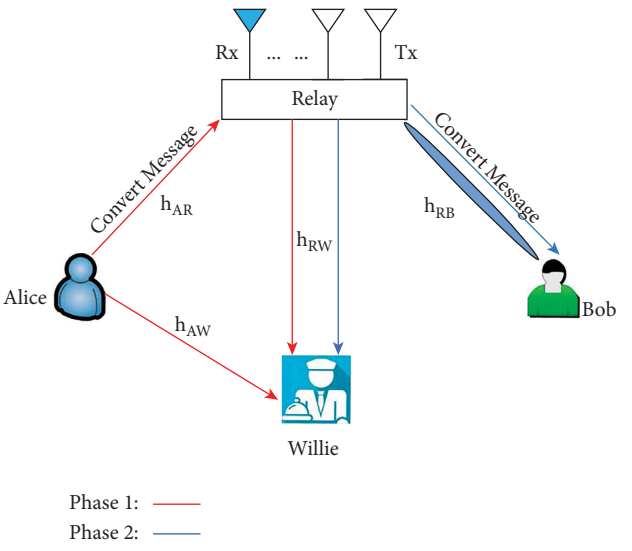


Figure 1: System model.

$$T_W = \begin{cases} P_{R_j} \left| h_{R_jW} \right|^2 + \sigma_W^2, & H_0, \\ P_A \left| h_{AW} \right|^2 + P_{R_j} \left| h_{R_jW} \right|^2 + \sigma_W^2, & H_1. \end{cases} \tag{7}$$

At the end of the communication slot, Willie has to make a decision. The false alarm rate is defined as the probability of Willie making D_1 decision under condition H_0, which is expressed by $\alpha = P(D_1 \mid H_0)$. Similarly, the miss detection rate is defined as the probability of Willie making D_0 decision under condition H_1, which is expressed by $\beta = P(D_0|H_1)$. Assuming that the prior probabilities of H_0 and H_1 are equal, Willie's detection performance can be judged by the detection error rate, which can be defined as follows:

$$\xi \triangleq \alpha + \beta. \tag{8}$$

3.3. Detection Performance at Willie. According to Theorems 1 and 2 in [37], Willie's optimal detection threshold, minimum detection error probability, and expected detection error probability can be obtained.

The optimal detection threshold is expressed as follows:

$$\tau^* = \begin{cases} [\rho_1, \rho_2], & \rho_1 < \rho_2, \\ [\rho_2, \rho_1], & \rho_1 \geq \rho_2. \end{cases} \tag{9}$$

The minimum detection error probability is expressed as follows:

$$\xi^* = \begin{cases} 0, & \rho_1 < \rho_2, \\ 1 - \dfrac{P_A \left|h_{AW}\right|^2}{P_{R_j}^{\max} \left|h_{R_j W}\right|^2}, & \rho_1 \geq \rho_2, \end{cases} \tag{10}$$

where $\rho_1 \triangleq P_{R_j}^{\max} |h_{R_j W}|^2 + \sigma_W^2$, $\rho_2 \triangleq P_A |h_{AW}|^2 + \sigma_W^2$, $\rho_3 \triangleq P_A |h_{AW}|^2 + P_{R_j}^{\max} |h_{R_j W}|^2 + \sigma_W^2$.

The expected detection error probability is expressed as follows:

$$\overline{\xi^*}(t) = -t^2 + t \ln t + 1, \tag{11}$$

where

$$t \triangleq \frac{P_A \left|h_{AW}\right|^2}{P_A \left|h_{AW}\right|^2 + P_{R_j}^{\max} \left|h_{R_j W}\right|^2}. \tag{12}$$

It can be seen from [2] that the expected detection error rate $\overline{\xi^*}$ is a monotone increasing function of $P_{R_j}^{\max}$.

3.4. Covert Performance. In general, the constraint for covert transmission can be defined as $\xi^* \geq 1 - \varepsilon$, where ε is a predetermined value, and there is $\varepsilon \in [0, 1]$. Therefore, the maximized effective covert rate can be expressed as follows:

$$\begin{aligned} &\max_{P_Z} \overline{R_C}, \\ &s.t.\ \overline{\xi^*} \geq 1 - \varepsilon, \end{aligned} \tag{13}$$

where $\overline{R_C} \triangleq R(1 - \delta)$.

3.4.1. The Transmission Outage Probability of Alice to Relay. Assuming that the transmission rate R is known, according to formula (1), $|h_{AR_i}|^2$, $|h_{R_j R_j}|^2$ and P_{R_j} are random variables, which can still cause transmission interruption.

Since the wireless channel is subject to independent quasi-static Rayleigh fading and independent identically distributed, the cumulative distribution function (*CDF*) of $\max_{1 \leq i \leq N_R} (|h_{AR_i}|^2)$ is $F_{\max\limits_{1 \leq i \leq N_R} (|h_{AR_i}|^2)}(x) = (1 - e^{-x})^{N_R}$; then its PDF is

$$f_{\max\limits_{1 \leq i \leq N_R} \left(\left|h_{AR_i}\right|^2\right)}(x) = \sum_{m=0}^{N_R} \binom{N_R}{m} (-1)^m \left(e^{-x}\right)^m. \tag{14}$$

Theorem 1. *For the first stage, the transmission outage probability from Alice to Relay is derived as follows:*

$$\delta_1 = 1 + \frac{P_A}{P_{R_j}^{\max} \varphi\left(2^{2R} - 1\right)} \sum_{m=1}^{N_R} \binom{N_R}{m} (-1)^m \frac{1}{m}, \exp\left[-\frac{m\left(2^{2R} - 1\right)\sigma_R^2}{P_A}\right] \times \ln\left[1 + m\varphi\left(2^{2R} - 1\right)\frac{P_{R_j}^{\max}}{P_A}\right]. \tag{15}$$

Proof. see Appendix A. □

Remark 1. The comments are as follows:

(1.1) In the first stage, the maximum power $P_{R_j}^{\max}$ of transmitting AN of a single antenna randomly selected by Relay can have a direct influence on the transmission outage probability δ_1. The larger $P_{R_j}^{\max}$ is, the greater δ_1 will be.

(1.2) In the first stage, the transmission outage probability δ_1 is a monotonically increasing function of channel noise σ_R^2 and transmission rate R, that is, the larger σ_R^2 and R are, the larger δ_1 will be.

(1.3) The configuration of antenna number N_R has a direct influence on the transmission outage probability δ_1 of the first stage. The larger N_R is, the smaller δ_1 is.

(1.4) In the first stage, the transmission outage probability δ_1 is a monotonically decreasing function of Alice's covert message sending power P_A, that is, the larger P_A is, the smaller δ_1 is. And the transmission outage probability δ_1 is a monotonically increasing function of the self-interference elimination coefficient ϕ, that is, the larger ϕ is, the larger δ_1 is.

3.4.2. The Optimal AN

Theorem 2. *Under any given fixed covert information power P_A sent by Alice, predetermined ε and transmission rate R, the optimal AN power sent by Relay is given by*

$$P_{R_j}^{\max^*} = \frac{(1-t_\varepsilon)P_A\left|h_{AW}\right|^2}{\left|h_{R_jW}\right|^2 t_\varepsilon}, \tag{16}$$

where t_ε is the solution of $\overline{\xi^*}(t) = 1-\varepsilon$.

Proof. see Appendix B. □

3.4.3. The Maximized Effective Covert Rate

Theorem 3. *The optimal effective convert rate is derived as follows:*

$$\overline{R_{C_1}^*} = -R\left|h_{R_jW}\right|^2 t_\varepsilon \sum_{m=1}^{N_R}\binom{N_R}{m}(-1)^m \times \exp\left[-\frac{m\left(2^{2R}-1\right)\sigma_R^2}{P_A}\right]\frac{\ln\left[1+m\varphi\left(2^{2R}-1\right)(1-t_\varepsilon)\left|h_{AW}\right|^2/t_\varepsilon\left|h_{R_jW}\right|^2\right]}{m\varphi\left(2^{2R}-1\right)(1-t_\varepsilon)\left|h_{AW}\right|^2}. \tag{17}$$

Proof. see Appendix C. □

Remark 2. The comments are as follows:

(2.1) The larger the predetermined convert constraint ε in the first stage, the greater the maximum effective convert rate of the first stage.

Due to $\overline{R_C} \triangleq R(1-\delta)$, then there is δ_1^*, so there will be $\overline{R_{C_1}^*} \triangleq R(1-\delta_1^*)$, and δ_1^* also has the property of δ_1, so we can obtain Remarks 2.2–2.4.

(2.2) In the first stage, the maximum effective covert rate $\overline{R_{C_1}^*}$ increases with the increase of the antenna number N_R and the power P_A of covert messages sent by Alice.

(2.3) In the first stage, the maximum effective covert rate $\overline{R_{C_1}^*}$ is the monotonic decreasing function of the channel noise σ_R^2 of Relay and its self-interference elimination coefficient ϕ, that is, the larger σ_R^2 and ϕ are, the smaller $\overline{R_{C_1}^*}$ is.

(2.4) In the first stage, the maximum effective covert rate $\overline{R_{C_1}^*}$ is the monotonic decreasing function of the channel coefficient $|h_{AW}|^2$, that is, the larger $|h_{AW}|^2$ is, the smaller $R_{C_1}^*$ is. And the maximum effective covert rate $\overline{R_{C_1}^*}$ is the monotonic increasing function of the $|h_{R_jW}|^2$, that is, the larger $|h_{R_jW}|^2$ is, the larger $\overline{R_{C_1}^*}$ is.

Corollary 1. *In the first stage, if the power P_A of Alice sending covert messages to Relay increases, the maximum effective covert rate $\overline{R_{C_1}^*}$ tends to a fixed value:*

$$\overline{R_{C_1}^*} = R\left|h_{R_jW}\right|^2 t_\varepsilon \sum_{m=1}^{N_R}\binom{N_R}{m}(-1)^m, \frac{\ln\left[1+m\varphi\left(2^{2R}-1\right)(1-t_\varepsilon)\left|h_{AW}\right|^2/t_\varepsilon\left|h_{R_jW}\right|^2\right]}{m\varphi\left(2^{2R}-1\right)(1-t_\varepsilon)\left|h_{AW}\right|^2}. \tag{18}$$

Proof. when P_A approaches ∞, then $\exp\{-[m(2^{2R}-1)\sigma_R^2]/P_A\}$ approaches 1, so the maximum effective convert rate approaches the result in formula (18) in the first stage. □

4. The Second Stage: Relay Forwards Covert Messages to Bob

Relay forwards Alice's information to Bob and sends zero-forcing AN signal to Willie. Relay selects the best antenna according to the CSI between it and Bob and decodes and forwards the convert messages sent by Alice to Bob, and the remaining $N_R - 1$ antennas are all used to send zero-forcing AN to Willie.

Relay works in the decoding and forwarding mode. In the second stage, Relay decodes and encodes the received signal and forwards it to Bob, whose transmitted signal is $X_R[i]$. The P_A is fixed, and Relay's forwarding mode is decoded forwarding, so it is assumed that Alice's covert messages power is also fixed by Relay forwarding; both Bob and Willie know this.

4.1. Reception at Bob. The instantaneous SINR at Bob is given by

$$\gamma_{Bob} = \frac{P_{R_k}\max_{1\le k\le N_R}\left(\left|h_{R_kB}\right|^2\right)}{\sigma_B^2}, \tag{19}$$

where h_{R_kB} is the channel coefficient between the k-th antenna selected by Relay and Bob; here, $k = 1, 2, \ldots, N_R$, P_{R_k} is the fixed power of Relay forwarding Alice covert messages; σ_B^2 is the channel variance of Relay.

In the second stage, when Relay sends covert messages, the signal received by Bob can be expressed as follows:

$$y_{Bob}[i] = \sqrt{P_{R_k}}\max_{1\le k\le N_R}\left(\left|h_{R_kB}\right|\right)x_R[i] + n_B[i], \tag{20}$$

where $x_R[i]$ is the covert signal forwarded by Relay, satisfying $E[x_R[i]x_R^\dagger[i]] = 1; i = 1, 2, \ldots, n$ represents the symbol index, n is the total number of channels used in each slot, and $n_B[i]$ is the AWGN at Bob with σ_B^2 as its variance, i.e., $n_B[i] \sim CN(0, \sigma_B^2)$.

4.2. Detection Metrics at Willie. In order to cause the uncertainty of Willie's detection power, Relay uses the remaining $N_R - 1$ antennas for zero-forcing beamforming, which interferes with Willie's transmission without affecting Bob's reception. The optimal weighted vector $\mathbf{W}_{ZF}$ is the solution of the following optimization problem:

$$\begin{aligned} &\max_{\mathbf{W}_{ZF}} \left|h_{RW}^{\dagger}\mathbf{W}_{ZF}\right|, \\ &s.t. \left|h_{RB}^{\dagger}\mathbf{W}_{ZF}\right| = 0 \quad \left\|\mathbf{W}_{ZF}\right\|_F = 1, \end{aligned} \tag{21}$$

where † is the conjugate transpose operator, $\|\bullet\|_F$ denotes the Frobenius norm, and h_{RW} or h_{RB} represents the n-dimensional channel vectors between Relay and Willie or Bob, respectively. According to the theorems in [40, 41], the solution of the optimization problem in formula (21), i.e., the precoding vector $\mathbf{W}_{ZF}$, can be described as follows:

$$\mathbf{W}_{ZF} = \frac{T^{\perp}h_{RW}}{\left\|T^{\perp}h_{RW}\right\|}, \tag{22}$$

where $T^{\perp} = (I - h_{RB}(h_{RB}^{\dagger}h_{RB})^{-1}h_{RB}^{\dagger})$ is the projection idempotent matrix with rank $N_R - 2$.

Let us define $Z_1 \triangleq P_Z|h_{RW}^{\dagger}W_{ZF}|^2/\sigma_W^2$ and $Z \triangleq P_Z|h_{RW}^{\dagger}W_{ZF}|^2$. According to equations (11) and (12) in [40], we have

$$f_{Z_1}(Z) = \frac{Z^{N_R-3}\exp\left(-\sigma_w^2 Z/P_Z\right)}{(N_R-3)!\left(P_Z/\sigma_w^2\right)^{N_R-2}}, \quad N_R \geq 3, Z \geq 0, \tag{23}$$

$$F_{Z_1}(Z) = 1 - \exp\left(-\frac{Z\sigma_W^2}{P_Z}\right)\sum_{l=0}^{N_R-3}\frac{1}{l!}\left(\frac{Z\sigma_W^2}{P_Z}\right)^{l}, \quad Z \geq 0. \tag{24}$$

The SINR at Willie is given by

$$\gamma_{willie2} = \frac{P_{R_k}\left|h_{R_kW}\right|^2}{P_Z\left|h_{RW}^{\dagger}W_{ZF}\right| + \sigma_W^2}, \tag{25}$$

where h_{R_kW} is the channel coefficient between the k-th antenna that Relay chooses to forward the covert messages and Bob and P_Z is the AN power of ZFB sent by Relay. In a communication slot of the second stage, Willie has to decide whether Relay has forwarded covert messages to Bob [42]. Therefore, Willie is faced with a binary hypothesis testing problem again, where the zero hypothesis M_0 means Relay has not forwarded covert messages, and the alternative hypothesis M_1 means Relay has forwarded covert messages to Bob. Based on these assumptions, Willie receives signals is given by

$$y_{willie2}[i] = \begin{cases} \sqrt{P_Z}\left|h_{RW}^{\dagger}W_{ZF}\right|V_{R_2}[i] + n_W[i], & M0, \\ \sqrt{P_{R_k}}\left|h_{R_kW}\right|x_R[i] + \sqrt{P_Z}\left|h_{RW}^{\dagger}W_{ZF}\right|V_{R_2}[i] + n_W[i], & M1. \end{cases} \tag{26}$$

Willie does not know the value of P_Z in the time slot, but the value of P_{R_K} is known. In the second stage, Willie attempts to detect whether $y_{willie2}$ is M_0 or M_1. Through the application of Neyman–Perason criterion and likelihood ratio test, the optimal decision rule for Willie to minimize his detection error is as follows:

$$T_{W2} \underset{U_1}{\overset{U_0}{\gtrless}} \tau, \tag{27}$$

where $T_{W2} = 1/n\sum_{i=1}^{n} y_{Willie2}[i]$ is the average power received at Willie in the slot, τ is a predefined Willie's detection threshold, and U_0 and U_1 are binary decisions that infer whether Relay forwards covert messages, respectively. In this paper, we consider an infinite number of channel uses, that is, $n \longrightarrow \infty$. Therefore, we have

$$T_{W2} = \begin{cases} P_Z\left|h_{RW}^{\dagger}W_{ZF}\right|^2 + \sigma_W^2, & U0, \\ P_{R_k}\left|h_{R_kW}\right|^2 + P_Z\left|h_{RW}^{\dagger}W_{ZF}\right|^2 + \sigma_W^2, & U1. \end{cases} \tag{28}$$

At the end of the communication slot, Willie has to make a decision. The false alarm rate is defined as the probability of Willie making U_1 decision under condition M_0, which is expressed by $\alpha_2 = P(U_1|M_0)$. Similarly, the miss detection rate is defined as the probability of Willie making U_0 decision under condition M_1, which is expressed by $\beta_2 = P(U_0|M_1)$. Assuming that the prior probabilities of M_0 and M_1 are equal, Willie's detection performance can be judged by detecting error rate, which can be defined as follows:

$$\xi_2 \triangleq \alpha_2 + \beta_2. \tag{29}$$

4.3. Detection Performance at Willie

4.3.1. False Alarm Rate and Miss Detection Rate

Theorem 4. *In the second stage, Willie's false alarm rate is derived as follows:*

$$\alpha_2 = \begin{cases} 1, & \tau \le \sigma_W^2, \\ xEi(-x) + \exp(-x), & \tau > \sigma_W^2 N_R = 3, \\ \exp(-x), & \tau > \sigma_W^2 N_R = 4, \\ \left\{1 + \sum_{l=2}^{N_R-3}\left[\frac{1}{l(l-1)}\sum_{k=0}^{l-2}\frac{x^{k+1}}{k!}\right]\right\}\exp(-x), & \tau > \sigma_W^2 N_R \ge 5, \end{cases} \tag{30}$$

and the miss detection rate is formula (31) where $x = (\tau - \sigma_w^2 / (N_R - 1)P_{R_j}^{\max})$, $y = (\tau - \sigma_w^2 - P_{R_k}|h_{R_kW}|^2/(N_R - 1)P_{R_j}^{\max})$.

The $P_{R_j}^{\max}$ is the maximum AN power transmitted by Relay's single antenna, which has been introduced in Section 3.1.

$$\beta_2 = \begin{cases} 0, & \tau \le \sigma_W^2 + P_{R_k}\left|h_{R_kW}\right|^2, \\ 1 - [yEi(-y) + \exp(-y)], & \tau > \sigma_W^2 + P_{R_k}\left|h_{R_kW}\right|^2 N_R = 3, \\ 1 - \exp(-y), & \tau > \sigma_W^2 + P_{R_k}\left|h_{R_kW}\right|^2 N_R = 4, \\ 1 - \left\{1 + \sum_{l=2}^{N_R-3}\left[\frac{1}{l(l-1)}\sum_{k=0}^{l-2}\frac{y^{k+1}}{k!}\right]\right\}\exp(-y), & \tau > \sigma_W^2 + P_{R_k}\left|h_{R_kW}\right|^2 N_R \ge 5. \end{cases} \tag{31}$$

Proof. see Appendix D. □

4.3.2. Optimal Detection Threshold and Minimum Detection Error Rate

Theorem 5. *According to Willie's hypothesis, the optimal detection threshold can be expressed as follows:*

$$\tau^* = \sigma_W^2 + P_{R_KW}\left|h_{R_KW}\right|^2. \tag{32}$$

The corresponding minimum detection error rate is derived as follows:

$$\xi_2^* = \begin{cases} x^*Ei(-x^*) + \exp(-x^*), & N_R = 3, \\ \exp(-x^*), & N_R = 4, \\ \left\{1 + \sum_{l=2}^{N_R-3}\left[\frac{1}{l(l-1)}\sum_{k=0}^{l-2}\frac{x^{*k+1}}{k!}\right]\right\}\exp(-x^*), & N_R \ge 5, \end{cases} \tag{33}$$

where

$$x^* = \frac{P_{R_k}\left|h_{R_kW}\right|^2}{(N_R - 1)P_{R_j}^{\max}\lambda_{RW}}. \tag{34}$$

Proof. see Appendix E. □

Remark 3. The comments are as follows:

(3.1) In the second stage, the minimum detection error rate ξ_2^* is a monotonic increasing function of the maximum AN power $P_{R_j}^{\max}$ of an antenna in Relay using ZFB method, that is, the larger $P_{R_j}^{\max}$ is, the larger ξ_2^* is. The minimum detection error rate ξ_2^* is also monotonically increasing with respect to the number of antennas N_R, that is, the larger N_R is, the larger ξ_2^* is.

(3.2) In the second stage, the minimum detection error rate ξ_2^* is a monotonic decreasing function of Relay forwarding covert message power P_{R_k}, that is, the larger P_{R_k} is, the smaller ξ_2^* is.

4.4. Covert Performance. In the second stage, the maximized effective covert rate can be expressed as follows:

$$\begin{aligned} &\arg\max_{P_{R_k}} && \overline{R_C}, \\ &s.t. && \xi_2^* \ge 1 - \varepsilon. \end{aligned} \tag{35}$$

4.4.1. The Transmission Outage Probability of Relay to Bob. Assuming that the transmission rate R is known, according to formula (19), $|h_{R_kB}|^2$ and P_{R_K} are random variables, which can still cause transmission interruption.

Since the wireless channel is subject to independent quasi-static Rayleigh fading and independent identically distributed, the ($C\ DF$) of $\max_{1\le i\le N_R}(|h_{R_kB}|^2)$ is $F_{\max_{1\le k\le N_R}(|h_{R_kB}|^2)}(x) = (1 - e^{-x})^{N_R}$; then its PDF is

$$f_{\max_{1\le k\le N_R}\left(\left|h_{R_kB}\right|^2\right)}(x) = \sum_{m=0}^{N_R}\binom{N_R}{m}(-1)^m\left(e^{-x}\right)^m. \tag{36}$$

Theorem 6. *For the second stage, the transmission outage probability from Relay to Bob is derived as follows:*

$$\delta = \left[1 - \exp\left(-\frac{(2^{2R}-1)\sigma_B^2}{P_{R_k}}\right)\right]^{N_R}. \tag{37}$$

Proof. see Appendix F. □

Remark 4. The comments are as follows:

(4.1) In the second stage, the larger the power P_{R_k} of Relay forwarding Alice covert messages is, the smaller the transmission outage probability δ_2 is

(4.2) In the second stage, the transmission outage probability δ_2 is a monotonically increasing function of σ_B^2 and R, that is, the larger σ_B^2 and R are, the greater δ_2 is

(4.3) In the second stage, the transmission outage probability δ_2 decreases with the increase of the number of antennas N_R, that is, the larger N_R is, the smaller δ_2 is

4.4.2. Optimal Transmission Power of Relay Forwarding Covert Messages

Theorem 7. *Under any given ε and transmission rate R, the optimal transmission power of Relay forwarding Alice's covert message is expressed as follows:*

$$P_{R_k}^* = \begin{cases} \dfrac{(N_R-1)P_{R_j}^{\max} t_{\varepsilon_1}}{\left|h_{R_k W}\right|^2}, & N_R = 3, \\ \dfrac{(N_R-1)P_{R_j}^{\max}}{\left|h_{R_k W}\right|^2}\ln\left(\dfrac{1}{1-\varepsilon}\right), & N_R = 4, \\ \dfrac{(N_R-1)P_{R_j}^{\max} t_{\varepsilon_2}}{\left|h_{R_k W}\right|^2}, & N_R \geq 5, \end{cases} \tag{38}$$

where t_{ε_1} is the solution of $xEi(-x) + \exp(-x) = 1 - \varepsilon$ and t_{ε_2} is the solution of $\left\{1 + \sum_{l=2}^{N_R-3}[1/l(l-1)\sum_{k=0}^{l-2}x^{k+1}/k!]\right\}\exp(-x) = 1 - \varepsilon$.

Proof:. see Appendix G. □

4.4.3. Optimal Effective Convert Rate

Theorem 8. *In the second stage, the optimal effective convert rate is derived as follows:*

$$\overline{R_{C_2}}^* = R - R\left\{1 - \exp\left[-\frac{(2^{2R}-1)\sigma_B^2}{P_{R_k}^*}\right]\right\}^{N_R}. \tag{39}$$

Proof. see Appendix H. □

Remark 5. The comments are as follows:

(5.1) In the second stage, the optimal effective convert rate $\overline{R_{C_2}^*}$ is a monotonic increasing function of the AN power $P_{R_j}^{\max}$, that is, the larger $P_{R_j}^{\max}$ is, the larger $\overline{R_{C_2}^*}$ is

(5.2) In the second stage, the optimal effective convert rate $\overline{R_{C_2}^*}$ increases with the increase of antenna number N_R and ε, that is, the larger N_R and ε, the greater $\overline{R_{C_2}^*}$

(5.3) In the second stage, the optimal effective convert rate $\overline{R_{C_2}^*}$ increases with the increase of the convert message power $P_{R_k}^*$ forwarded by Relay, and it decreases with the increase of channel noise σ_B^2 and channel coefficient $|h_{R_k W}|^2$

Corollary 2. *In the second stage, if the optimal power of Relay forwarding Alice's covert messages is increased, the maximum effective covert rate approaches a fixed value R.*

$$\lim_{P_{R_k}^* \longrightarrow \infty} \overline{R_{C_2}^*} = R. \tag{40}$$

Proof. according to formula (39), when $P_{R_k}^*$ tends to ∞, the $\exp\left\{-[(2^{2R}-1)\sigma_B^2]/P_{R_k}^*\right\}$ tends to 1; then $\{1 - \exp[-(2^{2R}-1)\sigma_B^2/P_{R_k}^*]\}^{N_R} \longrightarrow 0$, so $\overline{R_{C_2}^*} \longrightarrow R$. □

5. Total Performance

5.1. Transmission Outage Probability

Theorem 9. *The total transmission outage probability is formula (41).*

$$\delta = 1 + \frac{P_A}{P_{R_j}^{\max}\varphi\left(2^{2R}-1\right)}\sum_{m=1}^{N_R}\left\{\binom{N_R}{m}(-1)^m\frac{1}{m}\exp\left[-\frac{m\left(2^{2R}-1\right)\sigma_R^2}{P_A}\right]\ln\left[1 + m\varphi\left(2^{2R}-1\right)\frac{P_{R_j}^{\max}}{P_A}\right]\right\} \tag{41}$$

$$\times\left\{1 - \left[1 - \exp\left(-\frac{(2^{2R}-1)\sigma_B^2}{P_{R_k}}\right)\right]^{N_R}\right\}.$$

Proof. see Appendix I. □

5.2. Maximum Effective Covert Rate

Theorem 10. *The $\sigma_R^2 = \sigma_B^2, P_A = P_{R_k}^*$ is assumed; then the overall optimal effective covert rate is given by*

$$\overline{R_C^*} = \overline{R_{C_1}^*} = -R\left|h_{R_jW}\right|^2 t_\varepsilon \sum_{m=1}^{N_R} \binom{N_R}{m} (-1)^m, \exp\left[-\frac{m\left(2^{2R}-1\right)\sigma_R^2}{P_A}\right] \frac{\ln\left[1 + m\varphi\left(2^{2R}-1\right)\left(1-t_\varepsilon\right)\left|h_{AW}\right|^2 / t_\varepsilon \left|h_{R_jW}\right|^2\right]}{m\varphi\left(2^{2R}-1\right)\left(1-t_\varepsilon\right)\left|h_{AW}\right|^2}. \tag{42}$$

Proof. see Appendix J. □

6. Numerical Result

6.1. The First Stage. In this section, we present some detailed numerical results to illustrate the influence of system parameters on the detection and covert performance of the system. By observing the results in these figures, it can be found that the analysis results are completely consistent with the derivation results, which verifies the correctness of the theoretical analysis. If there are no special cases, we will set some general parameter values. Due to the channel does not change in a time slot, that is, it remains a constant, the value of the channel coefficient can be set to 0 dB. At the same time, the value of channel noise is set to 0 dB. The transmitting power of a single antenna (including nodes with a single antenna) is set to 10 dB. We define the predetermined covert communication condition value as 0.05 and determine the transmission rate as 1 or 1.5. Specific simulation data can be referred to in the instructions below each figure.

6.1.1. Transmission Outage Probability. Figures 2–4 illustrate the relationship between the transmission outage probability δ_1 and the maximum power $P_{R_j}^{\max}$ of a single antenna randomly selected by Relay based on different parameter values. It can be observed that the larger $P_{R_j}^{\max}$ is, the greater δ_1 is. This is because Relay will also generate self-interference while sending interference to Willie. Such self-interference will affect the communication between Alice and the i-th antenna of Relay, and the transmission is more likely to be interrupted. This is consistent with Remark 1.1. In Figure 2, it can be observed that the transmission outage probability δ_1 is a monotonic increasing function with respect to σ_R^2. This is because the larger the channel noise of Relay is, the smaller the SINR of the channel is, and the easier the transmission is interrupted. It can also be seen from Figure 2 that δ_1 is a monotonic increasing function of R, that is, the higher the transmission rate is, the easier the transmission is interrupted. These observations are consistent with Remark 1.2. In Figure 3, with the increase of the number of antennas, the transmission outage probability δ_1 is smaller. This is because the more the number of antennas is configured, the more security gain can be obtained, so that the more secure the transmission is, the less likely the transmission is interrupted. This confirms the correctness of Remark 1.3. In Figure 4, we can see that δ_1 is a monotonic increasing function of ϕ, that is, the larger the interference cancellation coefficient is, the easier the transmission is interrupted. This is because the self-interference coefficient determines the degree of self-interference. The greater the self-interference is, the less secure the transmission from Alice to Relay is. At the same time, it can be observed from Figure 4 that δ_1 is a monotonic decreasing function of P_A. The higher the power of hidden message transmission, the more favorable it is for Relay's reception, and the lower the transmission outage probability. This is consistent with Remark 1.4. In addition, since the expected detection error rate is cited in [2], it can be seen from Figure 3 of [2]: (1) the expected detection error rate $\overline{\xi^*}$ is a monotone increasing function of $P_{R_j}^{\max}$ and (2) the expected detection error rate $\overline{\xi^*}$ is a monotonic decreasing function of P_A.

6.1.2. Maximum Effective Convert Rate. Figures 5–7 shows the probability simulation curves of the maximum effective convert rate $\overline{R_{C_1}^*}$ under different power P_A that Alice sends covert messages. It can be observed that $\overline{R_{C_1}^*}$ increases with the increase of P_A. This is because the larger the transmission power of the covert message, the larger the SINR of the antenna used to receive the covert message at Relay, and the smaller the transmission outage probability in the first stage, so that $\overline{R_{C_1}^*}$ will be larger. It can be observed from Figure 5 that the larger ε is, the larger the maximum effective covert rate is. This is because the larger ε is and the smaller $1-\varepsilon$ is, the lower Willie's minimum detection error rate is required, and the constraint condition of the maximum effective covert rate is more relaxed. Therefore, the maximum effective covert rate $\overline{R_{C_1}^*}$ is a monotonic increasing function of ε, which is consistent with the content of Remark 2.1. It can also be observed from Figure 5 that with the increase of the antenna's number N_R, the maximum effective covert rate increases. The reason is that the more the number of antennas, the smaller the transmission outage probability, but the larger the $\overline{R_{C_1}^*}$. This is consistent with Remark 2.2. It can be seen from Figure 6 that the maximum effective covert rate $\overline{R_{C_1}^*}$ is a monotonic decreasing function of σ_R^2 and ϕ. The larger the channel noise of Relay, the smaller the SINR of Relay, the easier the transmission is interrupted, and the smaller the maximum effective covert rate will be. The larger the self-interference cancellation coefficient ϕ, the greater the self-interference to Relay, and the greater the influence

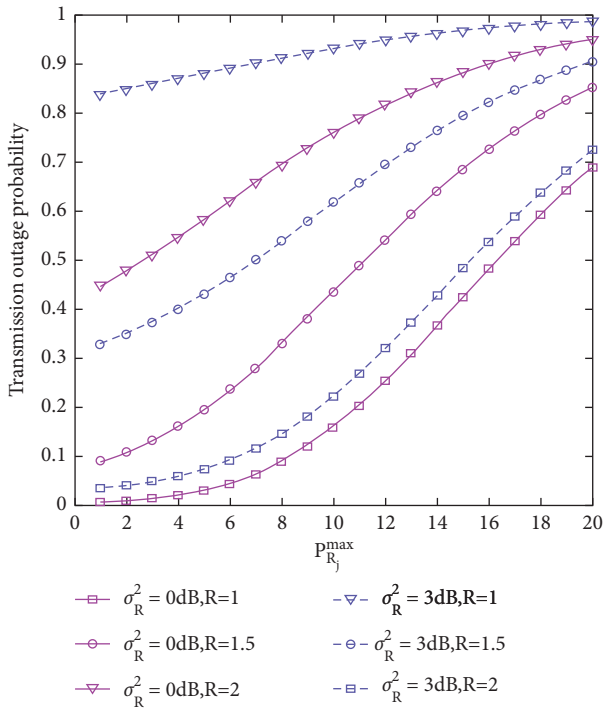


FIGURE 2: Transmission outage probability versus $P_{R_j}^{max}$ with different value of σ_R^2 and R, where $P_A = 10\ dB$, $N_R = 5$, and $\phi = 0.5$.

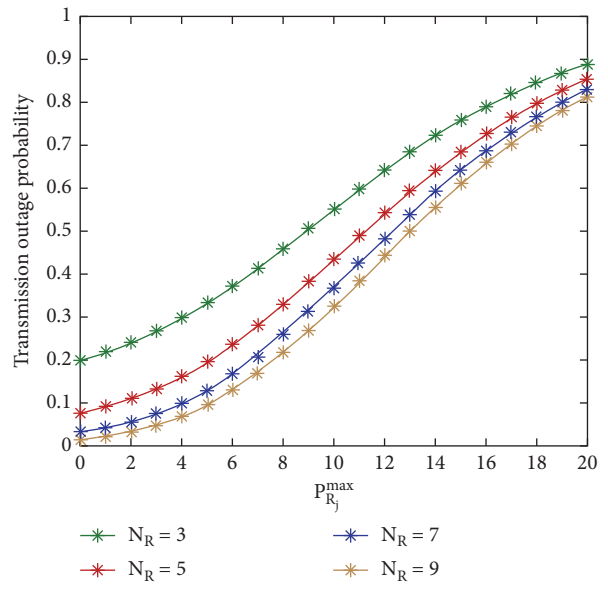


FIGURE 3: Transmission outage probability versus $P_{R_j}^{max}$ with different value of N_R, where $P_A = 10\ dB$, $\sigma_R^2 = 0\ dB$, $R = 1.5$, and $\phi = 0.5$.

on the covert transmission from Alice to Relay, so that the maximum effective covert rate $\overline{R_{C_1}^*}$ is smaller. This proves the correctness of inference Remark 2.3. Figure 7 illustrates the relationship between $|h_{AW}|^2$ and $|h_{R_jW}|^2$ and the maximum effective covert rate $\overline{R_{C_1}^*}$. When Alice sends covert messages, the larger the channel coefficient $|h_{AW}|^2$ between Alice and Willie is, the smaller the maximum effective covert rate $\overline{R_{C_1}^*}$ is. Because the larger the channel gain is, the more favorable

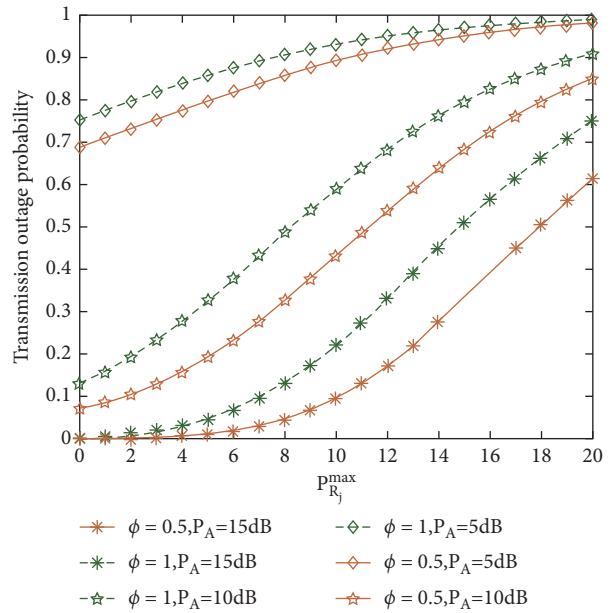


FIGURE 4: Transmission outage probability versus $P_{R_j}^{max}$ with different value of P_A and ϕ, where $N_R = 5$, $\sigma_R^2 = 0\ dB$, and $R = 1.5$.

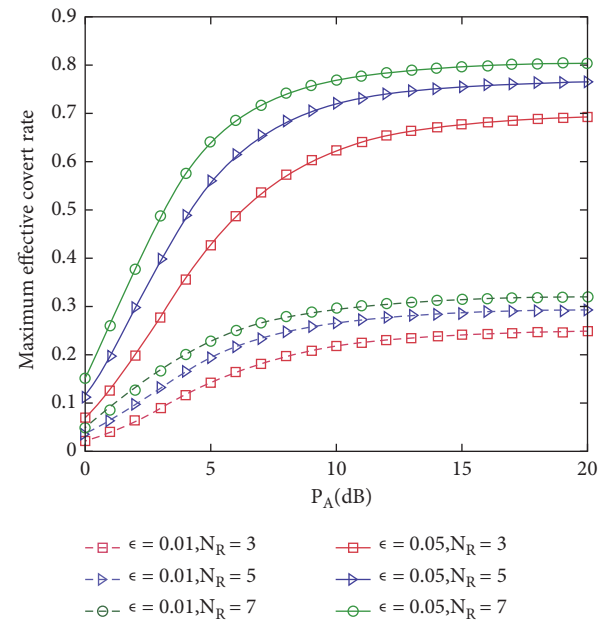


FIGURE 5: Maximum effective covert rate versus P_A with different value of ε and N_R, where $N_R = 5$, $\sigma_R^2 = 0\ dB$, $R = 1$, $|h_{AW}|^2 = 1$, $|h_{R_jW}|^2 = 1$, and $\phi = 0.01$.

Willie's detection is, the smaller the maximum effective covert rate is. And it can be observed that the larger the channel coefficient $|h_{R_jW}|^2$ between Relay randomly selected single antenna and Willie, the greater the maximum effective covert rate. The reason is that the more the channel gain of transmit interference is, the more difficult Willie is to detect, and the easier Willie is to perform covertness. This is consistent with Remark 2.4. In addition, a common point can be seen from Figures 5–7, that is, when increased to a certain value, the maximum effective covert rate of the first

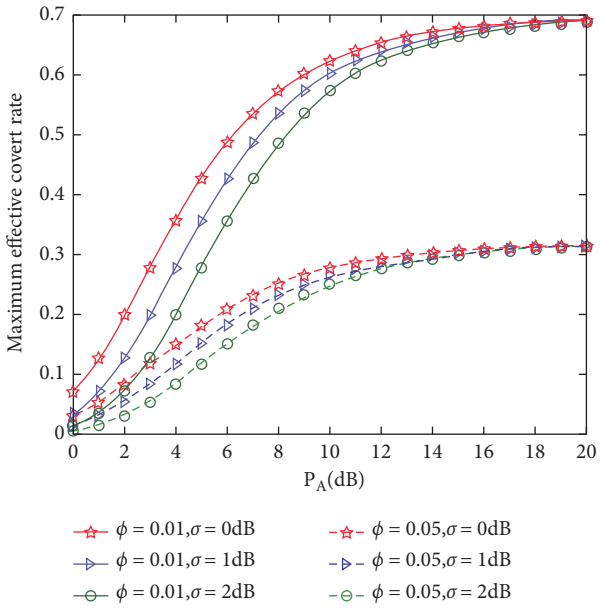


FIGURE 6: Maximum effective covert rate versus P_A with different value of σ_R^2 and ϕ, where $R = 1$, $|h_{AW}|^2 = 1$, $|h_{R_jW}|^2 = 1$, $\varepsilon = 0.05$, and $N_R = 3$.

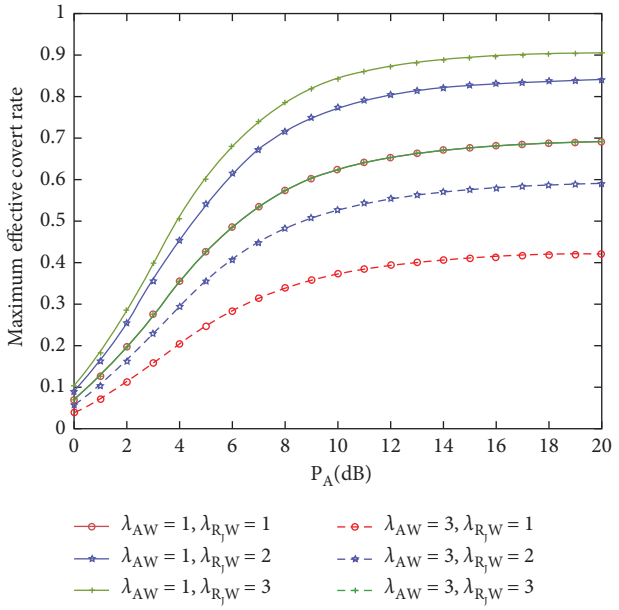


FIGURE 7: Maximum effective covert rate versus P_A with different value of $|h_{AW}|^2$ and $|h_{R_jW}|^2$, where $R = 1$, $\sigma_R^2 = 0\ dB$, $\phi = 0.01$, $\varepsilon = 0.05$, and $N_R = 3$.

stage tends to be a fixed value, which confirms the correctness of Corollary 1.

6.2. The Second Stage

6.2.1. Detection Error Probability. Figure 8 shows the relationship between the false alarm rate α_2, missed detection rate β_2, detection error probability ξ_2, and the threshold τ. It can be observed that the smaller the false alarm rate, the larger the missed detection rate, the smaller the missed

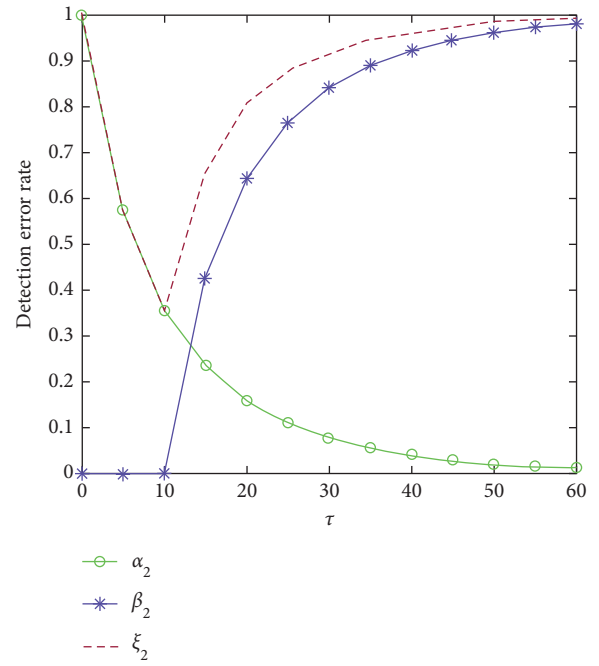


FIGURE 8: Minimum detection error rate, false alarm rate, and missed detection rate versus τ, where $\sigma_B^2 = 0\ dB$, $N_R = 3$, $P_{R_j}^{\max} = P_{R_k} = 10\ dB$, and $|h_{R_kW}|^2 = 1$.

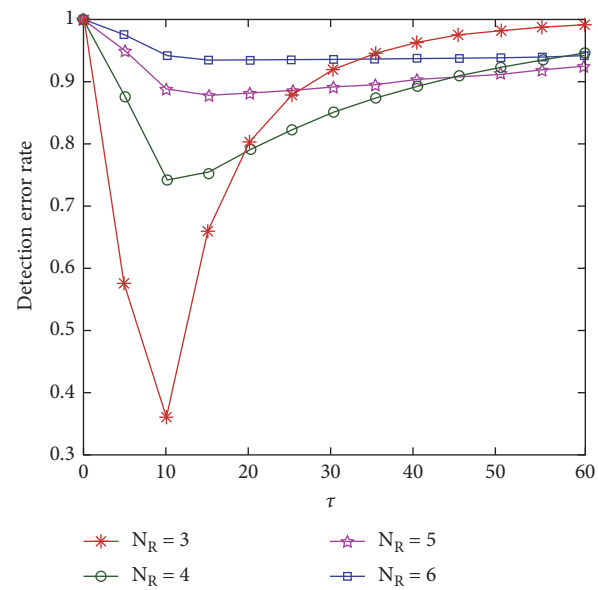


FIGURE 9: Minimum detection error rate versus τ with different value of N_R, where $\sigma_B^2 = 0\ dB$, $P_{R_j}^{\max} = P_{R_k} = 10\ dB$, and $|h_{R_kW}|^2 = 1$.

detection rate, and the larger the false alarm rate. It is difficult for Willie to detect the covert transmission with probability 1, which is consistent with our expectation, and there is a minimum detection error probability; Figure 9 draws the detection error rate curve according to different antenna number n and further illustrates the feasibility of minimum detection error rate under different parameters. This verifies the correctness of Theorem 5.

In Figures 10 and 11, the relationship between the minimum detection error probability ξ_2 and the maximum interference transmission power $P_{R_j}^{\max}$ of a single antenna formed by Relay using zero-forcing beamforming can be observed at the same time. The larger the $P_{R_j}^{\max}$ is, the greater the ξ_2 is. And Figure 10 also describes the relationship between the minimum detection error probability and the number of antennas N_R. It is easy to understand that the minimum detection error rate will increase with the increase in the number of antennas. This is because the larger the interference power and the number of antennas of a single antenna, the greater the total interference to Willie, and the greater the interference will hinder Willie's detection. Willie will be more prone to make errors, resulting in a higher error rate of Willie. This is consistent with the conclusion of Remark 3.1. In addition to describing the relationship between ξ_2 and $P_{R_j}^{\max}$, Figure 11 also plots the relationship between the minimum detection error probability ξ_2 and the power P_{R_K} of Relay forwarding Alice covert messages. The larger the P_{R_K} is, the smaller the ξ_2 is. This is because the higher the power of Relay forwarding covert messages is, the easier Willie is to detect the transmission of covert messages, and the smaller the probability of error is. This verifies the correctness of Remark 3.2.

6.2.2. Transmission Outage Probability. It can be seen from Figures 12 and 13 that the transmission outage probability δ_2 is a monotonic decreasing function of Relay covert message transmission power P_{R_K}, that is, the larger the covert message transmission power P_{R_K} is, the smaller the transmission interruption probability δ_2 of the second stage is. This is because increasing P_{R_K} will increase the SINR at Bob, so the more secure the transmission is, the less likely it is to be interrupted. This verifies the correctness of Remark 4.1. The relationship between the transmission outage probability δ_2 and Bob's channel noise σ_B^2 is also described in Figure 12. δ_2 increases with the increase of σ_B^2. This is because the increase of channel noise e will reduce the signal-to–interference-noise ratio at Bob, which makes the transmission easier to be interrupted. It can also be observed from Figure 12 that δ_2 is a monotonic increasing function of R, that is, the higher the transmission rate is, the easier the transmission is interrupted. These observations are consistent with Remark 4.2. It can be observed from Figure 13 that the transmission outage probability δ_2 is a monotonic decreasing function of the number of Relay antennas N_R, that is, the larger N_R is, the smaller the transmission outage probability δ_2 of the second stage is. This is because the increase in the number of antennas will increase the gain of the channel, thus making the transmission more secure. This verifies the correctness of Remark 4.3.

6.2.3. Maximum Effective Covert Rate. It can be seen from Figures 14 and 15 that the optimal effective covert rate $\overline{R_{C_2}^*}$ in the second stage is a monotonic increasing function of the interference transmission power $P_{R_j}^{\max}$ of a single Relay's antenna, that is, the greater the interference transmission power $P_{R_j}^{\max}$ of a single Relay's antenna, the greater the optimal effective covert rate $\overline{R_{C_2}^*}$ in the second stage. This is because the greater the interference, the more likely Willie's detection is to make mistakes, and the higher the effective covert rate is. This is consistent with Remark 5.1. Similar to the first stage, the optimal effective covert rate $\overline{R_{C_2}^*}$ in the second stage increases with the increase of the number of antennas N_R and ε. The reason has been described in the first stage, and it is not described here. This is consistent with the content of Remark 5.2.

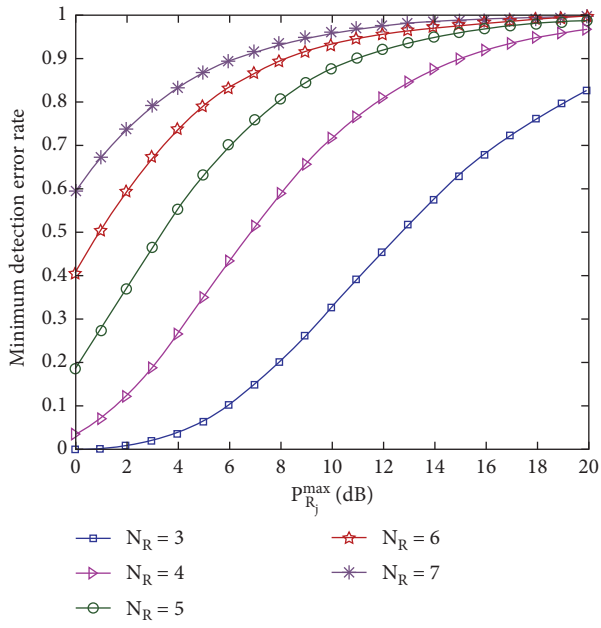


Figure 10: Minimum detection error rate versus $P_{R_j}^{\max}$ with different value of N_R, where $P_{R_k} = 10\ dB$ and $|h_{R_kW}|^2 = 1$.

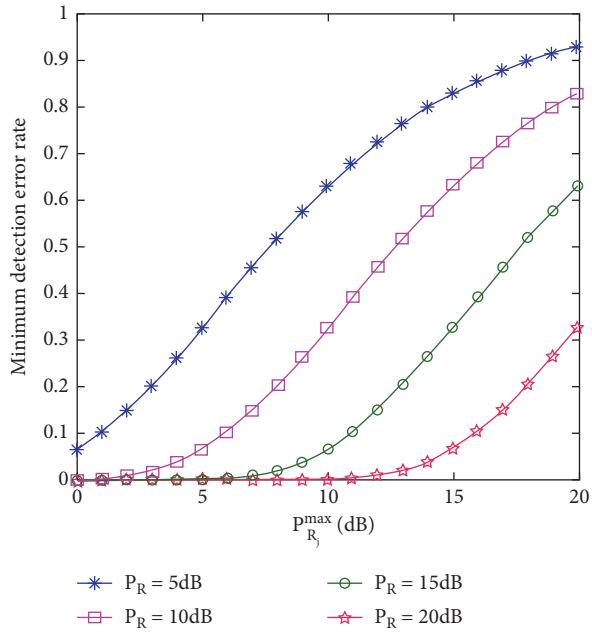


Figure 11: Minimum detection error rate versus $P_{R_j}^{\max}$ with different value of $P_{R_k} = 10\ dB$, where $N_R = 3$ and $|h_{R_kW}|^2 = 1$.

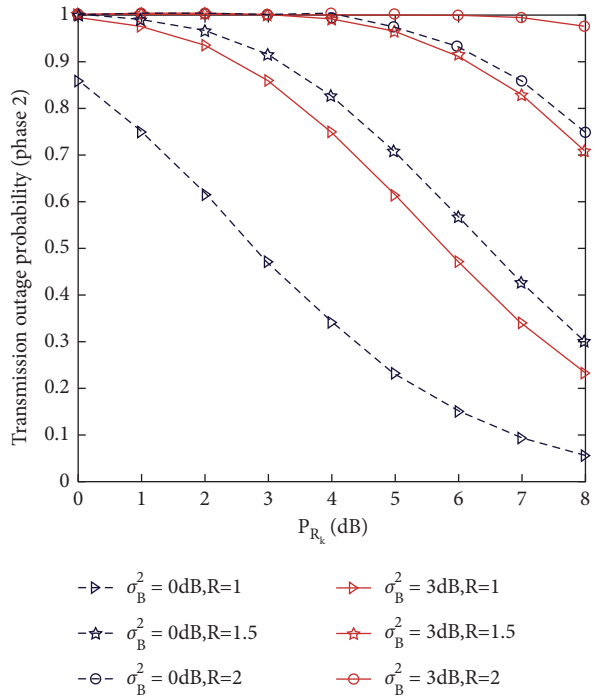


FIGURE 12: Transmission outage probability versus P_{R_k} with different value of σ_B^2 and R, where $N_R = 3$.

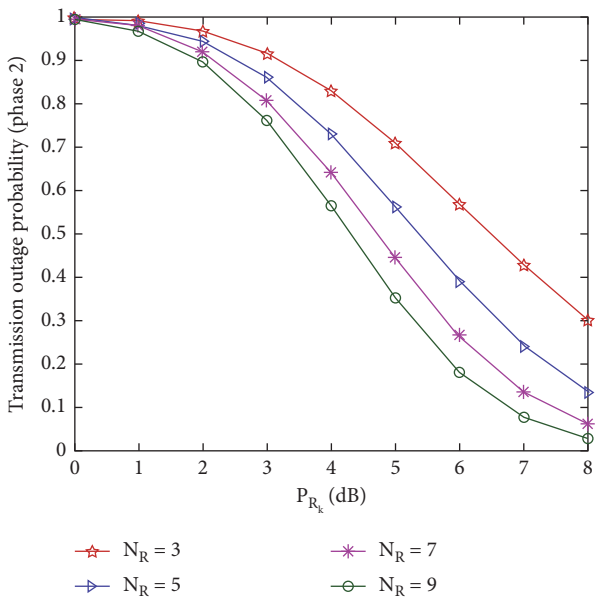


FIGURE 13: Transmission outage probability versus P_{R_k} with different value of N_R, where $R = 1.5$ and $\sigma_B^2 = 0$ dB.

Figure 16 shows the relationship between the optimal effective covert rate $\overline{R_{C_2}^*}$ and the power P_{R_K} of Relay forwarding covert messages and the noise σ_B^2 of Bob channel. The higher the power P_{R_K} of Relay forwarding covert messages, the higher the $\overline{R_{C_2}^*}$. The larger the noise σ_B^2 of Bob channel is, the smaller the $\overline{R_{C_2}^*}$ is. This is because the larger the P_{R_K} is, the larger Bob's SINR is, which is the more conducive to the transmission of covert messages, the larger the σ_B^2 is, the smaller the SINR of Bob will

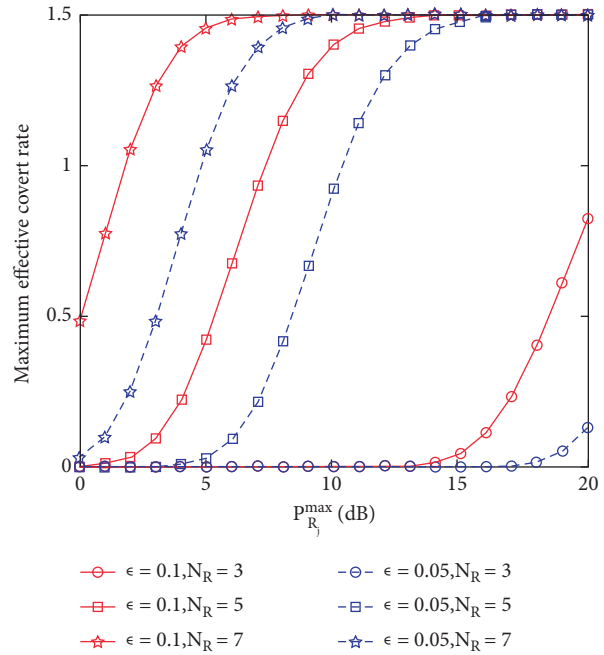


FIGURE 14: Maximum effective covert rate versus $P_{R_j}^{\max}$ with different value of N_R and ε, where $R = 1.5$, $|h_{R_jW}|^2 = 1$, and $\sigma_B^2 = 0$ dB.

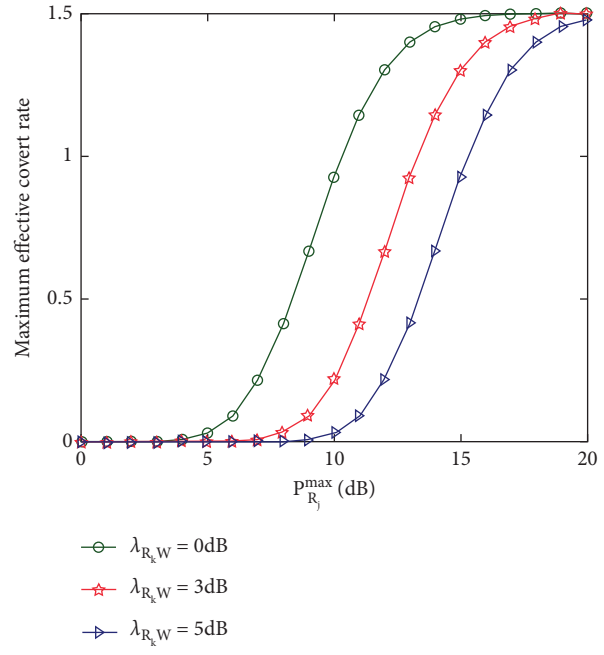


FIGURE 15: Maximum effective covert rate versus P_{R_k} with different value of σ_B^2, where $R = 1.5$, $|h_{R_jW}|^2 = 1$, $N_R = 5$, and $\varepsilon = 0.05$.

be, which is not conducive to covert message transmission. Figure 15 also describes the relationship between the optimal effective covert rate $\overline{R_{C_2}^*}$ and the channel coefficient $|h_{R_kW}|^2$. The larger the channel coefficient $|h_{R_kW}|^2$ is, the smaller the $\overline{R_{C_2}^*}$ is. Because it is the channel gain between Relay and Willie when Relay forwards Alice's covert messages, it is obvious that the

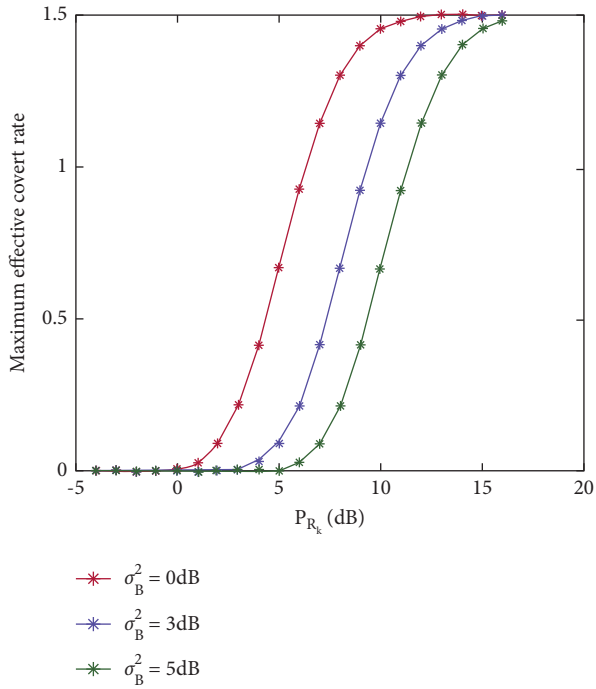


Figure 16: Maximum effective covert rate versus $P_{R_j}^{\max}$ with different value of $|h_{R_jW}|^2$, where $R = 1.5$, $N_R = 5$, $\varepsilon = 0.05$, and $\sigma_B^2 = 0$ dB.

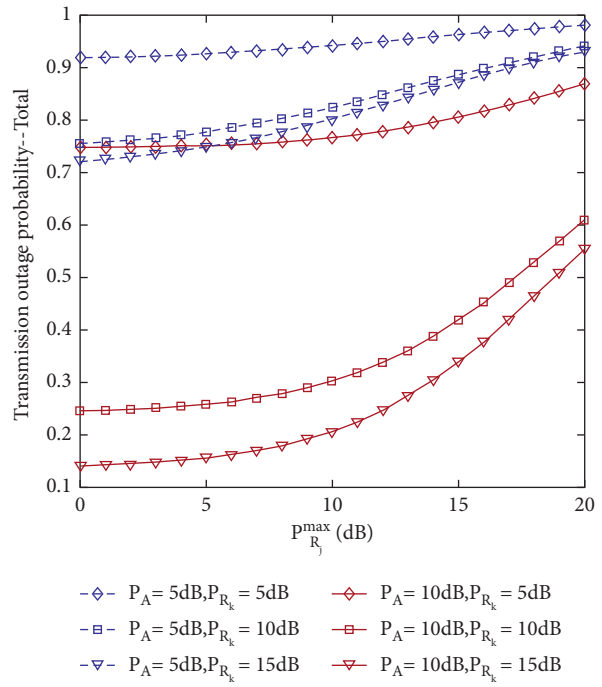


Figure 17: Total transmission outage probability versus $P_{R_j}^{\max}$ with different value of P_{R_k} and P_A, where $\sigma_R^2 = 0$ dB, $\sigma_B^2 = 0$ dB, $R = 1.5$, $N_R = 3$, and $\phi = 0.05$.

larger the coefficient is, the worse the transmission of covert messages will be. This is not only similar to the first stage conclusion but also verifies the correctness of Remark 5.3. In addition, Figures 14–16 shows that the effective covert rate of the second stage tends to a fixed value, which verifies the correctness of Corollary 2.

6.3. Total Performance

6.3.1. Transmission Outage Probability. As in the first and second stages, we also graphically plot the relevant parameters of the total transmission outage probability, as shown in Figures 17–20. By comprehensively comparing the transmission outage probability in the first and second stages, we can get the contents in Table 2. Among them, the upward front represents the increase of the value, and the downward arrow represents the decrease of the value.

It can be observed from Table 2 that for the same parameters, the first stage, the second stage, and the total transmission outage probability have the same conclusion, which is consistent with our expectation. The causes of these graph phenomena have been described in detail in the previous stages and will not be repeated here.

6.3.2. Maximum Effective Covert Rate. Because the total maximum effective covert rate is smaller in the first and second stages, the maximum effective covert rates of the first and second stages are drawn synthetically to facilitate the comparison. The details are shown in Figure 21. Obviously,

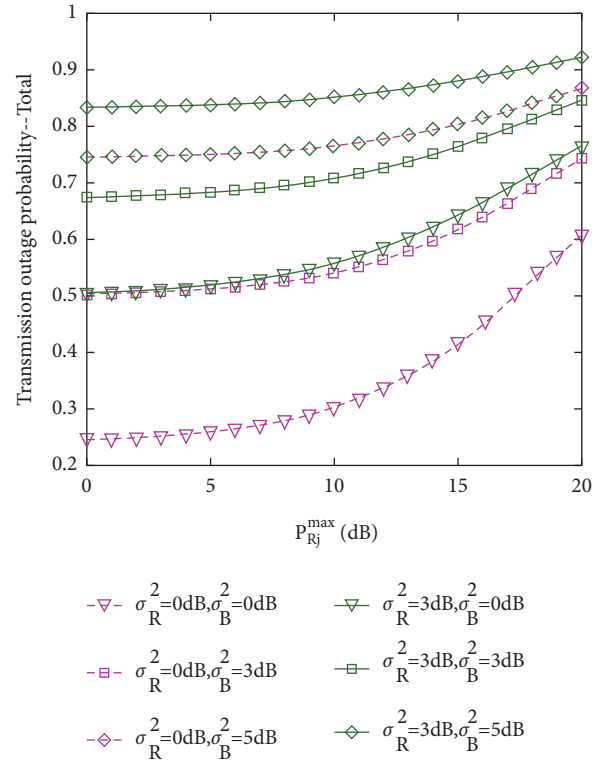


Figure 18: Total transmission outage probability versus $P_{R_j}^{\max}$ with different value of σ_R^2 and σ_B^2, where $P_{R_k} = 10$ dB, $P_A = 10$ dB, $R = 1.5$, $N_R = 3$, and $\phi = 0.05$.

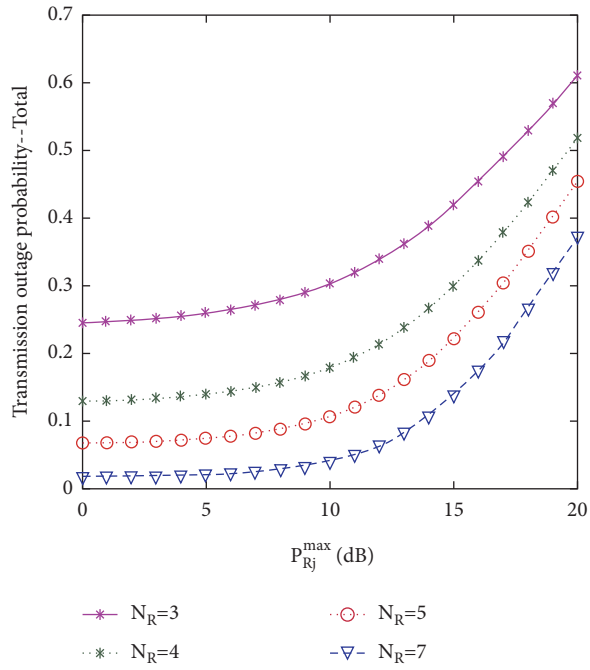


FIGURE 19: Total transmission outage probability versus $P_{R_j}^{\max}$ with different value of N_R, where $P_{R_k} = 10\ dB$, $P_A = 10\ dB$, $R = 1.5$, $\phi = 0.05$, $P_{R_k} = 10\ dB$, and $P_A = 10\ dB$.

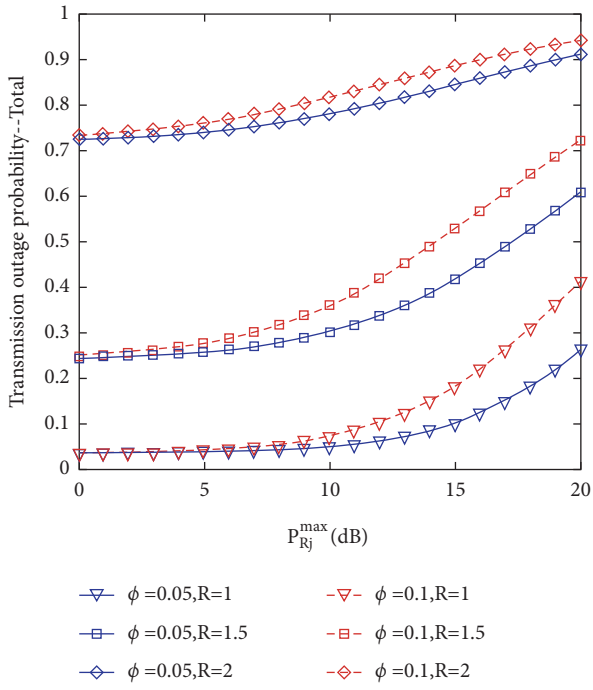


FIGURE 20: Total transmission outage probability versus $P_{R_j}^{\max}$ with different value of ϕ and R, where $P_{R_k} = 10\ dB$, $P_A = 10\ dB$, $N_R = 3$, $\sigma_R^2 = 0\ dB$, and $\sigma_B^2 = 0\ dB$.

the maximum effective covert rate of the first stage is less than that of the second stage.

By comprehensively comparing the maximum effective covert rate in the first and second stages, we can get the contents in Table 3.

It can be seen from Table 3 that for the same parameters, the total maximum effective concealment rate has the same conclusion as the maximum effective convert rate in one stage. This is because zero-forcing beamforming is used in the second stage to interfere with Willie without any impact on Bob, the transmission interruption probability will be smaller, and the corresponding maximum effective convert rate will be larger, so the maximum effective convert rate in the second stage is significantly higher than that in the first stage. This also verifies the correctness of Theorem 10.

Table 2: Transmission outage probability results.

Parameter	Parameter variation	δ_1	δ_2	δ
Interference power	↑	↑	↑	↑
Channel noise from the receiver	↑	↑	↑	↑
Transmission rate	↑	↑	↑	↑
Antennas number	↑	↓	↓	↓
Covert message sending power	↑	↓	↓	↓
Self-interference coefficient	↑	↑	—	↑

Table 3: Maximum effective covert rate results.

Parameter	Parameter variation	$\overline{R^*_{C_1}}$	$\overline{R^*_{C_2}}$	$\overline{R^*_C}$
Covert message sending power	↑	↑	↑	↑
Covert constraints, ε	↑	↑	↑	↑
Antennas number	↑	↑	↑	↑
Channel noise from the receiver	↑	↓	↓	↓
Self-interference coefficient	↑	↓	—	↓
The channel coefficient between covert message sender and Willie	↑	↓	↓	↓
The channel coefficient between the AN sender and Willie	↑	↑	—	↑
Interference power	↑	—	↑	—

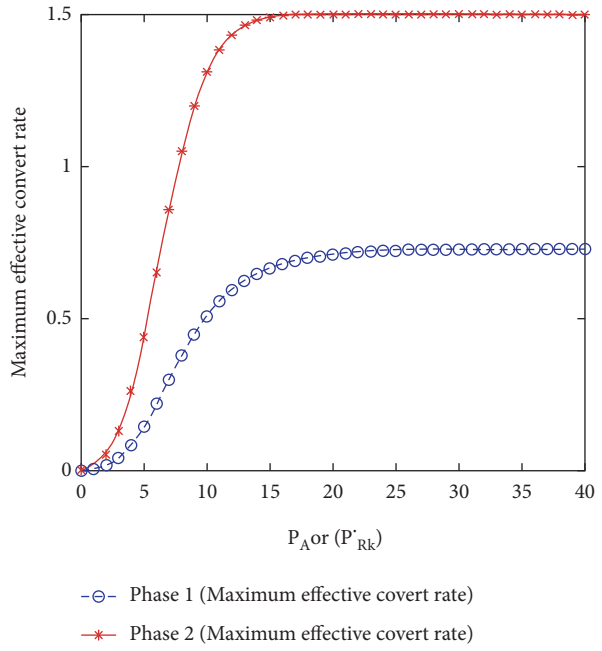


Figure 21: Comparison of stage 1 and stage 2 maximum effective covert rates, where $\sigma_R^2 = 0\ dB$, $\sigma_B^2 = 0\ dB$, $P_{R_k} = 10\ dB$, $R = 1.5$, $\phi = 0.01$, $N_R = 3$, $\varepsilon = 0.05$, $|h_{R_jW}|^2 = 1$, and $h_{AW} = 1$.

7. Conclusion

In this work, we consider a wireless network covert communication system achieved by multiantenna full-duplex relay. The system is divided into two stages; each stage needs to complete the transmission of covert messages. In the first stage, Relay receives covert messages and sends interference to Willie. In the second stage, Relay decodes and forwards covert messages while interfering with Willie's detection. The system must ensure that neither stage can be detected. The two stages and the total system performance are derived and analyzed. Under the constraint of covert transmission, the maximum effective covert rate of multiantenna single relay decode and forward relay network is studied, and the simulation results are analyzed in detail, the system can achieve excellent covert performance. In the future, multiantenna and multirelay wireless network covert communication systems can be further considered.

Appendix

A. Proof of Theorem 1

When the channel capacity C from Alice to Relay is less than the transmission rate R, the transmission from Alice to Relay will be interrupted. The channel capacity formula is as follows:

$$C_{AR} = \frac{1}{2}\log_2(1+\gamma_{AR}). \tag{A.1}$$

In combination with formulas (1) and (A.1) and the definition of transmission interrupt probability, we have

$$\delta_1 = P\left\{\frac{P_A \max_{1\le i\le N_R}\left(\left|h_{AR_i}\right|^2\right)}{\varphi P_{R_j}\left|h_{R_jR_j}\right|^2+\sigma_R^2} < 2^{2R}-1\right\} = \int_0^{P_{R_j}^{\max}}\int_0^{\infty}\left[1-\exp\left(-\frac{\left(2^{2R}-1\right)\left(\varphi P_{R_j}\left|h_{R_jR_j}\right|^2+\sigma_R^2\right)}{P_A}\right)\right]^{N_R}\exp(-y)\frac{1}{P_{R_j}^{\max}}dx\,dy,$$
$$= 1+\frac{P_A}{P_{R_j}^{\max}\varphi\left(2^{2R}-1\right)}\sum_{m=1}^{N_R}\binom{N_R}{m}(-1)^m\frac{1}{m}\exp\left[-\frac{m\left(2^{2R}-1\right)\sigma_R^2}{P_A}\right]\times\ \ln\left[1+m\varphi\left(2^{2R}-1\right)\frac{P_{R_j}^{\max}}{P_A}\right]. \tag{A.2}$$

B. Proof of Theorem 2

According to formulas (11) and (13), t_ε is the solution of $-t^2+t\ln t+1=1-\varepsilon$, and there is the definition of t in formula (11), we have

$$t_\varepsilon = \frac{P_A\left|h_{AW}\right|^2}{P_A\left|h_{AW}\right|^2+P_{R_j}^{\max}\left|h_{R_jW}\right|^2}. \tag{B.1}$$

The proof can be completed by solving $P_{R_j}^{\max}$ in formula (B.1).

C. Proof of Theorem 3

According to formula (13), we have

$$\overline{R_C} = R(1-\delta) = R\begin{cases}1-1-\dfrac{P_A}{P_{R_j}^{\max}\varphi\left(2^{2R}-1\right)}\displaystyle\sum_{m=1}^{N_R}\binom{N_R}{m}(-1)^m\frac{1}{m},\\ \exp\left[-\dfrac{m\left(2^{2R}-1\right)\sigma_R^2}{P_A}\right]\ln\left[1+m\varphi\left(2^{2R}-1\right)\dfrac{P_{R_j}^{\max}}{P_A}\right],\end{cases} \tag{C.1}$$
$$= -\frac{RP_A}{P_{R_j}^{\max}\varphi\left(2^{2R}-1\right)}\sum_{m=1}^{N_R}\binom{N_R}{m}(-1)^m\frac{1}{m}\exp\left[-\frac{m\left(2^{2R}-1\right)\sigma_R^2}{P_A}\right]\times\ \ln\left[1+m\varphi\left(2^{2R}-1\right)\frac{P_{R_j}^{\max}}{P_A}\right].$$

By substituting formula (16) into formula (C.1), we can get the following results:

$$\overline{R_{C_1}^*} = -R\sum_{m=1}^{N_R}\binom{N_R}{m}(-1)^m\exp\left[-\frac{m\left(2^{2R}-1\right)\sigma_R^2}{P_A}\right]\times\frac{\ln\left[1+m\varphi\left(2^{2R}-1\right)(1-t_\varepsilon)\left|h_{AW}\right|^2/t_\varepsilon\left|h_{R_jW}\right|^2\right]}{\left[m\varphi\left(2^{2R}-1\right)(1-t_\varepsilon)\left|h_{AW}\right|^2\right]/t_\varepsilon\left|h_{R_jW}\right|^2},$$
$$= -R\left|h_{R_jW}\right|^2 t_\varepsilon\sum_{m=1}^{N_R}\binom{N_R}{m}(-1)^m\exp\left[-\frac{m\left(2^{2R}-1\right)\sigma_R^2}{P_A}\right]\times\frac{\ln\left[1+m\varphi\left(2^{2R}-1\right)(1-t_\varepsilon)\left|h_{AW}\right|^2/t_\varepsilon\left|h_{R_jW}\right|^2\right]}{m\varphi\left(2^{2R}-1\right)(1-t_\varepsilon)\left|h_{AW}\right|^2}. \tag{C.2}$$

Proof of Theorem 4

According to formulas (27) and (28), the false alarm rate is as follows:

$$\begin{aligned}\alpha_2 &= P\left[P_Z\left|h_{RW}^{\dagger}W_{ZF}\right|^2+\sigma_W^2>\tau\right],\\ &= P\left[P_Z\left|h_{RW}^{\dagger}W_{ZF}\right|^2>\tau-\sigma_W^2\right],\\ &= P\left[\frac{P_Z\left|h_{RW}^{\dagger}W_{ZF}\right|^2}{\sigma_W^2}>\frac{\tau-\sigma_W^2}{\sigma_W^2}\right].\end{aligned} \tag{D.1}$$

By using formula (23), we can get the result (D.2), shown at the top of the next page, where, by using the variable substitution of $P_Z = 1/x$, and with the help of [41], equation (3.351.2.11), and [41], equation (3.351.4), to solve the integral in formula (D.2), the result of formula (30) can be obtained as follows:

$$\alpha_2 = 1-\int_0^{(N_R-1)P_{R_j}^{\max}}\left[1-\exp\left(-\frac{\tau-\sigma_w^2}{P_Z}\right)\sum_{l=0}^{N_R-3}\frac{1}{l!}\left(\frac{\tau-\sigma_w^2}{P_Z}\right)^l\right]\frac{1}{(N_R-1)P_{R_j}^{\max}}dP_Z. \tag{D.2}$$

In the same way, the miss detection rate is as follows:

$$\begin{aligned}\beta_2 &= P\left[P_{R_k}\left|h_{R_kW}\right|^2+P_Z\left|h_{RW}^{\dagger}W_{ZF}\right|^2+\sigma_W^2<\tau\right],\\ &= P\left[P_Z\left|h_{RW}^{\dagger}W_{ZF}\right|^2<\tau-\sigma_W^2-P_{R_k}\left|h_{R_kW}\right|^2\right],\\ &= P\left[\frac{P_Z\left|h_{RW}^{\dagger}W_{ZF}\right|^2}{\sigma_W^2}<\frac{\tau-\sigma_W^2-P_{R_k}\left|h_{R_kW}\right|^2}{\sigma_W^2}\right].\end{aligned} \tag{D.3}$$

By using (23), we can get the result (D.4), where by using the same variable substitution and solving the integral of (D.4) in the same way, the result of formula (31) can be obtained as follows:

$$\beta_2 = \int_0^{(N_R-1)P_{R_j}^{\max}}\left[1-\exp\left(-\frac{\tau-\sigma_w^2-P_{R_kW}\left|h_{R_kW}\right|^2}{P_Z}\right)\sum_{l=0}^{N_R-3}\frac{1}{l!}\left(\frac{\tau-\sigma_w^2-P_{R_kW}\left|h_{R_kW}\right|^2}{P_Z}\right)^l\right]\frac{1}{(N_R-1)P_{R_j}^{\max}}dP_Z. \tag{D.4}$$

E. Proof of Theorem 5

According to formulas (30) and (31), Willie's detection error rate (E.1) is shown at the top of the next page, where

$$x_1 = \frac{\tau - \sigma_w^2}{(N_R - 1)P_{R_j}^{\max}},$$

$$y_1 = \frac{\tau - \sigma_w^2 - P_{R_k}\left|h_{R_kW}\right|^2}{(N_R - 1)P_{R_j}^{\max}},$$

$$\xi_2 = \begin{cases} 1, & \tau \le \sigma_W^2, \\ x_1 Ei(-x_1) + \exp(-x_1), & \sigma_W^2 < \tau \le \sigma_W^2 + P_{R_k}\left|h_{R_kW}\right|^2, \quad N_R = 3, \\ \exp(-x_1), & \sigma_W^2 < \tau \le \sigma_W^2 + P_{R_k}\left|h_{R_kW}\right|^2, \quad N_R = 4, \\ \left\{1 + \sum_{l=2}^{N_R-3}\left[\frac{1}{l(l-1)}\sum_{k=0}^{l-2}\frac{x_1^{k+1}}{k!}\right]\right\}\exp(-x_1), & \sigma_W^2 < \tau \le \sigma_W^2 + P_{R_k}\left|h_{R_kW}\right|^2, \quad N_R \ge 5, \\ 1 + x_1 Ei(-x_1) + \exp(-x_1) - y_1 Ei(-y_1) - \exp(-y_1), & \tau > \sigma_W^2 + P_{R_k}\left|h_{R_kW}\right|^2, \quad N_R = 3, \\ 1 + \exp(-x_1) - \exp(-y_1), & \tau > \sigma_W^2 + P_{R_k}\left|h_{R_kW}\right|^2, \quad N_R = 4, \\ 1 + \left\{1 + \sum_{l=2}^{N_R-3}\left[\frac{1}{l(l-1)}\sum_{k=0}^{l-2}\frac{x_1^{k+1}}{k!}\right]\right\}\exp(-x_1) - \left\{1 + \sum_{l=2}^{N_R-3}\left[\frac{1}{l(l-1)}\sum_{k=0}^{l-2}\frac{x_1^{k+1}}{k!}\right]\right\}\exp(-y_1), & \tau > \sigma_W^2 + P_{R_k}\left|h_{R_kW}\right|^2, \quad N_R \ge 5, \end{cases}$$

$$\xi_2 = \begin{cases} 1, & \tau \le \sigma_W^2, \\ x_1 Ei(-x_1) + \exp(-x_1), & \sigma_W^2 < \tau \le \sigma_W^2 + P_{R_k}\left|h_{R_kW}\right|^2, \quad N_R = 3, \\ \exp(-x_1), & \sigma_W^2 < \tau \le \sigma P_{R_k}\left|h_{R_kW}\right|^2, \quad N_R = 4, \\ \left\{1 + \sum_{l=2}^{N_R-3}\left[\frac{1}{l(l-1)}\sum_{k=0}^{l-2}\frac{x_1^{k+1}}{k!}\right]\right\}\exp(-x_1), & \sigma_W^2 < \tau \le \sigma_W^2 + P_{R_k}\left|h_{R_kW}\right|^2, \quad N_R \ge 5, \\ 1 + x_1 Ei(-x_1) + \exp(-x_1) - y_1 Ei(-y_1) - \exp(-y_1), & \tau > \sigma_W^2 + P_{R_k}\left|h_{R_kW}\right|^2, \quad N_R = 3, \\ 1 + \exp(-x_1) - \exp(-y_1), & \tau > \sigma_W^2 + P_{R_k}\left|h_{R_kW}\right|^2, \quad N_R = 4, \\ 1 + \left\{1 + \sum_{l=2}^{N_R-3}\left[\frac{1}{l(l-1)}\sum_{k=0}^{l-2}\frac{x_1^{k+1}}{k!}\right]\right\}\exp(-x_1) - \left\{1 + \sum_{l=2}^{N_R-3}\left[\frac{1}{l(l-1)}\sum_{k=0}^{l-2}\frac{x_1^{k+1}}{k!}\right]\right\}\exp(-y_1), & \tau > \sigma_W^2 + P_{R_k}\left|h_{R_kW}\right|^2, \quad N_R \ge 5. \end{cases} \tag{E.1}$$

(1) When $N_R = 3$, let us define $f(x) = xEi(-x) + \exp(-x)$, then

$$f'(x) = Ei(-x) < 0;\ f''(x) = \frac{e^{-x}}{x} > 0, \quad x > 0. \tag{E.2}$$

Therefore, the function $f(x) = xEi(-x) + \exp(-x)\,(x > 0)$ decreases monotonically, and its first derivative increases monotonically.

When $\sigma_W^2 < \tau \le \sigma_W^2 + P_{R_k}|h_{R_kW}|^2$, we have

$$\frac{d\xi_2}{d\tau} = \frac{df(x_1)}{dx_1} \times \frac{dx_1}{d\tau} = \frac{1}{(N_R - 1)P_{R_j}^{\max}} \frac{df(x_1)}{dx_1} < 0. \tag{E.3}$$

When $\tau > \sigma_W^2 + P_{R_k}|h_{R_kW}|^2$, we have

$$\frac{d\xi_2}{d\tau} = \frac{df(x_1)}{dx_1} \times \frac{dx_1}{d\tau} - \frac{df(y_1)}{dy_1} \times \frac{dy_1}{d\tau} = \frac{1}{(N_R - 1)P_{R_j}^{\max}} \left[\frac{df(x_1)}{dx_1} - \frac{df(y_1)}{dy_1} \right] > 0. \tag{E.4}$$

According to $x_1 > y_1$ and the monotonicity of $f(x) = xEi(-x) + \exp(-x)(x > 0)$, the $\min \xi$ is obtained at $\tau = \sigma_W^2 + P_{R_k}|h_{R_kW}|^2$.

(2) Similarly, when $N_R = 4$, let us define $g(x) = \exp(-x)$; the same derivation as $N_R = 3$ is used. According to $x_1 > y_1$ and the monotonicity of $g(x)$, the $\min \xi$ is obtained at $\tau = \sigma_W^2 + P_{R_k}|h_{R_kW}|^2$.

(3) When $N_R = 5$, the position transformation operation of the lower function $h(x)$ can be obtained as follows:

$$h(x) = \left\{ 1 + \sum_{l=2}^{N_R-3} \left[\frac{1}{l(l-1)} \sum_{k=0}^{l-2} \frac{x^{k+1}}{k!} \right] \right\} \exp(-x) = \left\{ 1 + \sum_{k=1}^{N_R-4} \left[\sum_{l=k+1}^{N_R-3} \frac{1}{l(l-1)} \frac{x^k}{(k-1)!} \right] \right\} \exp(-x). \tag{E.5}$$

The first derivative can be expressed as follows:

$$\begin{aligned} h'(x) &= \left\{ \begin{array}{l} \sum_{k=1}^{N_R-4} \left[\sum_{l=k+1}^{N_R-3} \frac{1}{l(l-1)} \frac{kx^{k-1}}{(k-1)!} \right] - 1 \\ - \sum_{k=1}^{N_R-4} \left[\sum_{l=k+1}^{N_R-3} \frac{1}{l(l-1)} \frac{x^k}{(k-1)!} \right] \end{array} \right\} \exp(-x) \\ &= \left\{ -\frac{1}{N_R - 3} - \frac{1}{N_R - 3} \sum_{k=1}^{N_R-4} \left[\frac{x^k}{k!} \right] \right\} \exp(-x) < 0, x > 0. \end{aligned} \tag{E.6}$$

Its second derivative can be expressed as follows:

$$\begin{aligned} h''(x) &= \left\{ \begin{array}{l} -\frac{1}{N_R - 3} \sum_{k=1}^{N_R-4} \left[\frac{x^{k-1}}{(k-1)!} \right] \\ + \frac{1}{N_R - 3} + \frac{1}{N_R - 3} \sum_{k=1}^{N_R-4} \left[\frac{x^k}{k!} \right] \end{array} \right\} \exp(-x) \\ &= \frac{x^{N_R-4}}{(N_R - 3)!} \exp(-x) > 0, x > 0. \end{aligned} \tag{E.7}$$

Therefore, the function $h(x)$ is monotonically decreasing, and the first derivative is monotonically increasing.

When $\sigma_W^2 < \tau \le \sigma_W^2 + P_{R_k}|h_{R_kW}|^2$, we have

$$\frac{d\xi_2}{d\tau} = \frac{dh(x_1)}{dx_1} \times \frac{dx_1}{d\tau} = \frac{1}{(N_R - 1)P_{R_j}^{\max}} \frac{dh(x_1)}{dx_1} < 0. \tag{E.8}$$

When $\tau > \sigma_W^2 + P_{R_k}|h_{R_kW}|^2$, we have

$$\frac{d\xi_2}{d\tau} = \frac{dh(x_1)}{dx_1} \times \frac{dx_1}{d\tau} - \frac{dh(y_1)}{dy_1} \times \frac{dy_1}{d\tau} = \frac{1}{(N_R - 1)P_{R_j}^{\max}} \left[\frac{dh(x_1)}{dx_1} - \frac{dh(y_1)}{dy_1} \right] > 0. \tag{E.9}$$

According to $x_1 > y_1$ and the monotonicity of $h(x)$, the $\min \xi$ is obtained at $\tau = \sigma_W^2 + P_{R_k}|h_{R_kW}|^2$.

In summary, for any $N_R \ge 3$, there are the following conclusions. When $\sigma_W^2 < \tau \le \sigma_W^2 + P_{R_k}|h_{R_kW}|^2$, there is

$d\xi_2/d\tau = d\xi_2/dxdx/d\tau < 0$. When $\tau > \sigma_W^2 + P_{R_k}|h_{R_kW}|^2$, there is $d\xi_2/d\tau = d\xi_2/dxdx/d\tau - d\xi_2/dydy/d\tau > 0$.

Therefore, the optimal detection threshold is $\tau^* = \sigma_w^2 + P_{R_kW}|h_{R_kW}|^2$, which can be substituted into formula (E.1) to obtain formula (33).

F. Proof of Theorem 6

When the channel capacity C from Alice to Relay is less than the transmission rate R, the transmission from Relay to Bob will be interrupted. The channel capacity formula is as follows:

$$C_{RB} = \frac{1}{2}\log_2(1 + \gamma_{RB}). \tag{F.1}$$

In combination with formulas (19) and (F.1) and the definition of transmission interrupt probability, we have

$$\delta = P\left\{\frac{P_{R_k}\max\limits_{1\le k\le N_R}\left|h_{R_kB}\right|^2}{\sigma_B^2} < 2^{2R} - 1\right\} = P\left\{\max_{1\le k\le N_R}\left|h_{R_kB}\right|^2 < \frac{\left(2^{2R}-1\right)\sigma_B^2}{P_{R_k}}\right\} = \left[1 - \exp\left(-\frac{\left(2^{2R}-1\right)\sigma_B^2}{P_{R_k}}\right)\right]^{N_R}. \tag{F.2}$$

G. Proof of Theorem 7

According to the minimum detection error rate of (33) and its proof process, it is easy to get that for any $N_R \ge 3$, ξ^* decreases monotonically with respect to P_{R_k}. So the optimal transmission power is determined by $\xi_2^* = 1 - \varepsilon$, and the combined formula (33) can get formula (38).

H. Proof of Theorem 8

According to formulas (35) and (37), we have

$$\overline{R_C} = R(1-\delta) = R - R\left\{1 - \exp\left[-\frac{\left(2^{2R}-1\right)\sigma_B^2}{P_{R_k}}\right]\right\}^{N_R}. \tag{H.1}$$

According to the remark of Theorem 6, the larger the P_{R_k} is, the smaller the δ_2 is, and the larger the $\overline{R_C}$ is. Therefore, $\overline{R_C}$ is a monotone increasing function of P_{R_k}. The maximum value of $\overline{R_C}$ should be the maximum P_{R_k} satisfying the condition $\xi_2^* \ge 1 - \varepsilon$, and for any $N_R > 3$, ξ^* decreases monotonically with respect to P_{R_k}, so the maximum P_{R_k} satisfying condition $\xi_2^* \ge 1 - \varepsilon$ is determined by $\xi_2^* \ge 1 - \varepsilon$, that is, $P_{R_k}^*$. By substituting formula (38) into formula (H.1), we can get formula (39).

I. Proof of Theorem 9

The transmission outage probability of the first stage is δ_1, and the transmission outage probability of the second stage is δ_2, so the transmission outage probability is given by

$$\delta = 1 - (1-\delta_1)(1-\delta_2) = 1 - \left\{\begin{array}{l} 1 - 1 + \dfrac{P_A}{P_{R_j}^{\max}\varphi\left(2^{2R}-1\right)}\displaystyle\sum_{m=1}^{N_R}\binom{N_R}{m}(-1)^m\frac{1}{m} \\ \exp\left[-\dfrac{m\left(2^{2R}-1\right)\sigma_R^2}{P_A}\right]\times\ln\left[1 + m\varphi\left(2^{2R}-1\right)\dfrac{P_{R_j}^{\max}}{P_A}\right] \end{array}\right\} \times \left\{1 - \left[1 - \exp\left(-\frac{\left(2^{2R}-1\right)\sigma_B^2}{P_{R_k}}\right)\right]^{N_R}\right\}. \tag{I.1}$$

By simple calculation of (I.1), we can get (41).

J. Proof of Theorem 10

The total optimal effective covert rate should be the minimum of the optimal effective covert rate in the first and second stages.

$$\min\left(\overline{R_{C_1}^*}, \overline{R_{C_2}^*}\right). \tag{J.1}$$

By changing formula (17), we can get the following results:

$$\overline{R_{C_1}^*} = -R\sum_{m=1}^{N_R}\binom{N_R}{m}(-1)^m\exp\left[-\frac{m\left(2^{2R}-1\right)\sigma_R^2}{P_A}\right]\times\frac{\ln\left[1 + m\varphi\left(2^{2R}-1\right)(1-t_\varepsilon)\left|h_{AW}\right|^2/t_\varepsilon\left|h_{R_jW}\right|^2\right]}{m\varphi\left(2^{2R}-1\right)(1-t_\varepsilon)\left|h_{AW}\right|^2/\left(\left|h_{R_jW}\right|^2 t_\varepsilon\right)}. \tag{J.2}$$

By analyzing formula (J.2), we have

$$\frac{\ln\left[1+m\varphi\left(2^{2R}-1\right)\left(1-t_\varepsilon\right)\left|h_{AW}\right|^2/t_\varepsilon\left|h_{R_jW}\right|^2\right]}{m\varphi\left(2^{2R}-1\right)\left(1-t_\varepsilon\right)\left|h_{AW}\right|^2/\left(\left|h_{R_jW}\right|^2 t_\varepsilon\right)}<1, \quad \text{(J.3)}$$

and

$$\sum_{m=1}^{N_R}\binom{N_R}{m}(-1)^m\exp\left[-\frac{m\left(2^{2R}-1\right)\sigma_R^2}{P_A}\right]\times\frac{\ln\left[1+m\varphi\left(2^{2R}-1\right)\left(1-t_\varepsilon\right)\left|h_{AW}\right|^2/t_\varepsilon\left|h_{R_jW}\right|^2\right]}{m\varphi\left(2^{2R}-1\right)\left(1-t_\varepsilon\right)\left|h_{AW}\right|^2/\left(\left|h_{R_jW}\right|^2 t_\varepsilon\right)}<0. \quad \text{(J.4)}$$

Therefore, formula (J.2) can be further derived as follows:

$$\overline{R_{C_1}^*}<-R\sum_{m=1}^{N_R}\binom{N_R}{m}(-1)^m\exp\left[-\frac{m\left(2^{2R}-1\right)\sigma_R^2}{P_A}\right]=R-R\left\{1-\exp\left[-\frac{\left(2^{2R}-1\right)\sigma_R^2}{P_A}\right]\right\}^{N_R}. \quad \text{(J.5)}$$

Due to $\sigma_R^2=\sigma_B^2, P_A=P_{R_k}^*$, we have

$$\overline{R_{C_1}^*}<R-R\left\{1-\exp\left[-\frac{\left(2^{2R}-1\right)\sigma_B^2}{P_{R_k}^*}\right]\right\}^{N_R}=\overline{R_{C_2}}^*. \quad \text{(J.6)}$$

That is, $\min\left(\overline{R_{C_1}^*},\overline{R_{C_2}^*}\right)=\overline{R_{C_1}^*}$.

Data Availability

Relevant data supporting the research results of this paper have been elaborated in this paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under grant no. 62071486, the National Natural Science Foundation of Hunan Province under grant no. 2021JJ50051, the Scientific Research Project of Education Department Hunan Province, China (Nos. 19C0561 and 18C0531), the Natural Science Foundation on Frontier Leading Technology Basic Research Project of Jiangsu under grant no. BK2021200, and Key R & D Program of Jiangsu Province (advanced research and key core technology on Industry) Key Project and Topics (BE2021095).

References

[1] M. Chernyshev, Z. Baig, O. Bello, and S. Zeadally, "Internet of things (IoT): research, simulators, and testbeds," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1637–1647, 2018.

[2] Y. Liu and H. Su, "Some necessary and sufficient conditions for containment of second-order multi-agent systems with sampled position data," *Neurocomputing*, vol. 378, no. 32, pp. 228–237, 2020.

[3] C. Xu, H. Xu, H. Su, and C. Liu, "Disturbance-observer based consensus of linear multi-agent systems with exogenous disturbance under intermittent communication," *Neurocomputing*, vol. 404, no. 3, pp. 26–33, 2020.

[4] Y. Liu and H. Su, "Containment control of second-order multi-agent systems via intermittent sampled position data communication," *Applied Mathematics and Computation*, vol. 362, no. 362, p. 124522, 2019.

[5] C. Xu, B. Li, and L. Yang, "Semi-global containment of discrete-time high-order multi-agent systems with input saturation via intermittent control," *IET Control Theory & Applications*, vol. 14, no. 16, pp. 2303–2309, 2020.

[6] Z. Xiang, W. Yang, Y. Cai, Z. Ding, and Y. Song, "Secure transmission design in HARQ assisted cognitive NOMA networks," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 36–43, 2020.

[7] T. Zgang, Y. Huang, Y. Cai, C. Zhong, Y. Yang, and K. George, "Secure multi-antenna cognitive wiretap networks," *IEEE Transactions on Vehicular Technology*, vol. 66, pp. 36–43, 2017.

[8] X. Sun, W. Yang, Y. Cai, Z. Xiang, and X. Tang, "Secure transmissions in millimeter wave SWIPT UAV-based relay networks," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 785–788, 2019.

[9] X. Lu, W. Yang, X. Guan, and Y. Cai, "DCE-based secure transmission for massive MIMO relay system against active eavesdropper," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 11, pp. 13045–13059, 2020.

[10] Z. Xiang, W. Yang, G. Pan, Y. Cai, Y. Song, and Y. Zou, "Secure transmission in HARQ-assisted non-orthogonal multiple access networks," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 20–25, 2020.

[11] X. Lu, W. Yang, X. Guan, Q. Wu, and Y. Cai, "Robust and secure beamforming for intelligent reflecting surface aided mmWave MISO systems," *IEEE Wireless Communications Letters*, vol. 9, no. 12, pp. 2068–2072, 2020.

[12] Y. Wang, S. Yan, W. Yang, and Y. Cai, "Covert communications with constrained age of information," *IEEE Wireless Communications Letters*, vol. 10, no. 2, pp. 368–372, 2021.

[13] L. Tao, W. Yang, S. Yan, D. Wu, X. Guan, and D. Chen, "Covert communication in downlink NOMA systems with random transmit power," *IEEE Wireless Communications Letters*, vol. 9, no. 11, pp. 2000–2004, 2020.

[14] L. Yang, W. Yang, S. Xu, L. Tang, and Z. He, "Achieving covert wireless communications using a full-duplex multi-antenna receiver," in *Proceedings of the International Conference on Computer and Communications (ICCC)*, IEEE, Chengdu, China, Dec. 2019.

[15] B. A. Bash, D. Goeckel, and D. Towsley, "Limits of reliable communication with low probability of detection on AWGN channels," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1921–1930, 2013.

[16] A. Abdelaziz and C. E. Koksal, "Fundamental limits of covert communication over MIMO AWGN channel," *Proc. IEEE Conf. Commun. Network Security (CNS)*, vol. 31, no. 9, pp. 1–9, 2017.

[17] L. Wang, G. W. Wornell, and L. Zheng, "Fundamental limits of communication with low probability of detection," *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 3493–3503, 2016.

[18] K. S. K. Arumugam and M. R. Bloch, "Keyless covert communication over multiple-access channels," *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, vol. 62, no. 6, pp. 2229–2233, 2016.

[19] P. H. Che, M. Bakshi, and S. Jaggi, "Reliable deniable communication: hiding messages in noise," *Proc. IEEE Int. Symp. Inf. Theory*, vol. 13, no. 9, pp. 2945–2949, 2013.

[20] P. Mukherjee and S. Ulukus, "Covert bits through queues," *Proc. IEEE Conf. Commun. Netw. Security (CNS)*, vol. 60, no. 17, pp. 626–630, 2016.

[21] Z. Liu, J. Liu, Y. Zeng, and J. Ma, "Covert wireless communications in IoT systems:hiding information in interference," *IEEE Wireless Communications*, vol. 18, no. 33, pp. 46–52, 2018.

[22] G. Shabsigh and V. Frost, "Stochastic geometry for the analysis of wireless covert networks," *Milcom Track five-Selected Topics in Communications*, vol. 18, no. 33, pp. 46–52, 2016.

[23] T. G. Dvorkind and A. Cohen, "Maximizing miss detection for covert communication under practical constraints," *IEEE Workshop on Statistical Signal Processin*, vol. 18, no. 33, pp. 46–52, 2018.

[24] F. Shu, T. Xu, J. Hu, and S. Yan, "Delay-constrained covert communications with a full-duplex receiver," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 813–816, 2019.

[25] J. Wang, W. Tang, Q. Zhu, X. Li, H. Rao, and S. Li, "Covert communication with the help of relay and channel uncertainty," *IEEE Wireless Communications Letters*, vol. 8, no. 1, pp. 317–320, 2019.

[26] K. Shahzad, X. Zhou, S. Yan, J. Hu, F. Shu, and J. Li, "Achieving covert wireless communications using a full-duplex receiver," *IEEE Transactions on Wireless Communications*, vol. 17, no. 12, pp. 8517–8530, 2018.

[27] J. Hu, S. Yan, X. Zhou, F. Shu, J. Li, and J. Wang, "Covert communication achieved by a greedy relay in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 7, pp. 4766–4779, 2017.

[28] J. Hu, S. Yan, F. Shu, and J. Wang, "Covert transmission with a self-sustained relay," *IEEE Transactions on Wireless Communications*, vol. 18, no. 18, pp. 4089–4102, 2019.

[29] M. Forouzesh, P. Azmi, A. Kuhestani, and P. L. Yeoh, "Covert communication and secure transmission over untrusted relaying networks in the presence of multiple wardens," *IEEE Transactions on Communications*, vol. 68, no. 6, pp. 3737–3749, 2020.

[30] A. Sheikholeslami, M. Ghaderi, D. Towsley, B. A. Bash, S. Guha, and D. Goeckel, "Multi-hop routing in covert wireless networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 6, pp. 3656–3669, 2018.

[31] H.-M. Wang, Y. Zhang, X. Zhang, and Z. Li, "Secrecy and covert communications against UAV surveillance via multi-hop networks," *IEEE Transactions on Communications*, vol. 68, no. 1, pp. 389–401, 2020.

[32] X. Lu, W. Yang, Y. Cai, and X. Guan, "Proactive eavesdropping via covert pilot spoofing attack in multi-antenna systems," *IEEE Access*, vol. 7, 2019.

[33] Y. Li, R. Zhao, Y. Deng, F. Shu, and Z. Nie, "Harvest-and-Opportunistically-Relay: analyses on transmission outage and covertness," *IEEE Transactions on Wireless Communications*, vol. 97, no. 17, pp. 148–162, Aug. 2020.

[34] F. Shu, T. Xu, J. Hu, and S. Yan, "Delay-Constrained covert communications with a full-duplex receiver," *IEEE Transactions on Wireless Communications*, vol. 68, no. 1, pp. 469–485, 2019.

[35] T. Zheng, L. Xu, X. Liu, and Z. Lu, ""Covert communication with A full-duplex Receiver based on channel distribution information," international symposium on antennas," *International Symposium on An-tennas*, vol. 798, no. 13, pp. 3891–3899, 2019.

[36] O. Shmuel, A. Cohen, O. Gurewitz, and A. Cohen, "Multi-antenna Jamming in covert Communication," in *Proceedings of the 2019 IEEE International Symposium on Information Theory (ISIT)*, IEEE, Paris, France, July 2019.

[37] J. Hu, K. Shahzad, S. Yan, X. Zhou, F. Shu, and J. Li, "Covert Communications with a Full-Duplex Receiver over Wireless Fading Channels," in *Proceedings of the IEEE Int. Conf. Commun. (ICC)*, pp. 1–6, IEEE, Kansas City, MO, USA, May 2018.

[38] J. Hu, S. Yan, X. Zhou, F. Shu, and J. Wang, "Covert Communication in Wireless Relay Networks," in *Proceedings of the IEEE Global Communications Conference*, IEEE, Singapore, December 2017.

[39] E. Lehmann and J. Romano, *Testing Statistical Hypotheses*, Springer Press, New York, NY, USA, 3rd edition, 2005.

[40] Basilevsky, *Applied Matrix Algebra in the Statistical Sciences*, North-Holland, New York, 1983.

[41] A. Afana, V. Asghari, A. Ghrayeb, and S. Affes, "Cooperative relaying in spectrum-sharing systems with beamforming and interference constraints," in *Proceedings of the 2012 IEEE 13th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, IEEE, Cesme, Turkey, June 2012.

[42] L. Yang, W. Yang, J. Tu, X. Lu, L. Tang, and Z. He, "Covert communication achieved by a full-duplex multi-antenna receiver in wireless networks," *Journal of Circuits, Systems, and Computers*, p. 2150258, 2021.

Hindawi
Complexity
Volume 2022, Article ID 1288357, 10 pages
https://doi.org/10.1155/2022/1288357

Research Article

Rationalizing Denominators Using Gröbner Bases

Dongmei Li, Man Wu, Jinwang Liu, and Yiman Gao

School of Mathematics and Computing Sciences, Hunan University of Science and Technology, Xiangtan 411201, Hunan, China

Correspondence should be addressed to Dongmei Li; dmli@hnust.edu.cn

Received 29 May 2021; Revised 22 November 2021; Accepted 29 November 2021; Published 6 January 2022

Academic Editor: Zhiwei Gao

The problem of rationalizing denominators for two types of fractions is discussed in the paper. By using the theory and algorithms of Gröbner bases, we first introduce a method to rationalize the denominators of fractions with square root and cube root, and then, for the denominators with higher radical of the general form, the problem of rationalizing denominators is converted into the related problem of finding the minimal polynomials. Some interesting results and an executable algorithm for rationalizing the denominator of these type fractions are presented. Furthermore, an example is also established to illustrate the effectiveness of the algorithm.

1. Introduction

A typical topic in algebra is rationalizing denominators [1]. Rationalizing the denominator avoids the problem known as "subtrative cancellation," deals with the problem of recognizing equivalent expressions, and is commonly used in many of computer algebra systems. Also, rationalizing expression has applications in calculus [2]. More importantly, by rationalizing the denominator, mathematical operations and practical problems can be approximated more accurately [3]. Thus, there are circumstances in which it is advantageous to rationalize an expression.

The general method of denominator rationalization is to seek the rationalized factor of the denominator first and then multiply the numerator and denominator by this factor at the same time [4]. Using this method, the denominator rationalization of quadratic radical fraction has been solved. However, it is difficult to find the rationalized factor for an expression which contains a root higher than a square root [5]. Therefore, a lot of literatures tend to deal with some special case of this problem. In 1929, Paradiso [6] showed that theoretically, in all cases, and practically, in many cases, a rationalizing factor may be found by the method of undetermined coefficients. In 1970, Fateman presented an algorithm named RADCAN that is implemented in MACSYMA for the simplification of expressions containing radicals [7]. Zhou rationalized the denominator for a class of fractions $1/g(u)$ by theory of minimal polynomials in 1986, where $g(u)$ is a polynomial whose coefficient are rational and u is a complex root of a nonzero rational polynomial [8]. In 1989, Ma showed the possibility of denominator's rationalization of the irrational expressions: $A/(a_1\sqrt[n_1]{b_1}+a_2\sqrt[n_2]{b_2}+\cdots+a_m\sqrt[n_m]{b_m}+c)$, where A is the combination of radical and rational addition, subtraction, and multiplication and $a_1\sqrt[n_1]{b_1}+a_2\sqrt[n_2]{b_2}+\cdots+a_m\sqrt[n_m]{b_m}+c$ is an algebraic element in the field of rational numbers [9]. In 2000, Liu rationalized the denominator for a class of algebraic fractions as follows $1/f(\sqrt[n+1]{c})=1/(a_0+a_1\sqrt[n+1]{c}+\cdots+a_n\sqrt[n+1]{c^n})$ (where $f(x)=a_0+a_1x+\cdots+a_nx^n, c, a_1, a_2, \ldots, a_n\in Q, c>0$) using the knowledge of determinants [10]. In 2002, using polynomials, Tang also discussed the algebraic fractions whose denominator is the same with the above fraction [11]. And in 2015, Berele and Catoiu produced an exact formula for rationalizing any fraction whose denominator is a linear combination with rational coefficients of square roots of rational numbers [1].

Existing results mentioned above mainly deal with several kinds of denominators rationalization of radical fractions by using related theory of polynomials. And many other kinds of denominators of irrational fraction are unsolved, such as $1/(x+5\sqrt[7]{8}+6\sqrt[4]{9})$. This kind of fraction is a very common form in mathematical calculation, so it is of

high application value to study a general method to rationalize the denominator. In this paper, we will investigate this by the theory of Gröbner bases as it is an important tool to solve many problems in polynomial ideal [12, 13]. And it has been implemented in many computational softwares including Singular, Maple, CoCoA, Mathematica, Macaulay 2, etc. Many fundamental problems in commutative algebra, computational algebraic number theory, algebraic geometry, graph theory, image processing, cryptography and encoding, and science and engineering can be solved by it algorithmically [13–23]. And the minimal polynomial can be obtained by the reduced Gröbner basis algorithm easily. Based on this and results mentioned above, we consider using the theory of Gröbner bases to explore the relationship between minimal polynomials and denominator rationalization and discuss mainly denominator rationalization of the fraction with the form as

$$\frac{1}{x + c_1 \cdot \sqrt[m]{a} + c_2 \cdot \sqrt[n]{b}}, \tag{1}$$

where $m, n \in N^+, m, n \geq 2, a, b, c_1, c_2 \in Q, a, b > 0$. We hope to establish a simplified method for rationalizing the denominator of this type fraction.

The rest of the paper is organized as follows. We present some preliminary knowledge, basic concepts, and a special method to rationalize the denominators with square root and cube root in Section 2. Main results on rationalizing the denominator of a type fraction are shown in Section 3. A simplified algorithm and an example established to illustrate the algorithm are given in Section 4. Section 5 concludes the paper.

2. Denominators with Square Root and Cube Root

In what follows, $K[x_1, x_2, \ldots, x_n]$ will denote the polynomial ring in n variables $x_1, x_2, \ldots, x_n$ with coefficients in a field K, L/K will be an algebraic extension of the field K, and $[L: K]$ will denote the extension times of the extension field L/K. Q will denote the rational number field, N will be the set of integers, and $\mathrm{lcm}(m, n)$ will denote the least common multiple of m and n. For a nonzero polynomial $f(x) \in K[x_1, x_2, \ldots, x_n]$, we use $lt(f)$, $lc(f)$, and $lp(f)$ to denote the leading term, the leading coefficient, and the leading monomial of $f(x)$, respectively. For a set $F \subset K[x_1, x_2, \ldots, x_n]$, we denote $lt(F) = \{lt(f) | f \in F\}$. Then, we introduce several related definitions and algorithm.

Definition 1. Let I be an ideal in $K[x_1, x_2, \ldots, x_n]$. A finite subset G of I is called a Gröbner basis of I if

$$\langle lt(G) \rangle = \langle lt(I) \rangle. \tag{2}$$

Definition 2 (see [13]). Let $f, g \in K[x_1, x_2, \ldots, x_n], f, g \neq 0$, and $L = lcm(lp(f), lp(g))$, and the S-polynomial of f and g is defined as (Algorithm 1)

$$S(f, g) = \frac{L}{lt(f)} f - \frac{L}{lt(g)} g. \tag{3}$$

This section focuses on the problem of rationalizing the denominator with square root and cube root. For the sake of convenience in researching the problem, we put the coefficient in the radical sign, and then, the fraction can be reduced to one of the following forms:

$$\begin{aligned} &(1) \frac{1}{x + \sqrt[2]{a} + \sqrt[3]{b}} \ (a > 0, \text{ and } a, b \in Q), \\ &(2) \frac{1}{x - \sqrt[2]{a} + \sqrt[3]{b}} \ (a > 0, \text{ and } a, b \in Q). \end{aligned} \tag{4}$$

We first consider the denominators rationalization of form (1).

Lemma 1 (see [13]). *Let $G = \{g_1, g_2, \ldots, g_s\}$ be a subset of an ideal $I \in K[x_1, x_2, \ldots, x_n]$, and then, G is a Gröbner bases for I if and only if $S(g_i, g_j) \xrightarrow{G} 0$, for all $i \neq j, 1 \leq i, j \leq s$.*

Theorem 1. *Let I be an ideal in $K[x_1, x_2, \ldots, x_n]$, and then, $fg - 1 \in I \Longleftrightarrow 1 \in \langle I, f \rangle$.*

Proof. If $fg - 1 \in I$, then there exists $h \in I$, such that $fg - 1 = h$, that is, $1 = fg - h \in \langle I, f \rangle$.

In turn, if $1 \in \langle I, f \rangle$, then there exists $g \in K[x_1, x_2, \ldots, x_n]$ and $h \in I$, such that $1 = fg + h$, that is, $fg - 1 = -h \in I$.

Next, we introduce how to rationalize the denominator of fraction as form (1). First, consider the following ideal:

$$I = \langle y_1^2 - a, y_2^3 - b \rangle \subseteq Q(x)[y_1, y_2], \tag{5}$$

where $Q(x)[y_1, y_2]$ denotes the polynomial ring in variables y_1 and y_2 on $Q(x)$ and $Q(x)$ is an extension of the field Q with x as a variable. Let

$$f_1(y_1, y_2) = y_1^2 - a, f_2(y_1, y_2) = y_2^3 - b. \tag{6}$$

In the following, multivariate polynomials $h(x, y)$ are denoted as h for convenience. And the term order of the multivariate polynomial ring is the lexicographical ordering, which is defined by $y_1 > y_2$.

Note that

$$S(f_1, f_2) = by_1^2 - ay_2^3 \xrightarrow{f_2} by_1^2 + ab \xrightarrow{f_1} 0. \tag{7}$$

By Lemma 1, we have that $\{f_1, f_2\}$ is a Gröbner basis of I. Let

Input: $F = f_1, f_2, \ldots, f_s$.
Output: $G = g_{1,2}, \ldots, g_s$, G is a Gröbner basis of $\langle f_1, f_2, \ldots, f_s \rangle$.
Initialization: $G \coloneqq F, \Omega \coloneqq \{\{f_i, f_j\} | f_i \neq f_j \in G\}$.
When $\Omega \neq \varnothing$
select $f, g \in \Omega$
$\Omega \coloneqq \Omega / \{\{f, g\}\}$
$S(f, g) \xrightarrow{G}_{+} h$, h is reductive with respect to G.
If $h \neq 0$, then
$\Omega \coloneqq \Omega \cup \{\{u, h\} | \forall u \in G\}$
$G \coloneqq G \cup \{h\}$

Algorithm 1: Buchberger algorithm [14].

$$f_3(y_1, y_2) = x + y_1 + y_2, \tag{8}$$

$y_1 = \sqrt{a}, y_2 = \sqrt[3]{b}$, and then, $f_3(\sqrt{a}, \sqrt[3]{b}) = x + \sqrt{a} + \sqrt[3]{b}$, which is the denominator we want to deal with. Replace f by f_3 in Theorem 1, and then, the condition $f_3 \cdot g - 1 \in I$ is equivalent to $f_3(\sqrt{a}, \sqrt[3]{b}) \cdot g(\sqrt{a}, \sqrt[3]{b}) = 1$, where $g(\sqrt{a}, \sqrt[3]{b}) \in Q[\sqrt{a}, \sqrt[3]{b}]$ is the polynomial we need.

So, the key issue is whether $1 \in \langle I, f_3 \rangle$, that is, whether $1 \in \langle f_1, f_2, f_3 \rangle$. If true, we can rationalize the denominator of $1/(x + \sqrt{a} + \sqrt[3]{b})$ by using this method.

Now, we use Buchberger algorithm to obtain the Gröbner basis of $J = \langle f_1, f_2, f_3 \rangle$ and then determine whether $1 \in \langle f_1, f_2, f_3 \rangle$.

Initialize
$f_1 = y_1^2 - a$, $f_2 = y_2^3 - b$, and $f_3 = x + y_1 + y_2$, and using the algorithm, we get the Gröbner basis of J is

$$\{f_1, f_2, f_3, f_4, f_5, f_6\}, \tag{9}$$

where

$$\begin{aligned} f_4 &= y_2^2 - x y_1 + x y_2 - a, \\ f_5 &= \left(a + 3x^2\right) y_2 + 2x^3 - 2ax - b, \end{aligned} \tag{10}$$

$$f_6 = \frac{-\left(2x^3 - 2ax - b\right)^3}{\left(3x^2 + a\right)^3} - b. \tag{11}$$

Observe f_6, and it is easy to find that f_6 is independent of y_1 and y_2, and thus, $f_6 \in Q(x)$. Note that $Q(x)$ is a field, and thus, $f_6^{-1} = 1/f_6 \in Q(x)$. Consequently, $1 = f_6 \cdot 1/f_6 \in J$. Therefore, we can calculate the rationalizing result of $1/(x + \sqrt{a} + \sqrt[3]{b})$.

Then, express f_6 as the combination of f_1, f_2, f_3, f_4, f_5:

$$f_6 = f_2 - \left(\frac{y_2^2}{3x^2 + a} + \frac{-2x^3 + 2ax + b}{\left(3x^2 + a\right)^2} + \frac{\left(2x^3 - 2ax - b\right)^2}{\left(3x^2 + a\right)^3} \right) \cdot f_5, \tag{12}$$

that is,

$$\frac{f_2 - \left(y_2^2 / \left(3x^2 + a\right) + \left(-2x^3 + 2ax + b\right) / \left(3x^2 + a\right)^2 + \left(2x^3 - 2ax - b\right)^2 / \left(3x^2 + a\right)^3 \right) \cdot f_5}{f_6} = 1. \tag{13}$$

Hence, divide both sides of equation (13) by $x + \sqrt{a} + \sqrt[3]{b}$, and substitute the value of f_2, f_5 and $y_1 = \sqrt{a}, y_2 = \sqrt[3]{b}$ into the equation, and then, we obtain

$$\begin{aligned}\frac{1}{x+\sqrt{a}+\sqrt[3]{b}} &= \frac{\left(2x^3-2ax-b\right)^2\cdot 2x^2+bx\cdot\left(3x^2+a\right)^2+b\left(2x^3+2ax+b\right)\cdot\left(3x^2+a\right)}{\left(2x^3-2ax-b\right)^3+b\left(3x^2+a\right)^3}\\ &+\frac{b\left(3x^2+a\right)^2-2x\cdot\left(2x^3-2ax-b\right)^2}{\left(2x^3-2ax-b\right)^3+b\left(3x^2+a\right)^3}\cdot\sqrt{a}\\ &+\frac{-b\left(3x^2+a\right)^2+2x^2\cdot\left(-2x^3+2ax+b\right)\cdot\left(3x^2+a\right)+x\left(2x^3-2ax-b\right)^2}{\left(2x^3-2ax-b\right)^3+b\left(3x^2+a\right)^3}\cdot\sqrt[3]{b}\\ &+\frac{2x^2\cdot\left(3x^2+a\right)^2+x\cdot\left(3x^2+a\right)\cdot\left(-2x^3+2ax+b\right)-\left(2x^3-2ax-b\right)^2}{\left(2x^3-2ax-b\right)^3+b\left(3x^2+a\right)^3}\cdot\sqrt[3]{b^2}\\ &+\frac{2x\cdot\left(2x^3-2ax-b\right)\cdot\left(3x^2+a\right)+\left(2x^3-2ax-b\right)^2}{\left(2x^3-2ax-b\right)^3+b\left(3x^2+a\right)^3}\cdot\sqrt{a}\cdot\sqrt[3]{b}\\ &+\frac{-2x\cdot\left(3x^2+a\right)^2+\left(-2x^3+2ax+b\right)\cdot\left(3x^2+a\right)}{\left(2x^3-2ax-b\right)^3+b\left(3x^2+a\right)^3}\cdot\sqrt{a}\cdot\sqrt[3]{b^2}.\end{aligned} \tag{14}$$

Using Matlab to simplify the equation above, we obtain

$$\begin{aligned}\frac{1}{x+\sqrt{a}+\sqrt[3]{b}} &= \frac{1}{x^6-3ax^4+2bx^3+3a^2x^2+6abx+b^2-a^3}\left(x^5-2ax^3+bx^2\right.\\ &+a^2x+ab+\left(-x^4+2ax^2+2bx-a^2\right)\cdot\sqrt{a}+\left(-x^4-bx+a^2\right)\\ &\cdot\sqrt[3]{b}+\left(x^3+3ax+b\right)\cdot\sqrt[3]{b^2}+\left(2x^3-2ax-b\right)\cdot\sqrt{a}\cdot\sqrt[3]{b}\\ &\left.+\left(-3x^2-a\right)\cdot\sqrt{a}\cdot\sqrt[3]{b^2}\right).\end{aligned} \tag{15}$$

For form (2), we just need to change $f_3(y_1, y_2) = x + y_1 + y_2$ to $f_3(y_1, y_2) = x - y_1 + y_2$, and the other steps are the same. □

3. Denominators with Higher Radical

In this section, we focus on rationalizing the denominator of general form as

$$\frac{1}{x+c_1\cdot\sqrt[m]{a}+c_2\cdot\sqrt[n]{b}}, \tag{16}$$

where $m, n \in N^+, m, n \geq 2, a, b, c_1, c_2 \in Q, a, b > 0$.

By the method in Section 2, we first construct the ideal $J = < f_1, f_2, f_3 > \in Q(x)[y_1, y_2]$, where $f_1 = y_1^m - a$, $f_2 = y_2^n - n$, and $f_3 = x + c_1y_1 + c_2y_2$. Then, compute the Gröbner basis G and the reduced Gröbner basis G_0 of J. If $G_0 = \{1\}$, then calculate $f_s = G \cap Q(x)$. Finally, express f_s as a combination of $G/\{f_s\}$.

However, there are two important uncertainties, one is whether $G_0 = \{1\}$ and the other is whether f_s is a combination of $G/\{f_s\}$. So, this method can only solve the case when c_1, c_2, m, n are specific values. Because of this limitation, we hope to find another method for the general form.

First, we introduce several related definitions and lemmas.

Definition 3 (see [24]). Let L/K be an algebraic extension of the field K, $[L: K] = n$. Suppose $\sigma_i: L \longrightarrow C\,(1 \leq i \leq n)$ is n K-insertion of L, where C is the complex field. For $\alpha \in L$, define

$$T_{L/K}(\alpha) = \sum_{i=1}^{n}\sigma_i(\alpha), \tag{17}$$

as the trace of α to the expansion L/K.

Definition 4 (see [24]). Let L/K be an algebraic extension of the field K. Suppose $\sigma_1, \ldots, \sigma_n$ is the K-insertion of L and $\alpha_1, \ldots, \alpha_n \in L$, and we define

$$d_{L/K}(\alpha_1, \ldots, \alpha_n) = \det^2\left(\sigma_i\left(\alpha_j\right)\right), 1 \leq i \leq n, 1 \leq j \leq n, \tag{18}$$

as the discriminant of $\{\alpha_1, \ldots, \alpha_n\}$ for the expansion L/K.

Definition 5. Let K be a field, $u \in K$, and $f(x) \in K[x]$ be a polynomial. We call $f(x)$ is the minimal polynomial of u in K if $f(x)$ satisfies the following:

(1) $f(x)$ is monic and $f(u) = 0$

(2) If $g(x) \in K[x]$ and $g(u) = 0$, then $f(x) | g(x)$

Lemma 2 (see [24]). *Let L/K be the expansion of number field and $[L: K] = n$. Suppose $\alpha \in L$, $f(x) = x^m - c_1 x^{m-1} + \cdots + (-1)^m c_m$ is the minimal polynomial of α in K, where $m = [K(\alpha): K]$, and then, $T_{L/K}(\alpha) = nc_1/m$.*

Lemma 3 (see [23]). *Let $d_{L/K}(\alpha_1, \ldots, \alpha_n)$ be the discriminant of element $\{\alpha_1, \ldots, \alpha_n\}$ for the expansion L/K, and then, $d_{L/K}(\alpha_1, \ldots, \alpha_n) \neq 0$ if and only if $\alpha_1, \ldots, \alpha_n$ is K-linearly independent.*

Lemma 4 (see [25], Eisenstein criterion). *Let $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ be an integral coefficient univariate polynomial. If there is a prime number p such that*

(1) $p \nmid a_n$

(2) $p \mid a_{n-1}, a_{n-2}, \ldots, a_0$

(3) $p \nmid a_0^2$

Then, $f(x)$ is irreducible over the rational number field Q.

Using the lemmas above, we can prove the following results.

Theorem 2. *If p is prime number, then $f(x) = x^m - p$ is the minimal polynomial for $\sqrt[m]{p}$ in rational number filed Q, where $m \in N^+$ and $m \geq 2$.*

Proof. It is straightforward that $\sqrt[m]{p}$ is a root of $f(x) = x^m - p = 0$. In the following, we prove that $f(x)$ is irreducible.

Let $g(x) = f(x + p) = x^m + mpx^{m-1} + \cdots + C_m^i p^i x^{m-i} + \cdots + p^m - p$, where $C_m^i = m!/(i!(m-i)!)$. It is obvious that $p \nmid 1$, $p | C_m^i p^i (1 \leq i \leq m-1)$, and $p | (p^m - p)$. In fact, we can prove that $p^2 \nmid (p^m - p)$. Suppose $p^2 | (p^m - p)$, and then, $p | (p^{m-1} - 1)$. Combining $p | p^{m-1} (m \in N^+, m \geq 2)$, we have that $p | (p^{m-1} - (p^{m-1} - 1))$, that is, $p | 1$, and this is a contradiction. So, $p^2 \nmid (p^m - p)$. By the Eisenstein criterion, $g(x)$ is irreducible, so $f(x)$ is irreducible, and it is the minimal polynomial for $\sqrt[m]{p}$ in Q. □

Theorem 3. *If $p_1, p_2, \ldots, p_s$ are different prime numbers, then*

$$\sqrt[m1]{p_1^{t_1}} \cdot \sqrt[m2]{p_2^{t_2}}, \cdots \sqrt[ms]{p_s^{t_s}}, (t_i < m_i, t_i, m_i \in N^+, i = 1, 2, \ldots, s), \tag{19}$$

is an irrational number.

Proof. It is straightforward that $\sqrt[m1]{p_1^{t_1}} \cdot \sqrt[m2]{p_2^{t_2}} \cdots \sqrt[ms]{p_s^{t_s}}$ is a real number. In the following, we prove that it is an irrational number. Suppose it is a rational number, and set $f(x) = x^{m_1 m_2 \cdots m_s} - p_1^{t_1 m_2 \cdots m_s} p_2^{m_1 t_2 \cdots m_s} \cdots p_s^{m_1 m_2 \cdots t_s}$. Then, $\sqrt[m1]{p_1^{t_1}} \cdot \sqrt[m2]{p_2^{t_2}} \cdots \sqrt[ms]{p_s^{t_s}}$ is a positive rational root of $f(x) = 0$. Note that $f(x)$ is an integral coefficient polynomial and $\mathrm{lc}(f) = 1$, and then, the positive rational root of $f(x)$ must be a factor of its constant term. So, it has the form as $p_1^{n_1} p_2^{n_2} \cdots p_s^{n_s}$, where $0 \leq n_i \leq m_1 \cdots m_{i-1} t_i m_{i+1} \cdots m_s$, $n_i \in N$, $i = 1, 2, \ldots, s$. Then, we have that

$$\sqrt[m1]{p_1^{t_1}} \cdot \sqrt[m2]{p_2^{t_2}} \cdots \sqrt[ms]{p_s^{t_s}} = p_1^{n_1} p_2^{n_2} \cdots p_s^{n_s}. \tag{20}$$

From the equation above, we see that some of n_i must be 0. Without loss of generality, we denote the elements whose power exponents are nonzero as $p_1, p_2, \ldots, p_q$, and equation (20) turns into the following:

$$\sqrt[m1]{p_1^{t_1}} \cdot \sqrt[m2]{p_2^{t_2}} \cdots \sqrt[ms]{p_s^{t_s}} = p_1^{n_1} p_2^{n_2} \cdots p_q^{n_q}, \tag{21}$$

that is,

$$p_1^{n_1 - t_1/m_1} p_2^{n_2 - t_2/m_2} \cdots p_q^{n_q - t_q/m_q} = p_{q+1}^{t_{q+1}/m_{q+1}} p_{q+2}^{t_{q+2}/m_{q+2}} \cdots p_s^{t_s/m_s}. \tag{22}$$

Multiply both sides of equation (22) by the $m_1 m_2 \cdots m_s$ power, and then,

$$p_1^{l_1} p_2^{l_2} \cdots p_s^{l_s} = p_{q+1}^{l_{q+1}} p_{q+2}^{l_{q+2}} \cdots p_s^{l_s}, \tag{23}$$

where $l_i \in N^+, i = 1, 2, \ldots, s$. Hence, $p_1 \mid p_{q+1}^{l_{q+1}} p_{q+2}^{l_{q+2}} \cdots p_s^{l_s}$, and this contradicts $p_1, p_2, \ldots, p_s$ are different prime numbers. So, the conclusion is correct. □

Theorem 4. *If $p_1, \ldots, p_s$ are different prime numbers and $m_1, \ldots, m_s$ are positive integers and no less than 2, then the finite extension times of $Q(\sqrt[m1]{p_1}, \ldots, \sqrt[ms]{p_s})/Q$ is no more than $m_1 m_2 \ldots m_s$.*

Proof. Let $F_i = Q(\sqrt[m1]{p_1}, \sqrt[m2]{p_2}, \ldots, \sqrt[mi]{p_i})$, $1 \leq i \leq s$ and $i \in N^+$, and then, $F_{i+1} = F_i(\sqrt[mi+1]{p_{i+1}})$, that is, F_{i+1} can be viewed as the single extension of F_i. So, $[F_{i+1}: F_i] = \partial(g_i)$, where $\partial(g_i)$ denotes the degree of the polynomial $g_i(x)$, and $g_i(x)$ is the minimal polynomial of $\sqrt[mi]{p_i}$ in the field F_i. Obviously, $\sqrt[mi]{p_i}$ is a root for $f_i(x) = x^{m_i} - p_i$, $1 \leq i \leq s$. By the definition of minimal polynomial, we see that $\partial(g_i) \leq \partial(f_i)$, $2 \leq i \leq s, i \in N^+$. From Theorem 3, we see that $f_1(x)$ is an irreducible polynomial in Q. By the property of domain extension, we have $[F_1: Q] = m_1$. Hence,

$$\begin{aligned}\left[Q\left(\sqrt[m_1]{p_1}, \sqrt[m_2]{p_2}, \ldots, \sqrt[m_s]{p_s}\right): Q\right] &= [F_s : F_{s-1}][F_{s-1} : F_{s-2}] \cdots [F_2 : F_1][F_1 : Q] \\ &= \partial(g_s) \cdot \partial(g_{s-1}) \cdots \partial(g_2) \cdot m_1 \\ &\le m_1 m_2 \cdots m_s.\end{aligned} \tag{24}$$

□

Theorem 5. *If $p_1, p_2, \ldots, p_s$ are different prime numbers and $m_1, m_2, \ldots, m_s$ are positive integers and no less than 2, then $x^{m_s} - p_s$ is the minimal polynomial of $\sqrt[m_s]{p_s}$ in the field $Q(\sqrt[m_1]{p_1}, \sqrt[m_2]{p_2}, \ldots, \sqrt[m_{s-1}]{p_{s-1}})$.*

Proof. Let $b_i = \sqrt[m_i]{p_i}, 1 \le i \le s, M = Q(\sqrt[m_1]{p_1}, \sqrt[m_2]{p_2}, \ldots, \sqrt[m_s]{p_s})$, and in the following, we prove that all the $m_1 \cdot m_2 \cdots m_s$ elements in set

$$\begin{aligned}A = \{&b_1^{n_1} b_2^{n_2} \cdots b_s^{n_s} \mid 0 \le n_i \\ &\le m_i - 1, n_i \in N, i = 1, 2, \ldots, s\} \subseteq M,\end{aligned} \tag{25}$$

are linearly independent. First, sort the elements in A. It is easy to observe that each element in A corresponds to such an array $(n_1, n_2, \ldots, n_s)$. So, we can turn the problem into sorting the exponents $(n_1, n_2, \ldots, n_s)$, where the term order is lexicographic order. We denote the element that corresponds to the largest exponential as c_1, the element that corresponds to the second largest exponential as $c_2, \ldots$, and the element that corresponds to the smallest exponential as $c_{m_1 m_2 \cdots m_s}$, and then, we have sorted out all the elements in A.

Now, we calculate $\det(T_{M/Q}(c_i c_j))$. Suppose $c_i = b_1^{n_{i_1}} b_2^{n_{i_2}} \cdots b_s^{n_{i_s}}$ and $c_j = b_1^{n_{j_1}} b_2^{n_{j_2}} \cdots b_s^{n_{j_s}}$. If one of the following two situations is true,

(1) $\forall k = 1, 2, \ldots, s, n_{i_k} + n_{j_k} = m_k$

(2) $\forall k = 1, 2, \ldots, s, n_{i_k} = n_{j_k} = 0$

We obtain that $c_i c_j \in Q$ by Theorem 3. Hence, $T_{M/Q}(c_i c_j) = n c_i c_j \ne 0$, where $n = [M: Q]$. Otherwise, there exists $k \in \{1, 2, \ldots, s\}$ does not satisfy either of the above two cases, and we can calculate $T_{M/Q}(c_i c_j)$ by Lemma 2. We first want to obtain the minimal polynomial of $c_i c_j$ in rational number field Q. Set

$$l = \operatorname{lcm}\left(\frac{m_1}{\gcd\left(m_1, n_{i_1} + n_{j_1}\right)}, \frac{m_2}{\gcd\left(m_2, n_{i_2} + n_{j_2}\right)}, \ldots, \frac{m_s}{\gcd\left(m_s, n_{i_s} + n_{j_s}\right)}\right). \tag{26}$$

Obviously, $l > 1$. Let $q = (c_i c_j)^l$ and $B = \left\{m \in N^+ \mid (c_i c_j)^m \in Q\right\}$, and by Theorem 3, we have that $m_k \mid (n_{i_k} + n_{j_k})m$, for any $m \in B$, where $k = 1, 2, \ldots, s$. Hence, p is the smallest element in B.

In the following, we prove that $f(x) = x^l - q$ is the minimal polynomial of $c_i c_j$ in the rational number field Q.

It is straightforward that $f(c_i c_j) = (c_i c_j)^l - q = 0$. We factorize $f(x)$ in the complex field C as

$$f(x) = \left(x - \sqrt[l]{q}\right)\left(x - \varepsilon\sqrt[l]{q}\right) \cdots \left(x - \varepsilon^{l-1}\sqrt[l]{q}\right), \tag{27}$$

where $\varepsilon = e^{2\pi i/l} = \cos(2\pi/n) + i \sin(2\pi/n)$.

Suppose $f(x)$ is reducible in Q, and then, some constant terms of the linear factor in the decomposition above are rational numbers, that is, there is a $t \in N^+$ and $k \in N$ such that $(-1)^t \varepsilon^k q^{t/l} \in Q \subseteq R$, i.e., $e^{2k\pi i/l} = \cos(2k\pi/l) + i \sin(2k\pi/l) \in R$. So, $\sin(2k\pi/l) = 0$, that is, $k = \ln/2, n \in Z$. Hence, $\varepsilon^k = \cos(2k\pi/l) = \cos(n\pi) = \pm 1, n \in Z$, and there is an $r \in Q$ such that $q^{t/l} = r$. Note that $q = (c_i c_j)^l$, and then, $(c_i c_j)^t = r \in Q$ $(t < l)$, and this contradicts the selection of l. Thus, $f(x) = x^l - q$ is irreducible in Q, and then, $f(x)$ is the minimal polynomial of $c_i c_j$ in rational number field Q. Combined with Lemma 2, we see that $T_{M/Q}(c_i c_j) = 0$.

Based on the discussion above, we obtain that $d_{M/Q}(c_1, c_2, \ldots, c_{m_1 m_2 \cdots m_s}) = \det(T_{M/Q}(c_i c_j)) \ne 0$. Then, all elements in set A are linearly independent in Q by Lemma 3. Combined with Theorem 4, we see that

$$\left[Q\left(\sqrt[m_1]{p_1}, \sqrt[m_2]{p_2}, \ldots, \sqrt[m_s]{p_s}\right): Q\right] = m_1 m_2 \cdots m_s. \tag{28}$$

So, $x^{m_s} - p_s$ is the minimal polynomial of $\sqrt[m_s]{p_s}$ in the field $Q(\sqrt[m_1]{p_1}, \ldots, \sqrt[m_{s-1}]{p_{s-1}})$.

Based on the results above, we can present the steps for rationalizing denominators of the following form:

$$\frac{1}{x + c_1 \sqrt[m]{a} + c_2 \sqrt[n]{b}}, \tag{29}$$

where m and n are integers and $a, b \in Q^+$ and $c_1, c_2 \in Q$.

Step 1: write a and b as fractions in the lowest term, $a = d_1/e_1$ and $b = d_2/e_2$. Then, decompose d_1, e_1, d_2, e_2 into power product of different prime factors, respectively. We use $p_1, p_2, \ldots, p_t$ to denote the common prime factors of d_1 and d_2, $q_1, q_2, \ldots, q_k$ to denote the different prime factors in d_1, e_1, and $v_1, v_2, \ldots, v_s$ to denote the different prime factors in d_2, e_2.

Step 2: let $l = lcm(m, n)$, and

$$\begin{aligned}\alpha_1 = \sqrt[l]{p_1}, \alpha_2 = \sqrt[l]{p_2}, \ldots, \alpha_t = \sqrt[l]{p_t}, \\ \alpha_{t+1} = \sqrt[m]{q_1}, \alpha_{t+2} = \sqrt[m]{q_2}, \ldots, \alpha_{t+w} = \sqrt[m]{q_w}, \\ \alpha_{t+w+1} = \sqrt[n]{r_1}, \alpha_{t+w+2} = \sqrt[n]{r_2}, \ldots, \alpha_{t+w+s} = \sqrt[n]{r_s}.\end{aligned} \tag{30}$$

We construct the rational extension field:

$$M = Q(\alpha_1, \ldots, \alpha_t, \alpha_{t+1}, \ldots, \alpha_{t+w}, \ldots, \alpha_{t+w+s}). \tag{31}$$

Let $\alpha = c_1 \sqrt[m]{a} + c_2 \sqrt[n]{b}$, and there are $f(x_1, \ldots, x_{t+w+s}), g(x_1, \ldots, x_{t+w+s}) \in Q[x_1, \ldots, x_{t+w+s}]$ such that

$$\alpha = f(\alpha_1, \ldots, \alpha_{t+w+s}) / g(\alpha_1, \ldots, \alpha_{t+w+s}). \tag{32}$$

According to Theorem 2, we have that $h_1 = x_1^l - p_1 \in Q[x_1]$ is the minimal polynomial of α_1 in the rational number field Q. It is also known, by Theorem 5, $h_i = x^l - p_i \in Q(\alpha_1, \ldots, \alpha_{i-1})[x_i]$ is the minimal polynomial of α_i in $Q(\alpha_1, \ldots, \alpha_{i-1})$ for $2 \le i \le t$, $h_i = x^m - q_{i-t} \in Q(\alpha_1, \ldots, \alpha_{i-1})[x_i]$ is the minimal polynomial of α_i in $Q(\alpha_1, \ldots, \alpha_{i-1})$ for $t+1 \le i \le t+s$, and $h_i = x^n - v_{i-t-w} \in Q(\alpha_1, \ldots, \alpha_{i-1})[x_i]$ is the minimal polynomial of α_i in $Q(\alpha_1, \ldots, \alpha_{i-1})$ for $t+w+1 \le i \le t+w+s$.

Next, we construct the following homomorphic maps:

$$\varphi_i : Q[x_1, x_2, \ldots, x_i] \longrightarrow Q(\alpha_1, \alpha_2, \ldots, \alpha_{i-1})[x_i], 2 \le i \le t+w+s. \tag{33}$$

For any $f = \sum \lambda_\beta x_1^{\beta_1} x_2^{\beta_2} \ldots x_i^{\beta_i} \in Q[x_1, x_2, \ldots, x_{i-1}]$, where $\lambda_\beta \in Q, \beta = (\beta_1^\beta, \beta_2, \ldots, \beta_i) \in N^i$. Define

$$\varphi_i(f) = \sum_\beta \lambda_\beta \alpha_1^{\beta_1} \alpha_2^{\beta_2} \cdots \alpha_{i-1}^{\beta_{i-1}} x_i^{\beta_i}. \tag{34}$$

Under this map, the corresponding preimage of $h_i \in Q(\alpha_1, \alpha_2, \ldots, \alpha_{i-1})[x_i]$ is itself, where $2 \le i \le t+w+s$.

Step 3: we find the minimal polynomial of α in Q according to [26].

The term order $\prec$ in $Q[x_1, \ldots, x_t, x_{t+1}, \ldots, x_{t+w+s}, y]$ is the lexicographical ordering defined by $y \prec x_{t+w+s} \prec \cdots \prec x_{t+1} \prec x_t \prec \cdots \prec x_1$. Under this term order, we calculate the reduced Gröbner basis G of the ideal

$$I = \langle h_1, h_2, \ldots, h_{t+w+s}, g \cdot y - f \rangle, \tag{35}$$

where

$$h_i = \begin{cases} x^l - p_i, & 1 \le i \le t, \\ x^m - q_{i-t}, & t+1 \le i \le t+w, \\ x^n - r_{i-t-w}, & t+w+1 \le i \le t+w+s. \end{cases} \tag{36}$$

Then, we compute $h(y) = G \cap Q[y]$ and take

$$m(y) = \frac{h(y)}{\lambda}, \lambda = lc(h(y)). \tag{37}$$

Hence, $m(y)$ is the minimal polynomial of α in Q. And then we can rationalize the denominator by applying the minimal polynomial.

Set $n(y) = x + y$ (x is a parameter) and divide $m(y)$ by $n(y)$ such that

$$m(y) = q(y)n(y) + r(y), \partial(r(y)) < \partial(n(y)). \tag{38}$$

Then, $\partial(r(y)) = 0$ and $r(y) \ne 0$, otherwise, it contradicts that $n(y)$ contains x. Correspondingly, equation (38) can be rewritten as

$$m(y) = q(y)n(y) + r(x), r(x) \in Q[x]. \tag{39}$$

Substitute $y = \alpha$ into the equation above, and then, $0 = (x + \alpha)q(\alpha) + r(x)$, that is,

$$\frac{1}{x + \alpha} = -\frac{q(\alpha)}{r(x)}, \tag{40}$$

which is the desired result of rationalizing denominators. □

4. Algorithm and Example

According to the theorems and discussion in Section 3, we obtain an algorithm for rationalizing the denominators of fractions with the form as $1/(x + c_1 \cdot \sqrt[m]{a} + c_2 \cdot \sqrt[n]{b})$. We describe this algorithm in more detail in Figure 1.

In the following, we construct an example to show the effectiveness of the algorithm.

Example 1. Rationalizing the denominator of

$$\frac{1}{x + \sqrt{6} + \sqrt[3]{4}}. \tag{41}$$

Step 1: we know $a = 6$ and $b = 4$, and the command "format rat" can be omitted here. Carry out the prime factorization of 6 and 4 by using the function "factorization," and we obtain

$$6 = 2 \times 3, 4 = 2 \times 2. \tag{42}$$

It is easy to see that $p_1 = 2$ is a common prime factor of 6 and 4. In addition, 6 also contains the factor $q_1 = 3$.

Step 2: note that $m = 2$ and $n = 3$, and then, using the function "min_GBS," we obtain

$$l = \mathrm{lcm}(m, n) = 6. \tag{43}$$

Let

$$\alpha_1 = \sqrt[6]{2}, \alpha_2 = \sqrt{3}. \tag{44}$$

We construct the extension field $Q(\alpha_1, \alpha_2)$ of the rational number field Q. Set $\alpha = \sqrt{6} + \sqrt[3]{4}$, and then, $\alpha = f(\alpha_1, \alpha_2) = \alpha_1^3 \alpha_2 + \alpha_1^4$.

Step 3: hence, $h_1 = x_1^6 - 2 \in Q[x_1]$ is a minimal polynomial of α_1 on Q and $h_2 = x_2^2 - 3 \in Q(\alpha_1)[x_2]$ is a minimal polynomial of α_2 on $Q(\alpha_1)$.

Set

$$I = \langle x_1^6 - 2, x_2^2 - 3, y - x_1^3 x_2 - x_1^4 \rangle, \tag{45}$$

take the lexicographical ordering $y \prec x_2 \prec x_1$ as the term order in $Q[x_1, x_2, y]$, and calculate the reduced Gröbner basis G of ideal I in the software singular using the built-in function "groebner." Then, the calculation results are as follows:

$$G = \{f_1, f_2, f_3\}, \tag{46}$$

where

An algorithm for rationalizing denominator	
Input:	$m, n \in N^+$, $a, b \in Q^+$, $c_1, c_2 \in Q$, the term orders over Q
Output:	$\alpha, q(x), r(x)$ such that $\frac{1}{x+\alpha} = -\frac{q(\alpha)}{r(x)}$
Step 1.	Write a, b as fractions in lowest term $a = d_1/e_1, b = d_2/e_2$ using the command "format rat". Then decompose $d_i, e_i (i = 1, 2)$ into power product of different prime factors respectively using the function "factorization", denote the common prime factors of d_1, d_2 by p_i, use q_j and r_k to denote the different prime factors in d_1, e_1 and d_2, e_2 respectively, $i, j, k = 1, 2, \ldots$
Step 2.	Compute $l = \mathrm{lcm}(m, n)$, set $\alpha_i = \sqrt[l]{p_i}$, $i = 1, 2, \ldots, t$; $\alpha_{t+j} = \sqrt[l]{q_j}$, $j = 1, \ldots, w$, and $\alpha_{t+w+k} = \sqrt[l]{r_k}$, $k = 1, \ldots, s$. Construct the extension field $M = Q(\alpha_1, \alpha_2, \ldots, \alpha_{t+w+s})$. Find $f(x_1, \ldots, x_{t+w+s})$, $g(x_1, \ldots, x_{t+w+s}) \in Q(x_1, \ldots, x_{t+w+s})$ such that $\alpha = \frac{f(\alpha_1, \ldots, \alpha_{t+w+k})}{g(\alpha_1, \ldots, \alpha_{t+w+k})}$.
Step 3.	Let $h_i = x^l - p_i$ for $i = 1, 2, \ldots, t$, $h_{t+j} = x^m - q_j$ for $j = 1, \ldots, w$, and $h_{t+w+k} = x^n - v_k$ for $k = 1, \ldots, s$, set $I = \langle h_1, \ldots, h_{t+w+k},\ g \cdot y - f \rangle$, calculate the reduce Gröbner basis G of I using the function "groebner". Then compute $h(y) = G \cap Q[y]$. Set $m(y) = h(y)/lc(h(y))$, $n(y) = x + y$ divide $m(y)$ by $n(y)$ using the function "polynomialReduce" such that $m(y) = q(y)n(y) + r(y), \quad \partial(r(y)) < \partial(g(y))$. Write $m(y) = q(y)n(y) + r(x)$, and substitute $y = \alpha$ in the equation.
Return:	$\alpha, q(x), r(x)$.

Figure 1: Denominator rationalized algorithm.

$$
\begin{aligned}
f_1 &= y^6 - 18y^4 + 108y^2 - 144y - 200, \\
f_2 &= x_2^2 - 3, \\
f_3 &= 7128x_1 + 18x_2^3y^4 - 144x_2^3y^3 - 294x_2^3y^2 + 4756x_2^3y + 64x_2^3 - 3x_2y^5 - 30x_2y^2 - 14904x_2y + 4536x_2.
\end{aligned} \tag{47}
$$

Therefore, $G \cap Q[y] = \{y^6 - 18y^4 - 8y^3 + 108y^2 - 144y - 200\}$. Hence, the minimal polynomial of α in the rational number field Q is $m(y) = y^6 - 18y^4 - 8y^3 + 108y^2 - 144y - 200$.

Next, divide $m(y)$ by $x + y$ in the software Matlab using the built-in function "polynomialReduce." Then,

$$
m(y) = q(y) \cdot (x + y) + r(x), \tag{48}
$$

where

$$
\begin{aligned}
q(y) &= y^5 - xy^4 + y^3\left(x^2 - 18\right) + y^2\left(-x^3 + 18x - 8\right) + y\left(x^4 - 18x^2 + 8x + 108\right) + \left(-x^5 + 18x^3 - 8x^2 - 108x - 144\right), \\
r(x) &= x^6 - 18x^4 + 8x^3 + 108x^2 + 144x - 200.
\end{aligned} \tag{49}
$$

Substitute $\alpha = \sqrt{6} + \sqrt[3]{4}$ into

$$
\frac{1}{x + \alpha} = -\frac{q(\alpha)}{r(x)}, \tag{50}
$$

then

$$
\begin{aligned}
\frac{1}{x + \sqrt{6} + \sqrt[3]{4}} &= \frac{x^5 - 12x^3 + 4x^2 + 36x + 24}{r(x)} + \frac{-x^4 + 12x^2 + 8x - 36}{r(x)}\sqrt{6} + \frac{-x^4 - 4x + 36}{r(x)}\sqrt[3]{4} \\
&\quad + \frac{2x^3 + 36x + 8}{r(x)}\sqrt[3]{2} + \frac{2x^3 - 12x - 4}{r(x)}\sqrt{6}\sqrt[3]{4} + \frac{-6x^2 - 12}{r(x)}\sqrt[3]{2}\sqrt{6}.
\end{aligned} \tag{51}
$$

Remark 1. Note that $m = 2$ and $n = 3$, and we can use the method in Section 2 to do this example. Replace $a = 6$ and $b = 4$ and do all the steps as in Section 2, and we obtain

$$\begin{aligned}\frac{1}{x+\sqrt{6}+\sqrt[3]{4}} = \frac{1}{x^6-18x^4+8x^3+108x^2+144x-200}\Big(x^5-12x^3+4x^2+36x+24+\Big(-x^4+12x^2+8x-36\Big)\\ \cdot\sqrt{6}+\Big(-x^4-4x+36\Big)\cdot\sqrt[3]{4}+\Big(2x^3+36x+8\Big)\cdot\sqrt[3]{2}\\ +\Big(2x^3-12x-4\Big)\cdot\sqrt{6}\cdot\sqrt[3]{4}+\Big(-6x^2-12\Big)\cdot\sqrt{6}\cdot\sqrt[3]{2}\Big).\end{aligned} \tag{52}$$

Through the two examples above, it is easy to find that the results obtained by both methods are the same, which proves that the algorithm we designed is correct and the application scope is wider.

Remark 2. Using the method in Section 2 or the method of undetermined coefficients, the problem of denominator rationalization of fractions with the form as $1/(x + c_1 \cdot \sqrt[m]{a} + c_2 \cdot \sqrt[n]{b})$ may not be solved or can be solved but will take a long time. However, from Example 1, it is straightforward that this problem can be solved efficiently and simply according to the three steps of Denominator rationalized algorithm by using the software Matlab and Singular.

5. Conclusion

In this paper, using theory of Gröbner bases, we have achieved in rationalizing denominators for two types of fractions, especially rationalizing the denominator of the fraction with the form as

$$\frac{1}{x + c_1 \cdot \sqrt[m]{a} + c_2 \cdot \sqrt[n]{b}}, \tag{53}$$

where $m, n \in N^+, m, n \geq 2, a, b, c_1, c_2 \in Q, a, b > 0$. We have presented some interesting results and an executable algorithm on rationalizing the denominators for this type fractions. Furthermore, we have established an example to illustrate the effectiveness of the algorithm.

The method that we proposed on rationalizing denominators can be realized in computer system such as Maple and Singular, which makes the related computation more quick. Furthermore, it can improve the performance of the related algorithms. For example, we find that, for some new swarm intelligence algorithms proposed in recent years, such as monarch butterfly optimization (MBO) [27], earthworm optimization algorithm (EWA) [28], elephant herding optimization (EHO) [29], etc. the fraction of the form $1/(x + c_1 \cdot \sqrt[m]{a} + c_2 \cdot \sqrt[n]{b})$ occurs during the implementation of these algorithms. The denominator rationalized algorithm we proposed can rationalize the denominator of this type fraction, which can improve the accuracy of these algorithms on calculation.

Future work will investigate in rationalizing denominators of more general forms of radical fraction. The type of fractions discussed in the paper is unnested radical expressions, and the problem of the nested radical fractions is not involved here, which is also what we will do.

Data Availability

Data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (11871207 and 11971161).

References

[1] A. Berele and S. Catoiu, "Rationalizing denominators," *Mathematics Magazine*, vol. 88, no. 2, pp. 121–136, 2015.

[2] Schuette and H. Paul, "Rationalizing the denominator: why bother?" *Mathematics and Computer Education*, vol. 32, no. 1, pp. 19–29, 1998.

[3] D. C. De, "Learn to realize the "solution" engendered by skills: ideals from a "denominator rationalization" teaching short video," *The World of Junior School Student: China Version*, vol. 12, pp. 44–88, 2021.

[4] E. H. Neville, "Partial fractions rationalizing the denominator," *The Mathematical Gazette*, vol. 42, no. 342, pp. 261–266, 1958.

[5] H. B. Li, "A method for denominator rationalzation of higher radical fraction," *Middle-school Mathematics: China version*, vol. 6, pp. 23-24, 1982.

[6] L. J. Paradiso, "Discussions: rationalizing factors and the method of undetermined coefficients," *The American Mathematical Monthly*, vol. 36, no. 2, pp. 87–89, 1929.

[7] B. F. Caviness and R. J. Fateman, "Simplification of radical expressions," in *Proceedings of the 1976 ACM Symposium on Proceedings of the third ACM Symposium on Symbolic and Algebraic Computation*, pp. 329–338, Association for Computing Machinery, New York, NY, United States, August 1976.

[8] J. Zhou, "One way to rationalize the denominator for 1/(g(u)), the middle school teaching," *Mathematics version*, vol. 3, pp. 38–40, 1986.

[9] H. C. Ma, "On the discussion of denominator's rationalization of irrational expressions," *Journal of Yangzhou Teachers college(Natural Science): China version*, pp. 10–13, 1989.
[10] J. Liu, "Use the determinant to rationalize the denominator of an algebraic expression," *Journal of Mathematics: China version*, vol. 11, pp. 41-42, 2000.
[11] M. M. Tang, "How to rationalize the deominators of a kind of algebraic expressions by multinomials," *Journal of Zhuzhou Teachers College: China Version*, vol. 5, pp. 31–33, 2002.
[12] B. Buchberger, "A criterion for detecting unnecessary reductions in the construction of Gröbner bases," *Symbolic and Algebraic Computation*, vol. 72, pp. 3–21, 1979.
[13] B. Buchberger, *An Algorithm for Finding a Basis for the Redidue Class Ring of a Zero Dimensiomal Polynomial*, phD thesis, Universität Innsbruck, Institut für Mathematik, 1965.
[14] W. Adams and P. Loustaunau, *An Introduction to Gröbner Bases*, American Mathematical Society, New York, 1994.
[15] T. Becker and V. Weispfenning, *Gröbner Bases-A Computational Approach to Commutative Algebra*, Vol. 141, GTM Springer, , New York, 1993.
[16] B. Buchberger and F. Winkler, *Gröbner Bases and Application. London Mathematical Society Lecture Note Series*, Vol. 251, Cambridge University Press, , Cambridge, 1998.
[17] D. Li, J. Liu, and L. Zheng, "On the equivalence of multivariate polynomial matrices," *Multidimensional Systems and Signal Processing*, vol. 28, no. 1, pp. 225–235, 2017.
[18] D. Li, J. Liu, and L. Zheng, "A zero-dimensional valuation ring is 1-Gröbner," *Journal of Algebra*, vol. 484, pp. 334–343, 2017.
[19] Y. Liu and H. Su, "Some necessary and sufficient conditions for containment of second-order multi-agent systems with sampled position data," *Neurocomputing*, vol. 378, pp. 228–237, 2020.
[20] Y. Liu and H. Su, "Containment control of second-order multi-agent systems via intermittent sampled position data communication," *Applied Mathematics and Computation*, vol. 362, p. 124522, 2019.
[21] C. Xu, H. Xu, H. Su, and C. Liu, "Disturbance-observer based consensus of linear multi-agent systems with exogenous disturbance under intermittent communication," *Neurocomputing*, vol. 404, pp. 26–33, 2020.
[22] C. Xu, B. Li, and L. Yang, "Semi-global containment of discrete-time high-order multi-agent systems with input saturation via intermittent control," *IET Control Theory and Applications*, vol. 14, no. 16, pp. 2303–2309, 2020.
[23] D. A. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*, Springer, New York, 2007.
[24] S. Lang, "Algebraic number theory," *Graduate Texts in Mathematics*, Springer, New York, Second Edition, 1994.
[25] D. Eisenbud, *Commutative Algebra with a View towards Algebraic Geometry*, Springer-Verlag, New York, 1995.
[26] D. Wang, B. Xia, and Z. Li, *Computer Algebra*, pp. 181–183, Tsinghua University Press, Beijing, 2007.
[27] Y. H. Feng, S. Deb, G. G. Wang, and H. A. Alavi, "Monarch butterfly optimization: a comprehensive review," *Expert Systems with Applications*, vol. 168, p. 2021.
[28] S. K. Kanna, Rajesh, K. Sivakumar, and N. Lingaraj, "Development of deer hunting linked earthworm optimization algorithm for solving large scale traveling salesman problem," *Knowledge-Based Systems*, vol. 227, 2021.
[29] J. Li, H. Lei, H. A. Alavi, and G. G. Wang, "Elephant herding optimization: variants, hybrids, and applications," *Mathematics*, vol. 8, no. 9, p. 2020.

Hindawi
Complexity
Volume 2022, Article ID 5093277, 19 pages
https://doi.org/10.1155/2022/5093277

WILEY | Hindawi

Research Article

A Q-Learning-Based Parameters Adaptive Algorithm for Formation Tracking Control of Multi-Mobile Robot Systems

Chen Zhang,[1] **Wen Qin**,[1] **Ming-Can Fan**,[2] **Ting Wang**,[1] **and Mou-Quan Shen**[1]

[1]*College of Electrical Engineering and Control Science, Nanjing Tech University, Nanjing 211816, Jiangsu, China*
[2]*School of Mathematics and Statistics, Huizhou University, Huizhou 516007, Guangdong, China*

Correspondence should be addressed to Wen Qin; qinwen.wts@163.com

Received 10 September 2021; Revised 7 November 2021; Accepted 15 December 2021; Published 5 January 2022

Academic Editor: Miaomiao Wang

This paper proposes an adaptive formation tracking control algorithm optimized by Q-learning scheme for multiple mobile robots. In order to handle the model uncertainties and external disturbances, a desired linear extended state observer is designed to develop an adaptive formation tracking control strategy. Then an adaptive method of sliding mode control parameters optimized by Q-learning scheme is employed, which can avoid the complex parameter tuning process. Furthermore, the stability of the closed-loop control system is rigorously proved by means of matrix properties of graph theory and Lyapunov theory, and the formation tracking errors can be guaranteed to be uniformly ultimately bounded. Finally, simulations are presented to show the proposed algorithm has the advantages of faster convergence rate, higher tracking accuracy, and better steady-state performance.

1. Introduction

A multi-mobile robot system can present intelligent behaviours through mutual cooperation and achieve work efficiency and fault tolerance that a single individual cannot provide, so that it can complete some more difficult tasks. The coordinated formation control of multiple mobile robots has received extensive attention due to its important applications in the industrial and medical field [1]. The most existing control methods dealing with formation control problems of multiple mobile robots mainly include behaviour-based control [2], virtual structures [3], and leader-follower architecture [4–6]. As a decentralized control strategy, the leader-follower formation structure has become the preferred control strategy due to its simplicity and scalability and requires less computation and communication resources than other strategies [7]. The movement types of mobile robots are divided into omnidirectional mobile robot (OMR) (holonomic) and nonholonomic one [8]. Concerning the uncertainty and disturbance of robot docking, reference [9] proposed a novel robust containment architecture for nonholonomic mobile robot formations with docking capabilities, which realized multirobot formation maintenance/switching, docking, and collision avoidance. In [10], a dynamic control law was developed for the cooperative target encircling problem of multiple unicycle mobile robots subject to heterogeneous input disturbances generated by the linear exogenous system.

Since nonholonomic mobile robots have fewer controllable degrees of freedom (DOFs) than holonomic mobile robots, geometric constraints are introduced on the robot's motion. Common examples with incomplete constraints are unicycles and car-like wheeled mobile robots. In contrast, holonomic mobile robots have the same number of controllable DOFs and total DOFs, which makes them very flexible and being able to move within the workspace without geometric constraints (e.g., they can perform both rotation and lateral translation). A typical example of this class of vehicles is the omnidirectional robot with mechanical wheels. More details on the types and configurations of mobile robots can be found in [11]. References [12, 13] studied the formation control problem of multiple omnidirectional mobile robots. Reference [12]

developed a distributed adaptive control law for a multirobot system by obtaining information from moving targets through some mobile robots and using backstepping control technology in formation control. Reference [13] proposed an improved collision avoidance and formation control to configure a multirobot system optimized for omnidirectional visual simultaneous localization and mapping. However, the performance of the control schemes in [12, 13] will deteriorate when there exist uncertainties in the kinematics and dynamics of the omnidirectional mobile robots.

In the research of formation control of complex nonlinear systems, sliding mode control (SMC) is an effective robust controller for suppressing disturbances because of its excellent characteristics of being insensitive to system parameter changes and external disturbance when the system enters the sliding mode. In [14], a nonsingular fast terminal SMC was proposed, which can drive the tracking error to zero in finite time. Reference [15] investigated the leader-follower formation control for multiwheel mobile robots by combining a motion controller with a dynamic controller based on sliding mode. Considering the bounded external disturbance and parameter uncertainty of mobile robots, reference [16] proposed a dual-loop attitude tracking robust controller for mobile robots, using SMC with modified arrival law to ensure that the actual speed converges within a finite time. The main disadvantage of SMC is chattering phenomenon due to the discontinuity of the control law. In order to alleviate chattering, adaptive SMC [17] and higher order SMC [18, 19] have been proposed. However, these control methods may bring serious chattering even leading to instability when the system is exposed to a dynamic environment with large uncertainties and disturbances. Moreover, the traditional SMC is conservative to some extent since it ignores the information of uncertainties and disturbances. An effective way to solve this problem is using the disturbance estimation and compensation to decrease the conservatism and improve the control performance. Reference [20] proposed a disturbance observer and super-twisting SMC for the multirobot formation. Reference [21] designed an adaptive high-gain observer for the robot to estimate the nonlinearity that appears in the dynamics of the wheeled robot. In [22], the active disturbance rejection control technology was used to estimate the external disturbance in the inner loop of the double closed-loop strategy. On the other hand, the velocities of omnidirectional mobile robots cannot often be measured due to the lack of sensors which are needed in controller design.

Extended state observer (ESO) based active disturbance rejection control (ADRC), proposed by Han [23], is a powerful tool to cope with uncertainties and external disturbances. The key idea of ADRC is that the total disturbances (including internal uncertain dynamics, cross-coupling, and external disturbances), regarded as an extended state of the system, can be estimated by the ESO and then compensated in the control signal. As a matter of fact, the ESO is a state observer to estimate both the system states and the total disturbances. Considering the advantage of it, ESO is adopted to estimate the total disturbances and then followed by the ESO-based controller constructed to compensate it. Reference [24] employed a nonlinear extended state observer (NESO) to estimate unknown states as well as uncertainties and designed a robust finite-time tracking control scheme to handle wheeled mobile robots with parameter uncertainties and disturbances. Reference [25] used NESO-based estimation and compensation signals into the closed-loop control method; an NESO-based decoupling control method was then proposed to solve the attitude control problem of hypersonic gliding vehicles. In [26], an NESO was used to estimate the system uncertainties, and a saturation-resistant adaptive SMC was designed based on the estimated values to achieve robust trajectory tracking for a wheeled mobile robot. However, it is difficult to find appropriate nonlinear functions to design an NESO in practice. For convenience of theoretical analysis, Gao proposed a linearized and bandwidth-based linear ADRC (LADRC) to simplify parameter tuning and standardize controller tuning [27]. Linear ESO (LESO) with nested inner and outer loops was used in [28] to actively estimate and eliminate generalized interference and improve the estimation accuracy in various practical models. The proposal of LADRC makes the design and adjustment of the controller easier and more effective, and the tracking error will be more decreased than some classic control structures [16], which greatly promotes the engineering application of ADRC. However, the LADRC is loss of design flexibility because the LADRC parameters are adjusted based on bandwidth. A general LADRC with more tuning parameters was proposed in [29], and nowadays LADRC has been employed to various cases in engineering application and becomes much more popular [30, 31].

Reinforcement learning (RL) has been developed rapidly in recent years. As one of the important algorithms of RL, Q-learning is off-policy, tabular, model-free, and based on temporal-difference methods [32]. It has the advantages of not relying on models and having good learning effects for complex systems. In order to improve the control performance, some scholars combined Q-learning with PID control and proposed many excellent control methods [33, 34]. In the autonomous underwater vehicle system, reference [35] proposed a Q-learning PID controller based on RBFNN to improve control performance, in which Q-learning neural network was used to adaptively optimize control parameters. Reference [36] combined model-based Q-learning into the predictive control setting to provide closed-loop stability in online learning and ultimately improve the performance of the limited range controller. Chen proposed an adaptive auto-disturbance-rejection controller parameter adaptation method based on Q-learning for ship heading control with multiple uncertainties due to wind, wave, and current interference [37].

Inspired by the above statements, this paper investigates the formation tracking control of a multiple omnidirectional mobile robot system. The considered mobile robots have internal modelling uncertainties and external

disturbances (considered as total disturbances). To handle the disturbances, an LESO is constructed, and the total disturbances can be effectively estimated through the ESO. Then, on the basis of distributed formation tracking control architecture, an LESO-based SMC (LESO-SMC) is designed for each OMR to ensure that the observer errors and the formation neighbourhood errors are uniformly ultimately bounded (UUB). However, LESO based control is not widely used in practice because there are not adequate methods for LESO parameter adjustment. In view of this, and furthermore for obtaining better control performance, Q-learning is employed to optimize the bandwidth parameters of LESO and the control parameters of SMC and to avoid the complex parameter tuning process. In addition, a simulation example is given to verify the effectiveness of the proposed method.

The main features of the proposed methods are summarized as follows:

(1) An LESO is constructed to estimate the 'total disturbances' in real time, including both internal parameter uncertainties and external disturbances, and then an LESO-SMC based formation protocol is developed for the OMR system. The LESO provides distinctly better robustness against 'total disturbances' by providing accurate input variables to the control system, including the states of the mobile robot at each order, as well as the extended state representing the 'total disturbances'. Then a corresponding improved strategy on the SMC is made to compensate the influence of the 'total disturbances', which ensures a faster convergence rate and decreases the conservatism of the traditional SMC.

(2) To take full advantages of the LESO-SMC, an adaptive method of LESO-SMC parameters optimized by Q-learning algorithm is proposed in the formation tracking control of the OMR system. Q-learning is introduced to perform online parameter adaptation (including the observer, sliding mode variables, and controller parameters), which exhibits better formation tracking performance and avoids the complex parameter tuning process.

The organization of this paper is as follows. In Section 2, the dynamic model and some preliminary knowledge are outlined. In Section 3, the proposed controller design based on LESO and SMC is presented. Both the Q-learning algorithm and the Q-learning parameter tuning process are also introduced. The results of numerical simulations are discussed in Section 4, followed by the conclusion of this paper.

Notations: A^T and A^{-1} represent the transpose and inverse of matrix A, respectively. R^n represents n-dimensional real column vector set. I_n denotes an $n \times n$ unit matrix. 0_n denotes an $n \times n$ zero matrix. $\otimes$ stands for Kronecker product. $\text{sign}(\cdot)$ is the sign function. $\|\cdot\|$ represents the Euclidean norm. $\text{diag}x_1, x_2, \ldots, x_n$ denotes the diagonal matrix with its diagonal entries being $x_1, x_2, \ldots, x_n$. $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ represent the maximum and the minimum eigenvalues of matrix A, respectively.

2. Problem Formulation and Preliminaries

2.1. Dynamic Model. The Euler–Lagrange equation of motion can be used to describe the dynamic behaviour of an OMR. The dynamic model of the ith OMR can be described as [31]

$$M_i\ddot{q}_i + C_i\dot{q}_i + G_i = \tau_i, \tag{1}$$

where $i = 1, \ldots, n$, $q_i = [x_i, y_i, \theta_i]^T \in R^3$ represents the position and orientation angle of the ith robot in the world coordinate frame. M_i is the inertia matrix, C_i is the Coriolis and centrifugal term, G_i is the gravitational force, and τ_i denotes the control input. It is assumed that the robots move on flat ground, where gravitational force G_i is 0.

Considering the unknown dynamic disturbances and model uncertainties, a new dynamic model is obtained as follows:

$$(M_i + \Delta M_i)\ddot{q}_i + (C_i + \Delta C_i)\dot{q}_i = \tau_i + d_i, \tag{2}$$

where $\Delta M_i \in R^{3\times3}$ and $\Delta C_i \in R^{3\times3}$ denote uncertainty terms, and $d_i \in R^3$ denotes the unknown disturbance term.

The above equation can be rewritten as

$$M_i\ddot{q}_i + C_i\dot{q}_i + \Delta_i = \tau_i + d_i, \tag{3}$$

where $\Delta_i = \Delta M_i\ddot{q}_i + \Delta C_i\dot{q}_i$ denotes the uncertain term.

Assumption 1. The unknown disturbance term d_i satisfies $\|d_i^{(j)}\| \le \lambda_1$, $j = 0, 1, 2$, where λ_1 is an unknown positive constant.

Assumption 2. The uncertain terms ΔM_i and ΔC_i are bounded. Thus, the uncertain term Δ_i satisfies $\|\Delta_i^{(j)}\| \le \lambda_2$, $j = 0, 1, 2$, where λ_2 is an unknown positive constant.

As for the LESO design in the following section, the total disturbance is extended as a new state. Define $x_{i,1} = q_i$, $x_{i,2} = \dot{q}_i$, $x_{i,3} = f_i = M_i^{-1}(d_i - \Delta_i) \in R^3$, where $x_{i,3}$ is the extended state of the total disturbances f_i. The dynamic model (3) can be transformed into the following system:

$$\begin{cases} \dot{x}_{i,1} = x_{i,2}, \\ \dot{x}_{i,2} = x_{i,3} + M_i^{-1}\tau_i - M_i^{-1}C_ix_{i,2}, \\ \dot{x}_{i,3} = \dot{f}_i. \end{cases} \tag{4}$$

The dynamic equation of the virtual leader is

$$\begin{cases} \dot{x}_l = v_l, \\ \dot{v}_l = u_l, \end{cases} \tag{5}$$

where $x_l = \left[x_{lx}, x_{ly}, x_{l\theta}\right]^T \in R^3$ and $v_l = \left[v_{lx}, v_{ly}, v_{l\theta}\right]^T \in R^3$ indicate the position and velocity of the virtual leader, respectively. u_l is the control input of the virtual leader which can be obtained by some followers.

Assumption 3. The derivative of the total disturbances f_i is bounded by an unknown constant ρ_i, i.e., $\|\dot{f}_i\| \le \rho_i$.

Remark 1. Note that f_i in the dynamic model (4) indicates an unknown term, such as external disturbances and model uncertainties for the mechanism of OMRs. In practical applications, the total disturbances mainly include wheel-ground sliding, modelling uncertainty due to robot load variation, etc.

Remark 2. In practice, both the speed of the DC motor which drives the OMR forward and its derivative have upper bounds, i.e., q_i, $\dot{q}_i$, and $\ddot{q}_i$ are all bounded. Δ_i is related to $\dot{q}_i$ and $\ddot{q}_i$, so we can conclude that Δ_i and its derivative are bounded. Similarly, f_i and its time derivative are bounded as well. Therefore, the assumptions of Δ_i and f_i in Assumptions 1 and 2 are reasonable.

2.2. Algebraic Graph Theory. Consider an omnidirectional mobile robot system consisting of one virtual leader and n followers. Assume that each robot is a node, and the information exchange among follower robots can be described by a directed graph G. The graph $G = (V, E, A)$ is composed of the node set $V = \{v_1, \ldots, v_n\}$, the edge set $E = \{(v_i, v_j) \in V \times V\}$, and the adjacency matrix $A = (a_{ij}) \in R^{n\times n}$. If there is an edge between agents i and j, i.e., $(v_i, v_j) \in E$, then $a_{ij} > 0$; otherwise $a_{ij} = 0$. The set of neighbours of node v_i is denoted by $N = \{j: (v_i, v_j) \in E\}$. The in-degree of node v_i is defined as $d_i = \sum_{j\in N} a_{ij}$. Then, the in-degree matrix of digraph G is $D = \mathrm{diag}\{d_1, \ldots, d_n\}$, and the Laplacian matrix of digraph G is $L = D - A$. A path in graph G from v_i to v_j is a sequence of distinct vertices starting with v_i and ending with v_j such that consecutive vertices are adjacent. Graph G is connected if there is a path between any two vertices. Graph G contains a directed spanning tree if there is a vertex (the root node) which can reach all the other vertices through a directed path. The virtual leader's adjacency weight matrix $B = \mathrm{diag}\{b_1, \ldots, b_n\}$, where b_i represents the topological weight of the communication between agent i and the virtual leader, $b_i > 0$, if there is communication between the agent i and the virtual leader; otherwise, $b_i = 0$.

Assumption 4. For the considered multiple mobile robot systems (2), graph G for the n follower robots contains a directed spanning tree; i.e., there is a vertex (the root node) which can reach all the other vertices through a directed path.

Lemma 1 (see [38]). *If G is a directed graph which contains a directed spanning tree and at least the root-node agent has access to the virtual leader, then the matrix $L + B$ is of full rank.*

3. Main Results

In this section, in order to achieve better formation tracking control performance, the LESO-SMC scheme will be designed for system (3) in the presence of the unknown disturbances and model uncertainties, such that all follower OMRs can track the virtual leader with the given formation configuration in advance and maintain the same speed with the virtual leader. Furthermore, a parameter adaptation method based on Q-learning algorithm is involved in LESO-SMC to avoid the complex parameter tuning process, which displays a better formation tracking control performance.

3.1. Linear Extended State Observer Design. In this section, we use the LESO to estimate the OMR's total disturbances f_i, which include unknown disturbances and model uncertainties.

The LESO for system (4) is designed as follows:

$$\begin{cases} \dot{\hat{x}}_{i,1} = \hat{x}_{i,2} + \beta_1\left(x_{i,1} - \hat{x}_{i,1}\right), \\ \dot{\hat{x}}_{i,2} = \hat{x}_{i,3} + M_i^{-1}\tau_i - M_i^{-1}C_i\hat{x}_{i,2} + \beta_2\left(x_{i,1} - \hat{x}_{i,1}\right), \\ \dot{\hat{x}}_{i,3} = \beta_3\left(x_{i,1} - \hat{x}_{i,1}\right), \end{cases} \tag{6}$$

where $\hat{x}_{i,1}, \hat{x}_{i,2}, \hat{x}_{i,3}$, and $\hat{h}_i$ are the estimations of $x_{i,1}, x_{i,2}, x_{i,3}$, and h_i, respectively, and β_k $(k = 1, 2, 3)$ is the observer gain to be determined. $\beta_1 = 3\omega_0, \beta_2 = 3\omega_0^2, \beta_3 = \omega_0^3$, where ω_0 is the ESO bandwidth [39].

Define the estimation errors as follows: $\tilde{x}_{i,j} = x_{i,j} - \hat{x}_{i,j}$ $(j = 1, 2, 3)$; then the estimation error equation is given by

$$\begin{cases} \dot{\tilde{x}}_{i,1} = \tilde{x}_{i,2} - 3\omega_0\tilde{x}_{i,1}, \\ \dot{\tilde{x}}_{i,2} = \tilde{x}_{i,3} + M_i^{-1}C_i\tilde{x}_{i,2} - 3\omega_0^2\tilde{x}_{i,1}, \\ \dot{\tilde{x}}_{i,3} = \dot{f}_i - \omega_0^3\tilde{x}_{i,1}. \end{cases} \tag{7}$$

Let $\xi_i = [\xi_{i,1}, \xi_{i,2}, \xi_{i,3}]^T = \left[\tilde{x}_{i,1}\ \tilde{x}_{i,2}/\omega_0\ \tilde{x}_{i,3}/\omega_0^2\right]^T \in R^9$; equation (7) can be rewritten as

$$\dot{\xi}_i = \omega_0 A_1\xi_i + \frac{N_2M_i^{-1}C_i\tilde{x}_{i,2}}{\omega_0} + \frac{N_1\dot{f}_i}{\omega_0^2}, \tag{8}$$

where $A_1 = \begin{bmatrix} -3I_3 & I_3 & 0_3 \\ -3I_3 & 0_3 & I_3 \\ -I_3 & 0_3 & 0_3 \end{bmatrix} \in R^{9\times 9}$ is Hurwitz, $N_1 = \begin{bmatrix} 0_3 \\ 0_3 \\ I_3 \end{bmatrix}$, $N_2 = \begin{bmatrix} 0_3 \\ I_3 \\ 0_3 \end{bmatrix}$.

Slightly different to the proof in [40], the convergence analysis of LESO (6) is given below.

Lemma 2. *Considering the estimation error dynamics (8), the LESO (6) is bounded stable if the observer bandwidth ω_0 is designed to satisfy the condition $\omega_0 - 2\|P\| \cdot \|\kappa\omega_0\|/\omega_0 - 1 > 0$.*

Proof. Since matrix A_1 is Hurwitz, there exists a positive definite matrix P_1 which satisfies $A_1^TP_1 + P_1A_1 = -I_9$.

$$\left\|N_2M_i^{-1}C_i\tilde{x}_{i,2}\right\| \le \kappa\left\|\tilde{x}_{i,2}\right\| \le \kappa\omega_0\left\|\xi_{i,2}\right\|, \tag{9}$$

where κ is an unknown constant.

Consider a Lyapunov function candidate for (8) as

$$V_{eso} = \sum_{i=1}^{n} \xi_i^T P_1 \xi_i. \tag{10}$$

Then we obtain the time derivative of V_{eso} as

$$\begin{aligned}
\dot{V}_{eso} &= \sum_{i=1}^{n}\left(\dot{\xi}_i^T P_1 \xi_i + \xi_i^T P_1 \dot{\xi}_i\right) \\
&= \left(\sum_{i=1}^{n} \omega_0 \xi_i^T A_1^T P_1 \xi_i + \frac{\left(N_2 M_i^{-1} C_i \tilde{x}_{i,2}\right)^T P_1 \xi_i}{\omega_0} + \frac{\left(N_1 \dot{f}_i\right)^T P_1 \xi_i}{\omega_0^2} + \omega_0 \xi_i^T P_1 A_1 \xi_i + \frac{\xi_i^T P_1 \left(N_2 M_i^{-1} C_i \tilde{x}_{i,2}\right)}{\omega_0} + \frac{\xi_i^T P_1 \left(N_1 \dot{f}_i\right)}{\omega_0^2}\right) \\
&= \sum_{i=1}^{n}\left(-\omega_0 \xi_i^T \xi_i + 2\frac{\xi_i^T P_1 N_2 M_i^{-1} C_i \tilde{x}_{i,2}}{\omega_0} + 2\frac{\xi_i^T P_1 N_1 \dot{f}_i}{\omega_0^2}\right) \\
&= -\omega_0 \xi^T \xi + 2\frac{\xi^T P \overline{N} M_i^{-1} C_i \tilde{x}_{i,2}}{\omega_0} + 2\frac{\xi^T P N \dot{f}}{\omega_0^2} \\
&\le -\left(\omega_0 - 2\frac{\|P\| \cdot \|\kappa\omega_0\|}{\omega_0}\right)\|\xi\|^2 + 2\frac{\xi^T P N \dot{f}}{\omega_0^2},
\end{aligned} \tag{11}$$

where

$$\begin{aligned}
\xi &= \left[\xi_1^T, \ldots, \xi_n^T\right]^T, \\
P &= 1_n \otimes P_1, \\
N &= 1_n \otimes N_1, \\
\overline{N} &= 1_n \otimes N_2, \\
1_n &= [1, 1, \ldots, 1]^T \quad \in R^n, \\
\dot{f} &= \left[\dot{f}_1^T, \ldots, \dot{f}_n^T\right]^T.
\end{aligned} \tag{12}$$

Using Young's inequality, one has

$$2\frac{\xi^T P N \dot{f}}{\omega_0^2} \le \|\xi\|^2 + \frac{\|PN\|^2 \|\dot{f}\|^2}{\omega_0^4}. \tag{13}$$

From Assumption 3, we have $\dot{f} \le \rho$, where ρ is an unknown constant. It can be obtained from (11) that

$$\dot{V}_{eso} \le -\left(\omega_0 - 2\frac{\|P\| \cdot \|\kappa\omega_0\|}{\omega_0} - 1\right)\|\xi\|^2 + \frac{\|PN\|^2 \rho^2}{\omega_0^4}. \tag{14}$$

If $\omega_0 - 2\|P\| \cdot \|\kappa\omega_0\|/\omega_0 - 1 > 0$, then $\|PN\|^2\rho^2/\omega_0^4$ is bounded; hence the proposed LESO is bounded stable. The proof is completed.

To achieve a better tracking control performance, in this section, an LESO-SMC-based formation control scheme will be proposed to ensure the tracking performance based on the dynamic models introduced in Section 2. □

3.2. Sliding Mode-Based Formation Controller Design. Define the system neighbourhood errors as

$$\begin{cases}
e_{i,1} = \sum_{j \in N} a_{ij}\left(\left(\hat{x}_{i,1} - \delta_i\right) - \left(\hat{x}_{j,1} - \delta_j\right)\right) + b_i\left(\hat{x}_{i,1} - x_l - \delta_i\right), \\
e_{i,2} = \sum_{j \in N} a_{ij}\left(\hat{x}_{i,2} - \hat{x}_{j,2}\right) + b_i\left(\hat{x}_{i,2} - v_l\right),
\end{cases} \tag{15}$$

where $\delta_i = [\delta_{ix}, \delta_{iy}, \delta_{i\theta}]^T \in R^3$ denotes the desired relative position for the ith robot and the virtual leader.

According to (15), we design the ESO-based sliding mode surface of the overall formation for the ith agent as

$$S_i = \mu e_{i,1} + e_{i,2}, \tag{16}$$

where μ is a positive constant.

The time derivative of S_i is

$$\begin{aligned}
\dot{S}_i &= \mu \dot{e}_{i,1} + \dot{e}_{i,2} \\
&= \mu e_{i,2} + \sum_{j \in N} a_{ij}\left(\dot{\hat{x}}_{i,2} - \dot{\hat{x}}_{j,2}\right) + b_i\left(\dot{\hat{x}}_{i,2} - u_l\right).
\end{aligned} \tag{17}$$

The formation tracking controller based on ESO-SMC algorithm for the ith agent is designed as

$$\tau_i = M_i\left(M_i^{-1} C_i \hat{x}_{i,2} - \hat{x}_{i,3} + u_l + \left(\sum_{j \in N} a_{ij} + b_i\right)^{-1} \cdot \left(\sum_{j \in N} a_{ij}\tau_j - \mu e_{i,2} - c_1 S_i - c_2 \mathrm{sgn}\left(S_i\right)\right)\right), \tag{18}$$

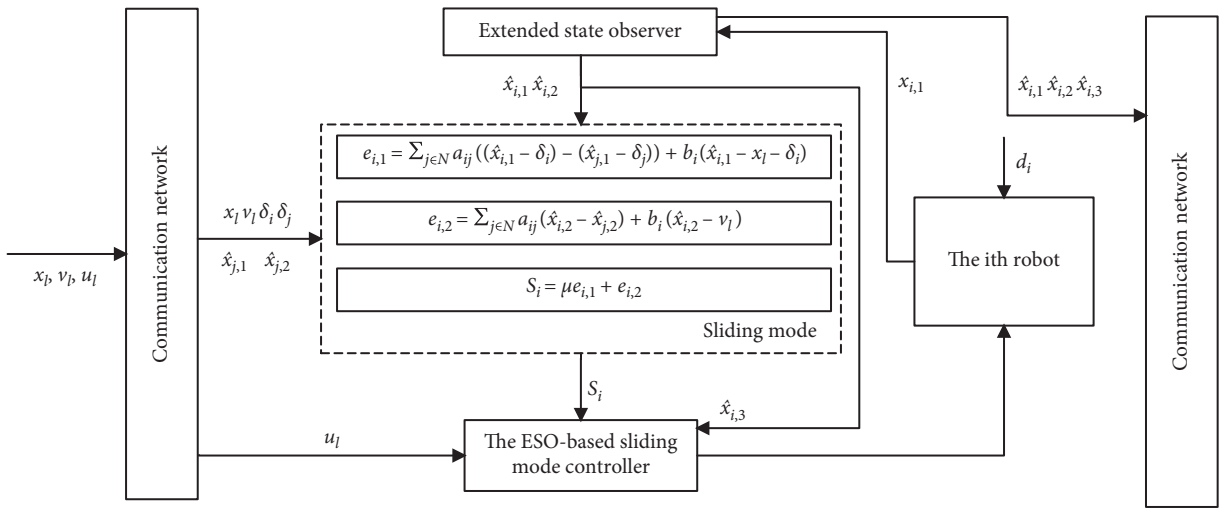


Figure 1: The proposed control block diagram.

where c_1 and c_2 are positive constants, and u_l is the control input of the virtual leader which can be obtained by some followers.

The designed control structure block diagram is shown in Figure 1.

3.3. Stability Analysis of State Tracking Error Dynamics. We have introduced LESO to estimate the states and uncertainties of each order of OMR; next we will propose the SMC-based formation controller.

Theorem 1. *Consider the OMR system composed of (4) and (5) that satisfies Assumptions 1–4. Under the adaptive formation tracking control algorithm optimized by Q-learning scheme with the LESO (6), the SMC (18) and the Q-learning-based parameter adaptive algorithm (see Section 3.4), the estimation errors of the proposed LESO, and the formation tracking errors are guaranteed to be UUB.*

Similar to [41], we comprehensively consider the closed-loop system composed of the estimation errors of the observer (6) and the sliding variable (16). The stability analysis of the closed-loop system is given below.

Proof. Define the positive definite Lyapunov function as

$$V=\frac{1}{2}\sum_{i=1}^{n}\xi_i^TP_1\xi_i+\frac{1}{2}\sum_{i=1}^{n}S_i^TS_i, \tag{19}$$

with the time derivative of the second term in the right-hand side of (19) being

$$\begin{aligned}\sum_{i=1}^{n}S_i^T\dot{S}_i&=\sum_{i=1}^{n}S_i^T\left(\mu\dot{e}_{i,1}+\dot{e}_{i,2}\right)\\&=\sum_{i=1}^{n}S_i^T\left(\mu e_{i,2}+\sum_{j\in N}a_{ij}\left(\dot{\hat{x}}_{i,2}-\dot{\hat{x}}_{j,2}\right)+b_i\left(\dot{\hat{x}}_{i,2}-u_l\right)\right).\end{aligned} \tag{20}$$

The system neighbourhood error (15) can be expressed in matrix form as

$$\begin{cases}e_1=(L+B)\otimes I_3\cdot\left(\widehat{x}_1-1_n\otimes x_l-\delta\right),\\e_2=(L+B)\otimes I_3\cdot\left(\widehat{x}_2-1_n\otimes v_l\right),\end{cases} \tag{21}$$

where $e_1=[e_{1,1}^T,\ldots,e_{n,1}^T]^T\in R^{3n}$, $e_2=[e_{1,2}^T,\ldots,e_{n,2}^T]^T\in R^{3n}$, $\widehat{x}_1=[\widehat{x}_{1,1}^T,\ldots,\widehat{x}_{n,1}^T]^T\in R^{3n}$, $\widehat{x}_2=[\widehat{x}_{1,2}^T,\ldots,\widehat{x}_{n,2}^T]^T\in R^{3n}$, $\delta=[\delta_1^T,\ldots,\delta_n^T]^T\in R^{3n}$, $1_n=[1,1,\ldots,1]^T\in R^n$.

The sliding surface (16) can be rewritten as

$$S=\mu e_1+e_2. \tag{22}$$

Then the formation tracking controller (18) can be rewritten as

$$\tau=M\left(M^{-1}C\widehat{x}_2-\widehat{x}_3+1_n\otimes u_l+(L+B)^{-1}\otimes I_3\cdot\left(-\mu e_2-c_1S-c_2\mathrm{sgn}(S)\right)\right), \tag{23}$$

where $M=\mathrm{diag}\{M_i\}$, $C=\mathrm{diag}\{C_i\}$.

Substituting (21)–(23) into (20), one can obtain

$$\begin{aligned}\sum_{i=1}^{n}S_i^T\dot{S}_i&=S^T\left(\mu e_2+(L+B)\otimes I_3\cdot\left(\widehat{x}_3+M^{-1}\tau-1_n\otimes u_l\right)\right)\\&=-c_1S^TS-c_2S^T\mathrm{sgn}(S)\\&=-c_1\|S\|^2-c_2\|S\|<0.\end{aligned} \tag{24}$$

Substituting (14) and (24) into (19), then we have

$$\begin{aligned}\dot{V}&\le-\frac{1}{2}\left(\omega_0-2\frac{\|P\|\cdot\|\kappa\omega_0\|}{\omega_0}-1\right)\|\xi\|^2+\frac{\|PN\|^2\rho^2}{2\omega_0^4}\\&\quad-c_1\|S\|^2-c_2\|S\|\quad\le-\varphi^T\Gamma\varphi+\varsigma,\end{aligned} \tag{25}$$

where $\varphi = [\xi^T, S^T]$, $\Gamma = \begin{bmatrix} \Gamma_1 & 0 \\ 0 & \Gamma_2 \end{bmatrix}$, $\varsigma = \|PN\|^2\rho^2/2\omega_0^4 - c_2\|S\|$, with $\Gamma_1 = \operatorname{diag}\{\underbrace{\eta, \ldots, \eta}_{n}\}$, $\eta = 1/2(\omega_0 - 2\|P\| \cdot \|\kappa\omega_0\|/\omega_0 - 1)$, $\Gamma_2 = \operatorname{diag}\{\underbrace{c_1, \ldots, c_1}_{n}\}$. Since Γ_1 and Γ_2 are positive definite, one has

$$\dot{V} \le -\lambda_{\min}(\Gamma)\left(\|\xi\|^2 + \|S\|^2\right) + \varsigma \le -\lambda_{\min}(\Gamma)\left(\frac{1}{\lambda_{\max}(P)}\xi^T P\xi + S^T S\right) + \varsigma \le -\chi V + \varsigma, \tag{26}$$

where $\chi = 2\lambda_{\min}(\Gamma)\min\{1, 1/\lambda_{\max}(P)\}$.

Based on Gronwall's inequality, one has

$$V(t) \le V(0)e^{-\chi t} + \frac{\varsigma}{\chi}\left(1 - e^{-\chi t}\right). \tag{27}$$

Since $V(t) \ge 1/2\xi^T P\xi \ge 1/2\lambda_{\min}(P)\|\xi\|^2$, one has $1/2\lambda_{\min}(P)\|\xi\|^2 \le V(0)e^{-\chi t} + \varsigma/\chi(1 - e^{-\chi t})$.

Hence $\|\xi\| \le \sqrt{2[V(0)\chi e^{-\chi t} + \varsigma(1 - e^{-\chi t})]/\chi\lambda_{\min}(p)}$; then we obtain the bound of the observer error as

$$\|\xi\| \le \sqrt{\frac{2\varsigma}{\chi\lambda_{\min}(P)}}, \quad t \longrightarrow \infty. \tag{28}$$

Since $V(t) \ge 1/2S^T S = 1/2\|S\|^2$, one has $\|\xi\| \le \sqrt{2[V(0)\chi e^{-\chi t} + \varsigma(1 - e^{-\chi t})]/\chi}$. Similarly, we can conclude that $\|S\| \le \sqrt{2\varsigma/\chi}$, $t \longrightarrow \infty$. Therefore ξ and S are UUB. By choosing appropriate parameters, the neighbourhood error can be attracted into a small stable region containing the origin. The proof is completed.

3.4. Q-Learning Based Parameter Optimization Process. The learning process of Q-learning is of continuous environmental interaction. First, at time t, select an action value $\overline{A}_t$. Then the agent will transfer from the original state $\overline{S}_t$ to a new state $\overline{S}_{t+1}$ with a probability of $P(\overline{S}_{t+1}|\overline{S}_t, a)$. At this time, due to environmental interaction, the agent can get a feedback return R, then the time variable is updated, and the agent restarts the above steps in the new state until the optimal strategy is finally obtained. The Q-learning algorithm is shown below.

Parameters: $\alpha \in (0, 1]$, small $\varepsilon > 0$, $\gamma \in (0, 1]$.

Step 1. For each state-action pair (s, a), $\forall s \in \overline{S}$, $a \in \overline{A}$, initialize the table entry $Q(s, a)$ to zero. Observe the current state s.

Step 2. Loop for each episode: Select an action a for state s using ε-greedy policy and execute it, use the same method to choose a_t for state s_t, then receive immediate reward r_t, and observe the new state s_{t+1}. Update the table entry for $Q(s_t, a_t)$ as follows:

$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_t + \gamma \max Q \quad (s_{t+1}, a) - Q(s_t, a_t)]$, $s_t \leftarrow s_{t+1}$, until s is the terminal state and all of $Q(s, a)$ are convergent.

Step 3. Output the final policy $\pi(s) = \arg\max Q(s, a)$.

In this paper, we consider the application of Q-learning ideas to the optimization of observers and controllers. Regarding the formation error and the derivative of the error as the state of Q-learning, dividing the controller parameters selection into a reasonable range and combining the divided values as an optional action, a Q-learning algorithm for observer and controller parameters optimization can be obtained. The specific steps of the algorithm are as follows:

(a) States set are composed of $\overline{S}_t(\varepsilon_i)$, $(i = 1, 2)$.

(b) Actions set are described as $\omega_0 \in [2, 4]$, $c_1 \in [5, 7]$, $c_2 \in [5, 7]$, $\mu \in [3, 5]$. Reasonably divide the interval of the four parameters and randomly select and combine them to obtain the actions set.

(c) In this paper, the stage performance index, i.e., the average of the squared errors of 10 sampling times, is used to design the reward function. Rewards are given as follows:

$$\begin{cases} J = \sum_{k=1}^{10} \dfrac{e_{i,1}^2(k)}{10}, \\ r = \begin{cases} 100, & J \le \sigma, \\ -100, & J > \sigma. \end{cases} \end{cases} \tag{29}$$

Perform the learning of the value function Q according to the algorithm flow mentioned above. Then we get the learned Q table and the optimal strategy $\pi(s) = \arg\max Q(s, a)$ for online parameter selection.

There are 3 termination conditions as follows. If any one of them is fulfilled then the Q-value training is terminated.

(a) It is not desired in practice that the formation error in the control is too large, which will make little sense to continue to iteratively calculate the Q table. So when $|e_{i,1}| > 10$, the training is terminated and reinitialized.

(b) The control process has reached a steady state; then the training is terminated, i.e., $|e_{i,1}| < 0.001$ and $|e_{i,2}| < 0.01$.

(c) The simulation time $t = 500s$; then the training is terminated.

Denote L_t as the training times. When $L_t = 800$, the training is terminated and the trained Q table can be obtained for online control.

Remark 3. Both the formation error and the error variation rate are divided into 7 linguistic values, {NB, NM, NL, Z, PL, PM, PB}, so that the number of state set elements is 49. Choose a reasonable range for each element in the parameter set $\Pi = [\omega_0 c_1 c_2 \mu]$. Based on the experience of adjusting the LESO-SMC controller parameters, set $\omega_0 \in [2, 4]$, $c_1 \in [5, 7]$, $c_2 \in [5, 7]$, $\mu \in [3, 5]$. Let $p_1 = p_2 = p_3 = p_4 = 3$; i.e., each parameter has 3 possible values; then the number of equivalent selectable actions is $\overline{A}_t = 81$. Therefore, the value function matrix $Q \in R^{81\times 49}$ for the formation system.

Remark 4. The introduction of Q-learning avoids the process of selecting and optimizing the controller, the observer, and the sliding mode parameters. After Q-learning optimization, the convergence performance of the controller and observer can be guaranteed.

4. Numerical Simulations

In this section, numerical simulation examples are used to illustrate the previous conclusions and the effectiveness of the proposed control scheme. Consider a scenario where a multiagent systems composed of three OMRs (followers) are simultaneously tracking a preset target (the virtual leader).

The communication topology G is given in Figure 2. The corresponding Laplacian matrix L and the adjacency weight matrix B can be described as

$$L = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ -1 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{30}$$

Each dynamics model of the three OMRs can be described by an Euler–Lagrange equation as follows:

$$M_i \ddot{q}_i + C_i \dot{q}_i + \Delta_i = \tau_i + d_i, \quad i = 1, 2, 3, \tag{31}$$

where $M_i = \begin{bmatrix} m & 0 & md \sin\theta_i \\ 0 & m & -md\cos\theta_i \\ md\sin\theta_i & 0 & I_v \end{bmatrix}$, $C_i = \begin{bmatrix} 0 & 0 & m\, d\dot{\theta}_i \cos\theta_i \\ 0 & 0 & m\, d\dot{\theta}_i \sin\theta_i \\ 0 & 0 & 0 \end{bmatrix}$. Physical parameters and their values of the OMR are shown in Table 1. (x_i, y_i) and θ_i denote the positions and the orientation angle of the ith OMR in the x and y directions in the ground coordinate system, respectively. $q_i = [x_i, y_i, \theta_i]^T \in R^3$ represents the position and orientation angle of the ith robot in the world coordinate frame, $\dot{q}_i = [\dot{x}_i, \dot{y}_i, \dot{\theta}_i]^T \in R^3$ represents the linear velocity and angular velocity of the ith robot in the world coordinate frame, and $\ddot{q}_i = [\ddot{x}_i, \ddot{y}_i, \ddot{\theta}_i]^T \in R^3$ represents the linear acceleration and angular acceleration of the ith robot in the world coordinate frame.

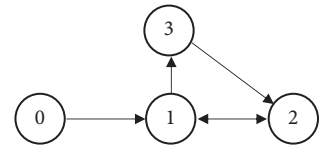


Figure 2: The communication topology of the systems. The root node 0 denotes the virtual leader.

We assume the initial position states of each OMR are randomly chosen within $[-5, 5] \times [-5, 5]$, the initial angle is 0, and the initial velocity is $[0, 0, 0]^T$. $\delta_1 = (-2, 0, 0)^T$, $\delta_2 = [1, \sqrt{3}, 0]^T$, and $\delta_3 = [1, -\sqrt{3}, 0]^T$. The comparison simulations are carried out by the proposed adaptive method of LESO-SMC parameters with and without Q-learning algorithm, respectively, denoted by 'Q-LESO-SMC′ and 'LESO-SMC′, respectively. The parameters of LESO-SMC are chosen as $\omega_0 = 2$, $c_1 = 6$, $c_2 = 7$, $\mu = 4$, and the parameters of Q-LESO-SMC are obtained after online optimization by the Q-learning algorithm in Section 3.4. The other parameters used in the simulation are chosen as $\alpha = 0.1, \gamma = 0.8, \sigma = 0.0001, \varepsilon = 0.2$.

In order to observe the performance of the two different controllers, the robot formation is implemented in two cases, including the case with constant disturbances and the case with sinusoidal disturbances.

Case A. The case with constant disturbances.

Consider the robot formation subject to constant external disturbances, denoted by step functions. The corresponding results are shown in Figures 3–9.

The estimation behaviours of the LESO in two dimensions for Case A with constant disturbances are shown in Figures 3–5. Here $\tilde{x}_{i,1x}, \tilde{x}_{i,2x}$, and $\tilde{x}_{i,3x}$ $(\tilde{x}_{i,1y}, \tilde{x}_{i,2y}, \tilde{x}_{i,3y})$ denote the position, velocity, and the extended state (total disturbances) estimation in x-direction (in y direction) for the ith agent, $i = 1,2,3$, respectively. In each figure, the error plots are locally enlarged to highlight the specific convergence time and steady errors. From the comparison results, the estimation errors of all three states converge to a small stable region, and the estimation error of Q-LESO has faster convergence time and smaller steady estimation error than that of LESO under the constant disturbances.

Figure 6 shows the trajectories of three mobile robots and the virtual leader. ☆ and ○ represent the end positions of each follower robot and the virtual leader, respectively. It can be seen from the figure that the robots form a preset formation with the virtual leader as the center.

In order to make the comparison of control performance between LESO-SMC and Q-LESO-SMC, the formation tracking performance of each follower under these two controllers is depicted in Figures 7–8. Here $e_{i,1x}$ and $e_{i,2x}$ ($e_{i,1y}$ and $e_{i,2y}$) denote the formation tracking position error

Table 1: Model parameters of the OMR.

Parameters	Definitions	Values
m	Mass of the OMR	4.5 kg
I_v	Moment of inertia	0.25 kg$*$m^2
d	Distance between the center of mass of the OMR and the intersection point of the symmetry axis of the OMR and the wheel axis	0.2 m
r	Wheel radius	0.0635 m
b	The distance between an actuated wheel and the symmetry axis	0.166 m

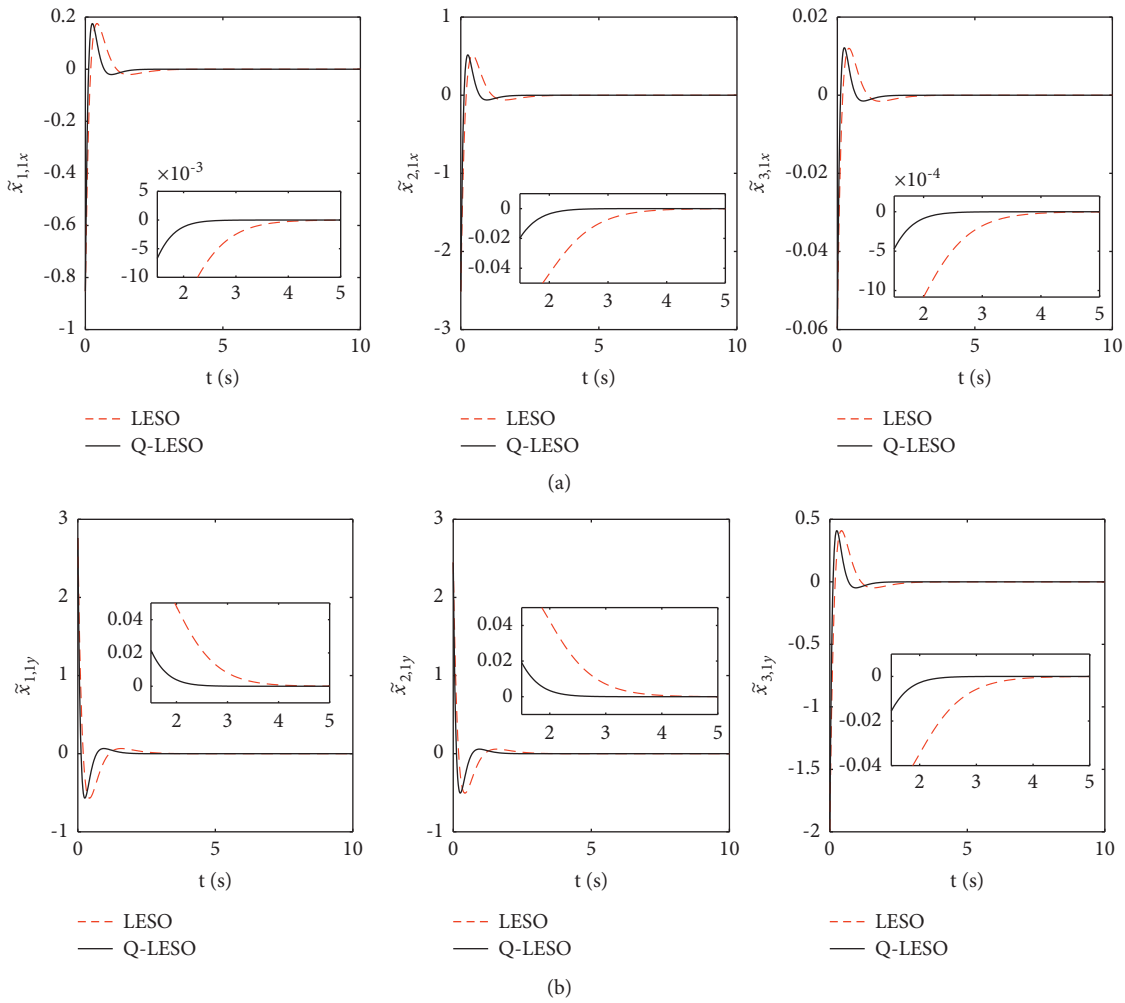


Figure 3: The position estimation error $\tilde{x}_{i,1}$ in Case A. (a) The position estimation error of the x-direction of the ith agent $i = 1, 2, 3$. (b) The position estimation error of the y-direction of the ith agent $i = 1, 2, 3$.

and velocity error in x-direction (in y direction) for the ith agent, $i = 1,2,3$, respectively. It can be seen that the convergence speed of formation tracking error with Q-LESO-SMC is faster than that with LESO-SMC method, and Q-LESO-SMC method has stronger ability to suppress disturbance by comparing the magnitudes of the steady-state errors.

The position trajectories for three OMRs are shown in Figure 9, where ○ represents the starting position of each robot. It can be seen that 3 followers quickly track the virtual leader under the proposed Q-LESO-SMC method.

Case B. The case with sinusoidal disturbances.

For simplicity of simulation, consider the robot formation subject to nonlinear external disturbances, such as sine functions, e.g., $f_i = [\,0.5e^{-0.02t}\cos(t),\ 0.2\sin(t),\ 0.5\cos(t)\,]^T$. The corresponding results are shown in Figures 10–16.

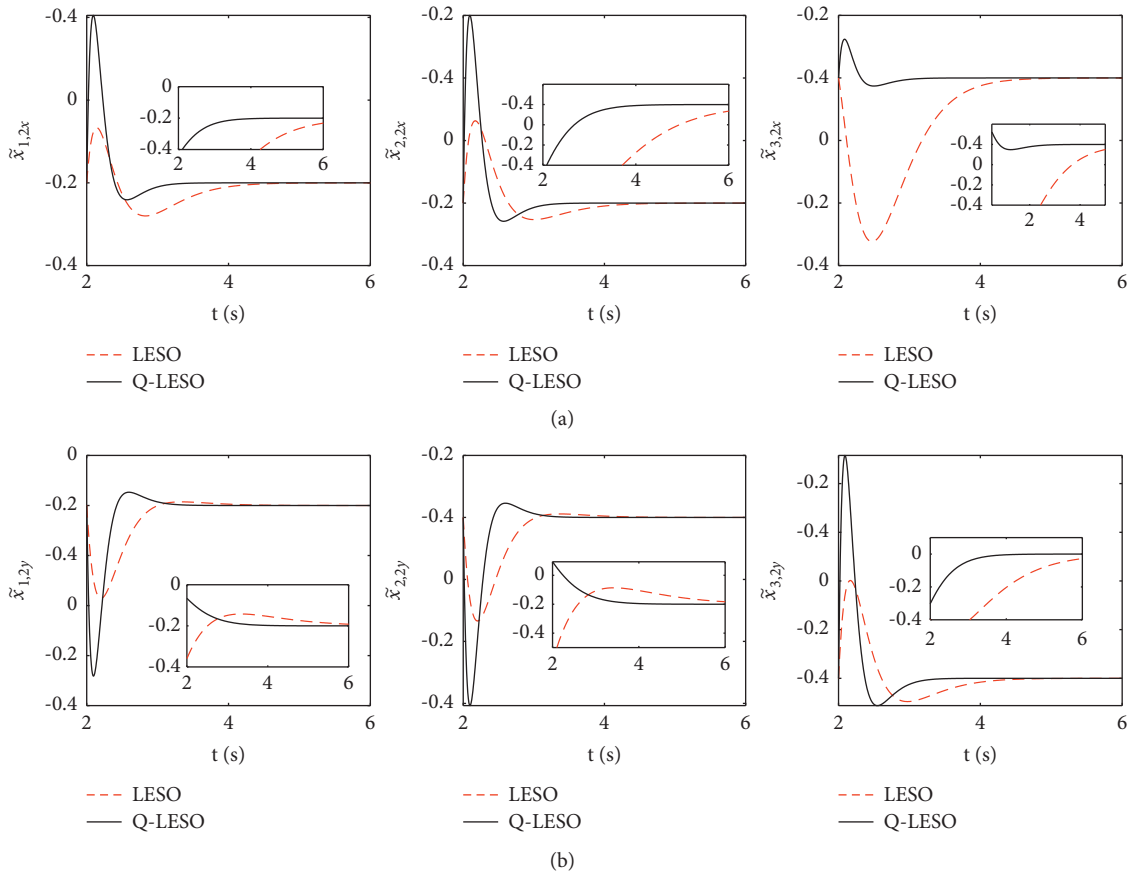


Figure 4: The velocity estimation error $\tilde{x}_{i,2}$ in Case A. (a) The velocity estimation error of the x-direction of the ith agent $i = 1, 2, 3$. (b) The velocity estimation error of the y-direction of the ith agent $i = 1, 2, 3$.

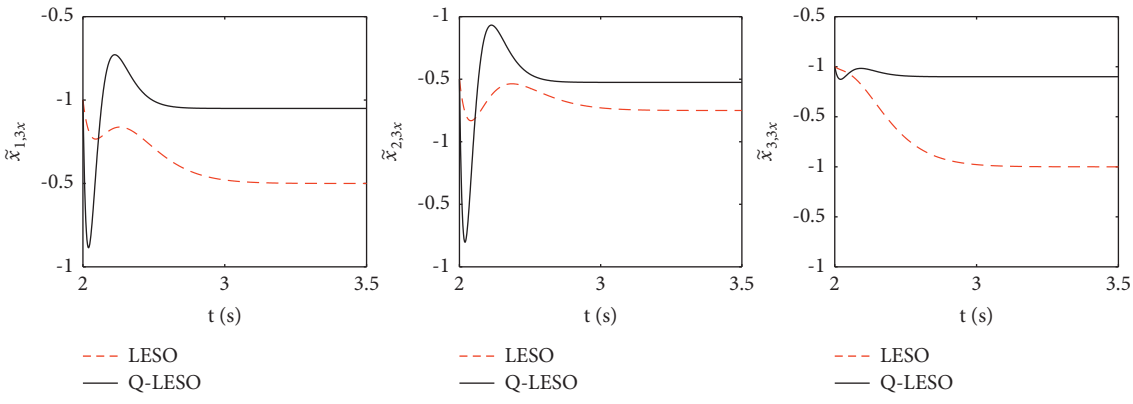
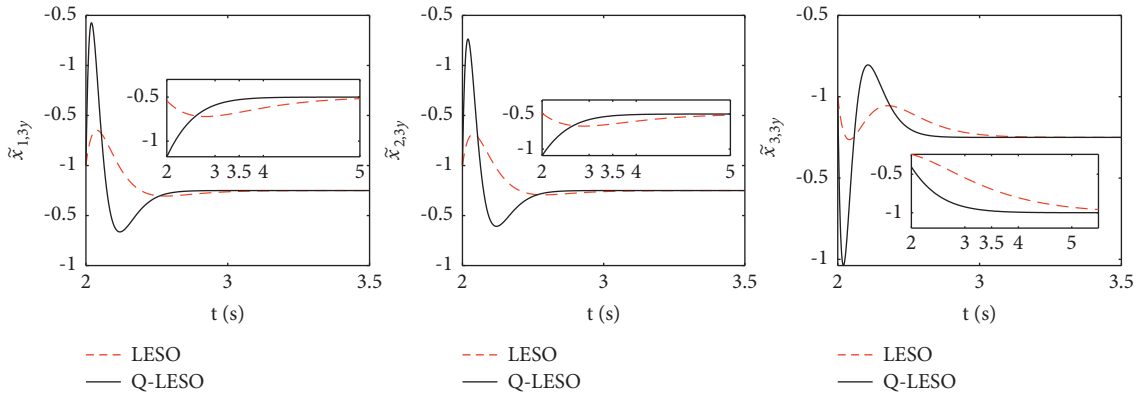


Figure 5: Continued.



(b)

Figure 5: The extended state estimation error $\tilde{x}_{i,3}$ in Case A. (a) The disturbance estimation error of the x-direction of the ith agent $i = 1, 2, 3$. (b) The disturbance estimation error of the y-direction of the ith agent $i = 1, 2, 3$.

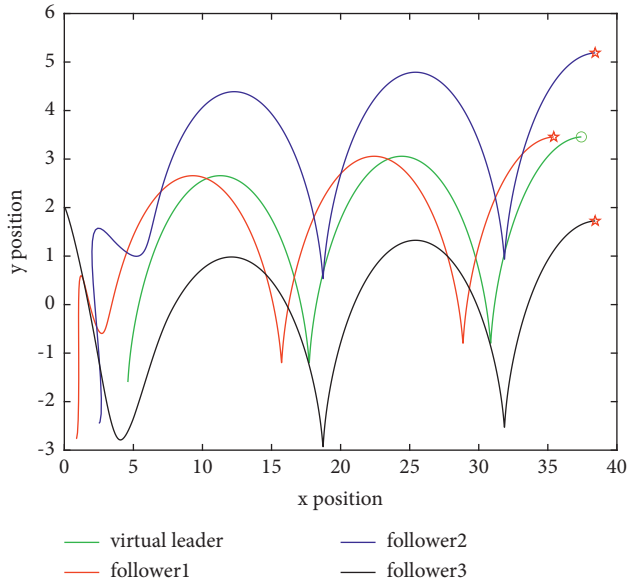
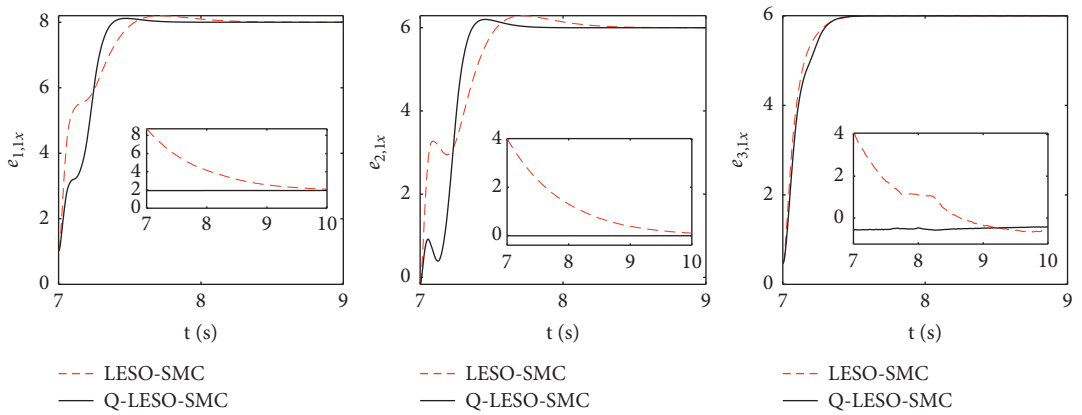


Figure 6: The formation trajectories of three followers and the virtual leader in Case A.



(a)

Figure 7: Continued.

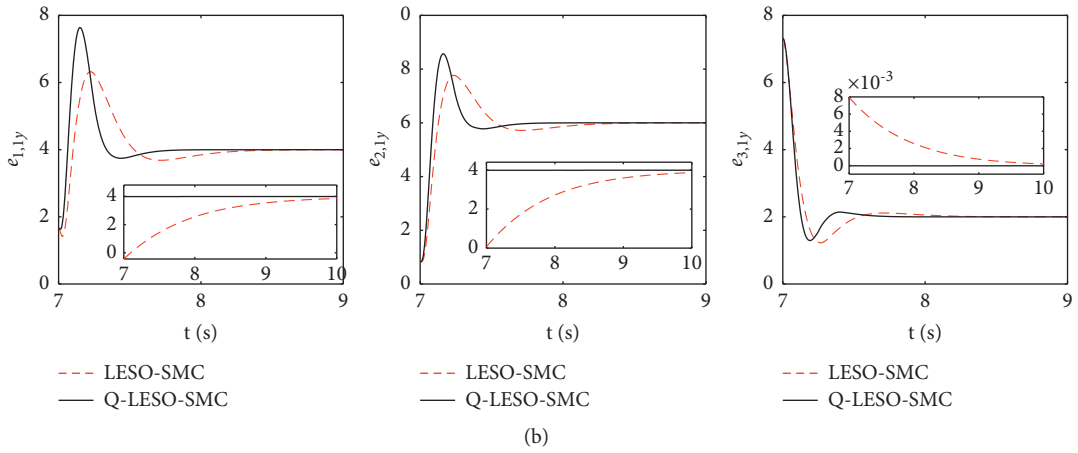


FIGURE 7: The formation tracking position error $e_{i,1}$ in Case A. (a) The formation tracking position error of the x-direction of the ith agent $i = 1, 2, 3$. (b) The formation tracking position error of the y-direction of the ith agent $i = 1, 2, 3$.

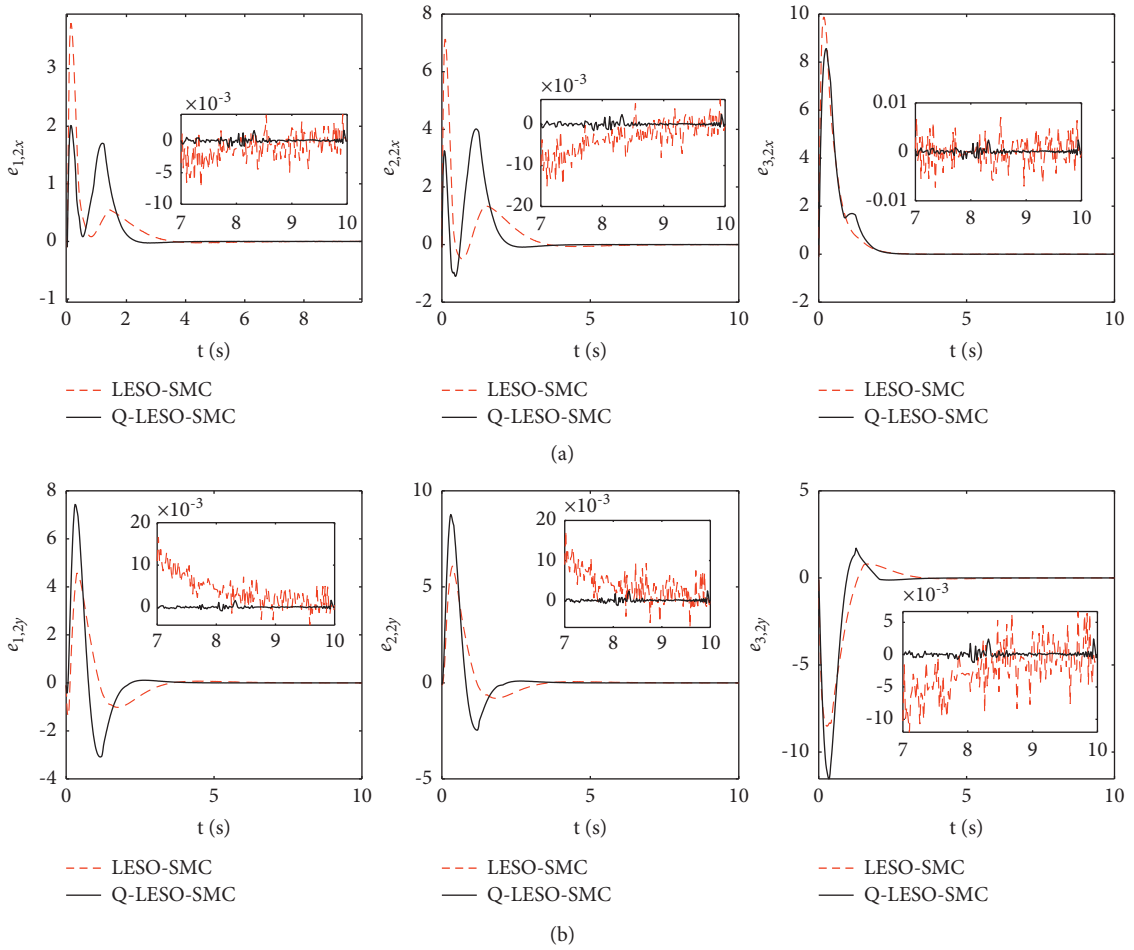


FIGURE 8: The formation tracking velocity error $e_{i,2}$ in Case A. (a) The formation tracking velocity error of the x-direction of the ith agent $i = 1, 2, 3$. (b) The formation tracking velocity error of the y-direction of the ith agent $i = 1, 2, 3$.

Figures 10–12 show the comparison of estimation performance between Q-LESO and LESO with sinusoidal disturbances. The comparison of disturbance estimation behaviour is shown in Figure 12, where the 'total disturbances' of each follower are approximated by Q-LESO in a shorter period. In addition, the amplitude of the

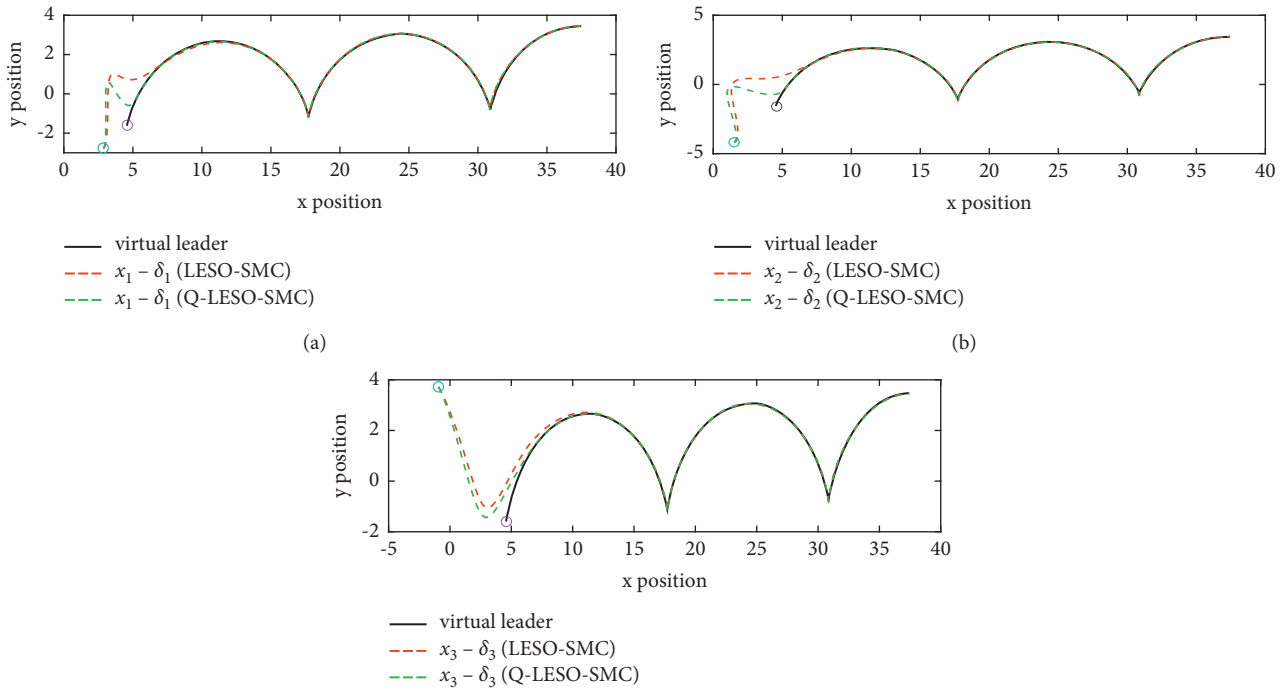


Figure 9: The tracking trajectories of the three followers and the virtual leader in Case A. (a) The tracking trajectories of the 1st follower for the virtual leader. (b) The tracking trajectories of the 2nd follower for the virtual leader. (c) The tracking trajectories of the 3rd follower for the virtual leader.

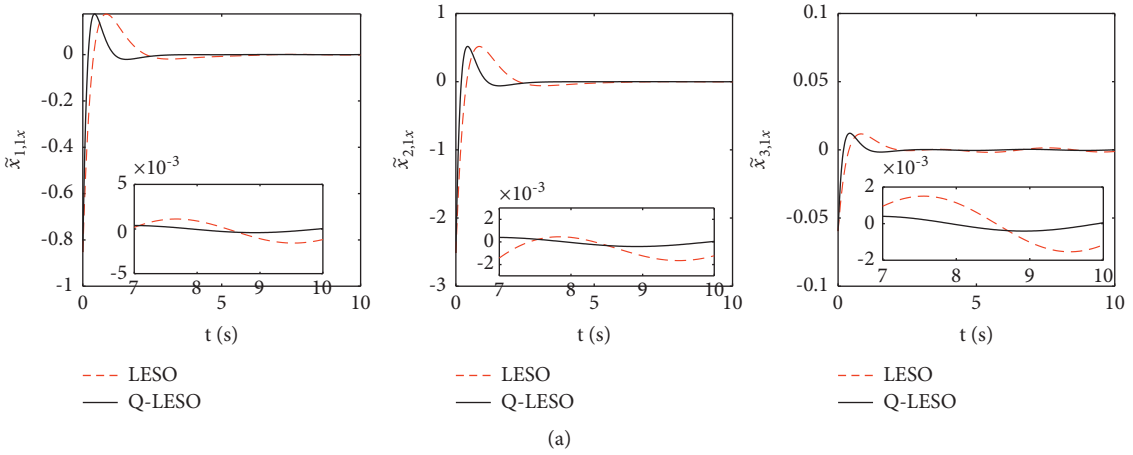


Figure 10: Continued.

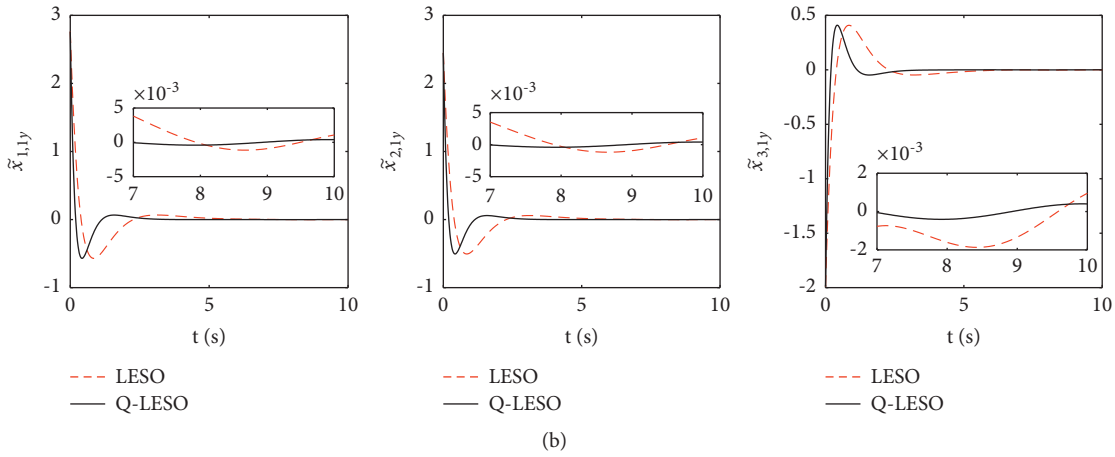


Figure 10: The position estimation error $\tilde{x}_{i,1}$ in Case B(a) The position estimation error of the x-direction of the ith agent $i = 1, 2, 3$. (b) The position estimation error of the y-direction of the ith agent $i = 1, 2, 3$.

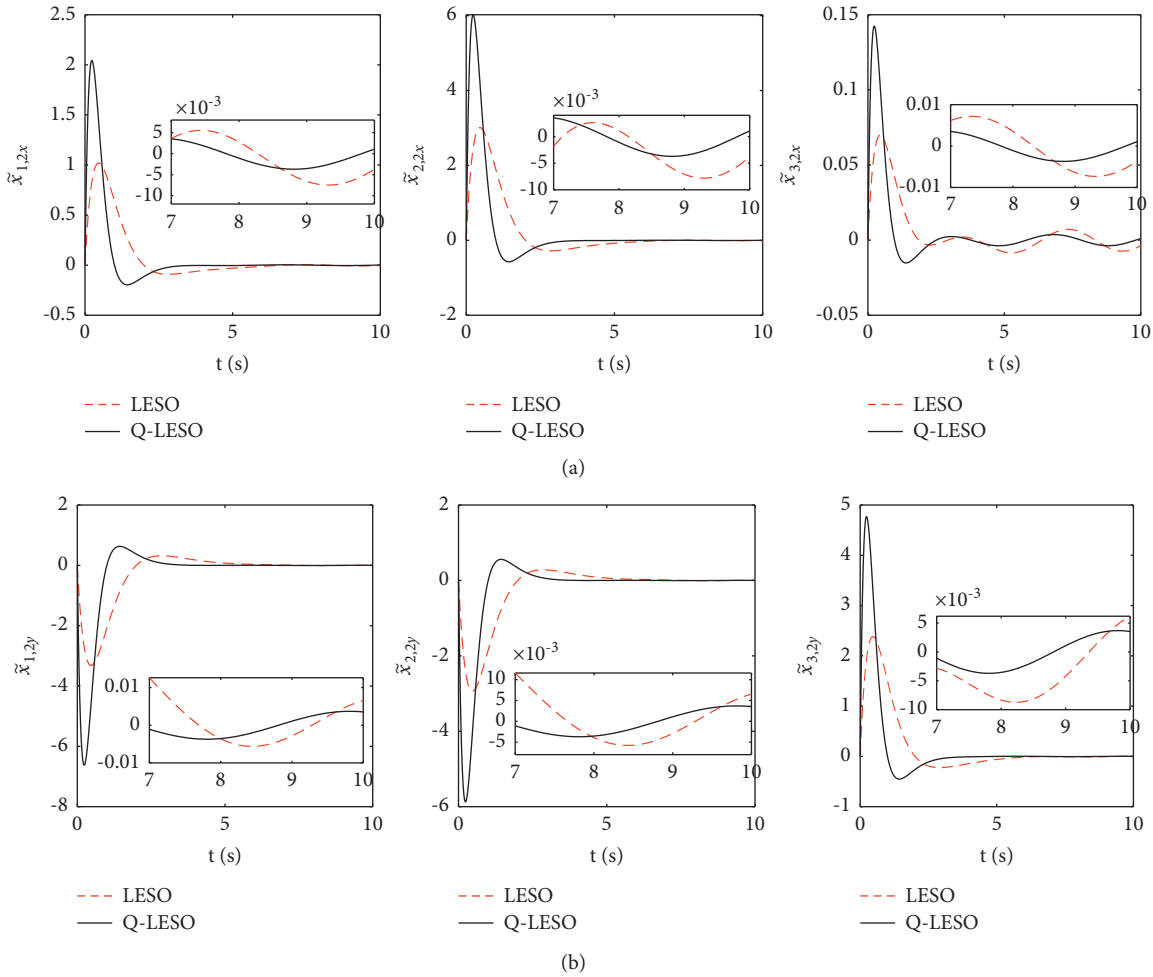


Figure 11: The velocity estimation error $\tilde{x}_{i,2}$ in Case B(a) The velocity estimation error of the x-direction of the ith agent $i = 1, 2, 3$. (b) The velocity estimation error of the y-direction of the ith agent $i = 1, 2, 3$.

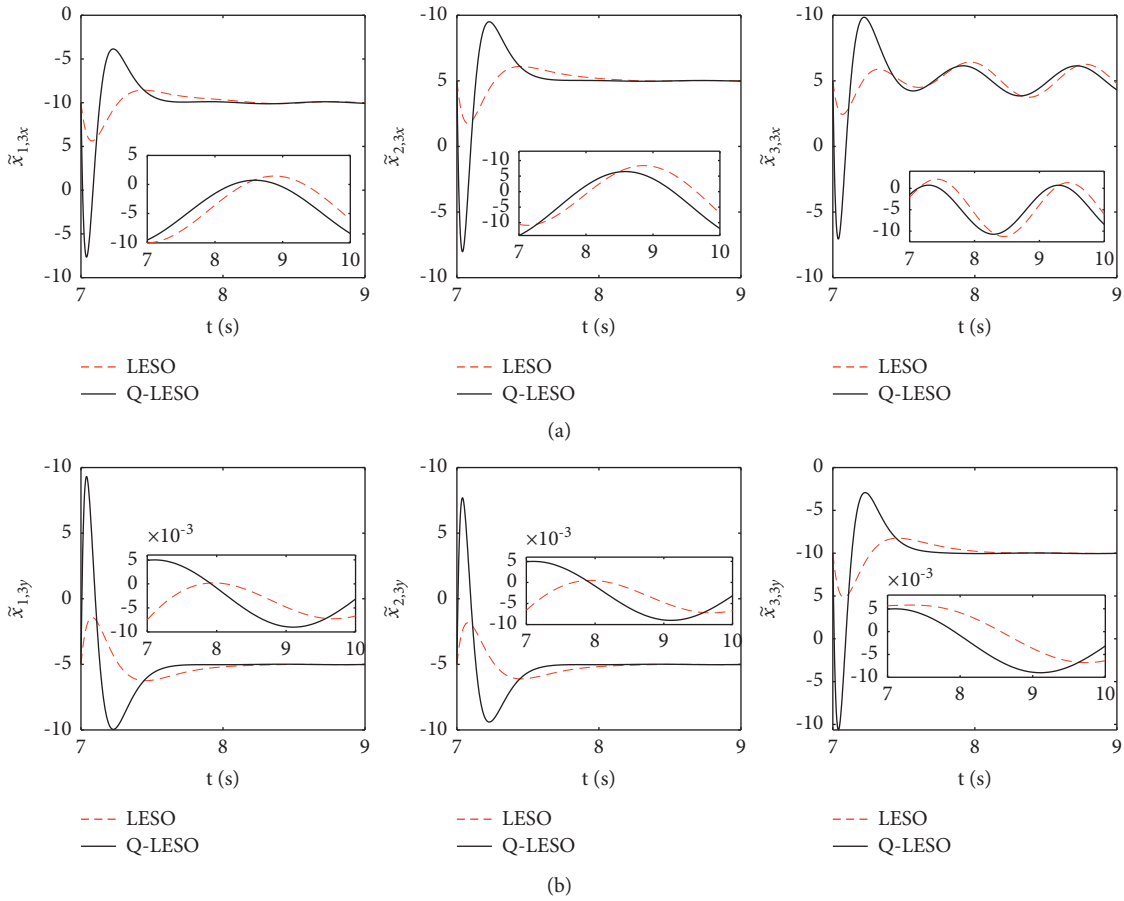


Figure 12: The disturbance estimation error $\tilde{x}_{i,3}$ in Case B(a) The disturbance estimation error of the x-direction of the ith agent $i = 1, 2, 3$. (b) The disturbance estimation error of the y-direction of the ith agent $i = 1, 2, 3$.

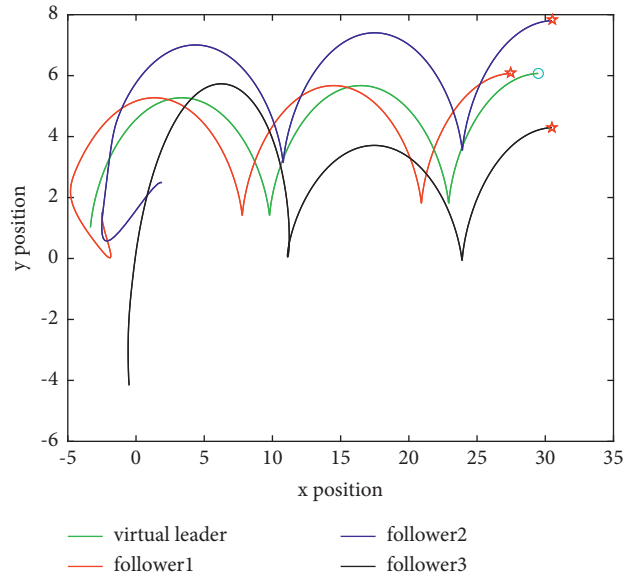


Figure 13: The formation trajectories of three followers and the virtual leader in Case B.

steady state of the estimation error of Q-LESO is smaller than that of LESO as can be seen in the locally enlarged plot.

Similar to Figure 6, Figure 13 shows the trajectories of the three mobile robots and the virtual leader in the same way. As can be seen from the figure, the target triangular

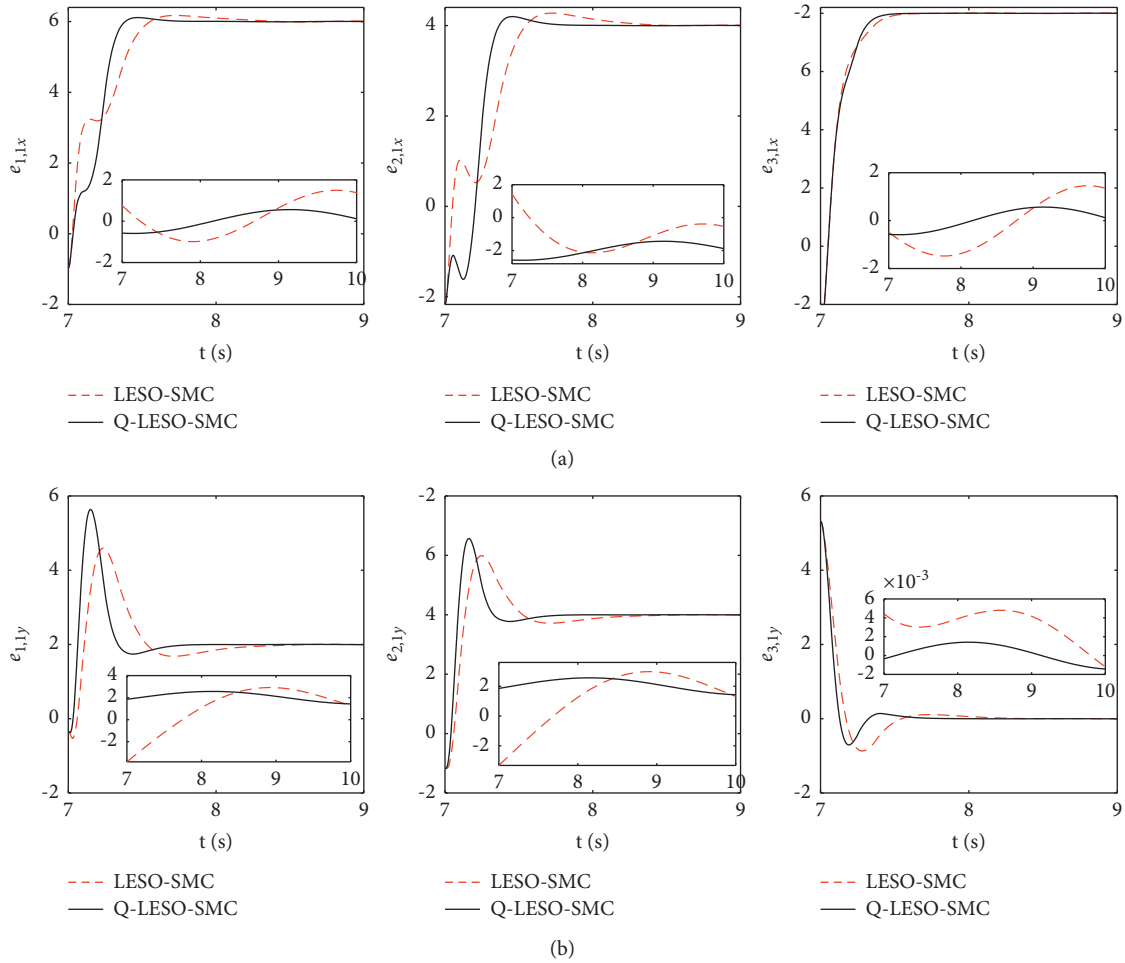


Figure 14: The formation tracking position error $e_{i,1}$ in Case B(a) The formation tracking position error of the x-direction of the ith agent $i = 1, 2, 3$. (b) The formation tracking position error of the y-direction of the ith agent $i = 1, 2, 3$.

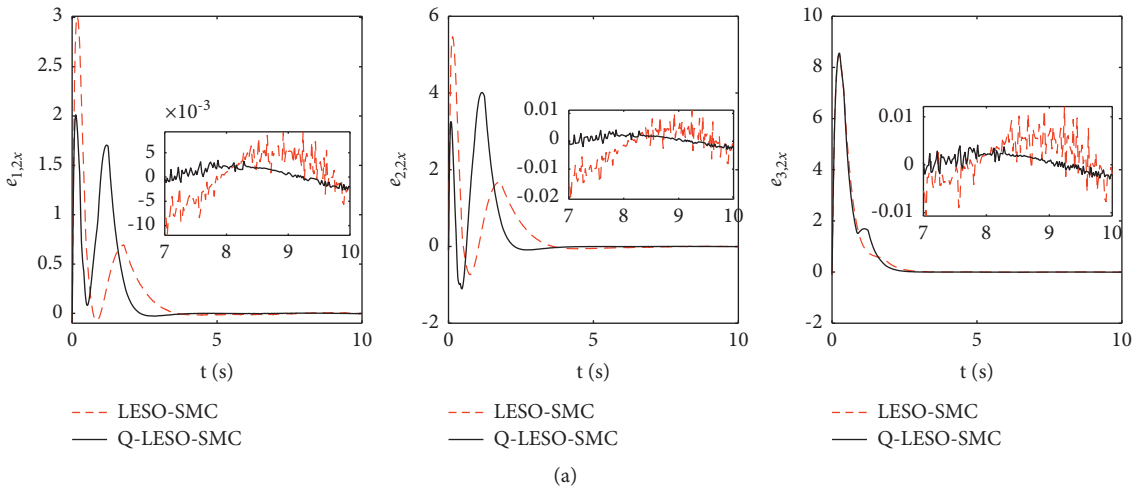
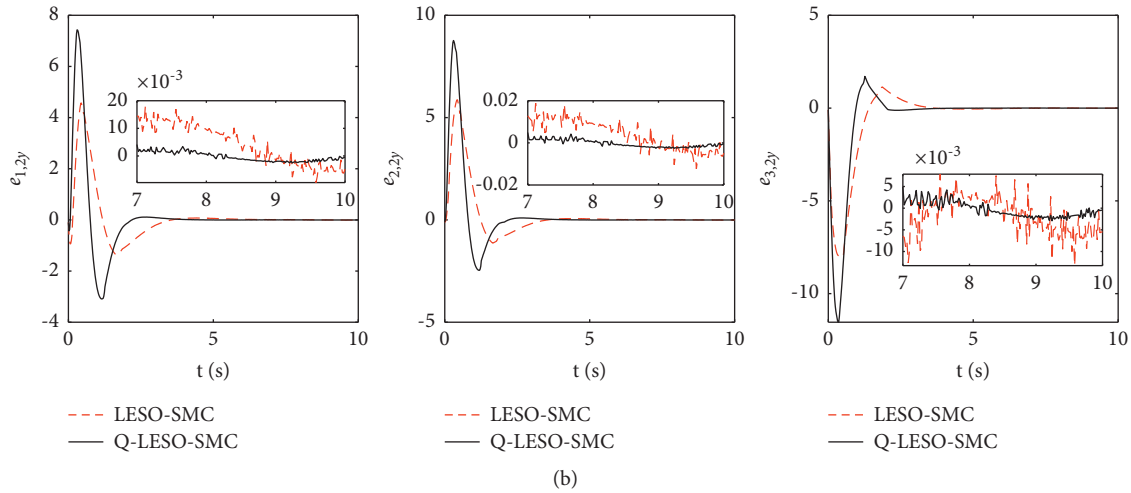


Figure 15: Continued.



(b)

FIGURE 15: The formation tracking velocity error $e_{i,2}$ in Case B(a) The formation tracking velocity error of the x-direction of the ith agent $i = 1, 2, 3$. (b) The formation tracking velocity error of the y-direction of the ith agent $i = 1, 2, 3$.

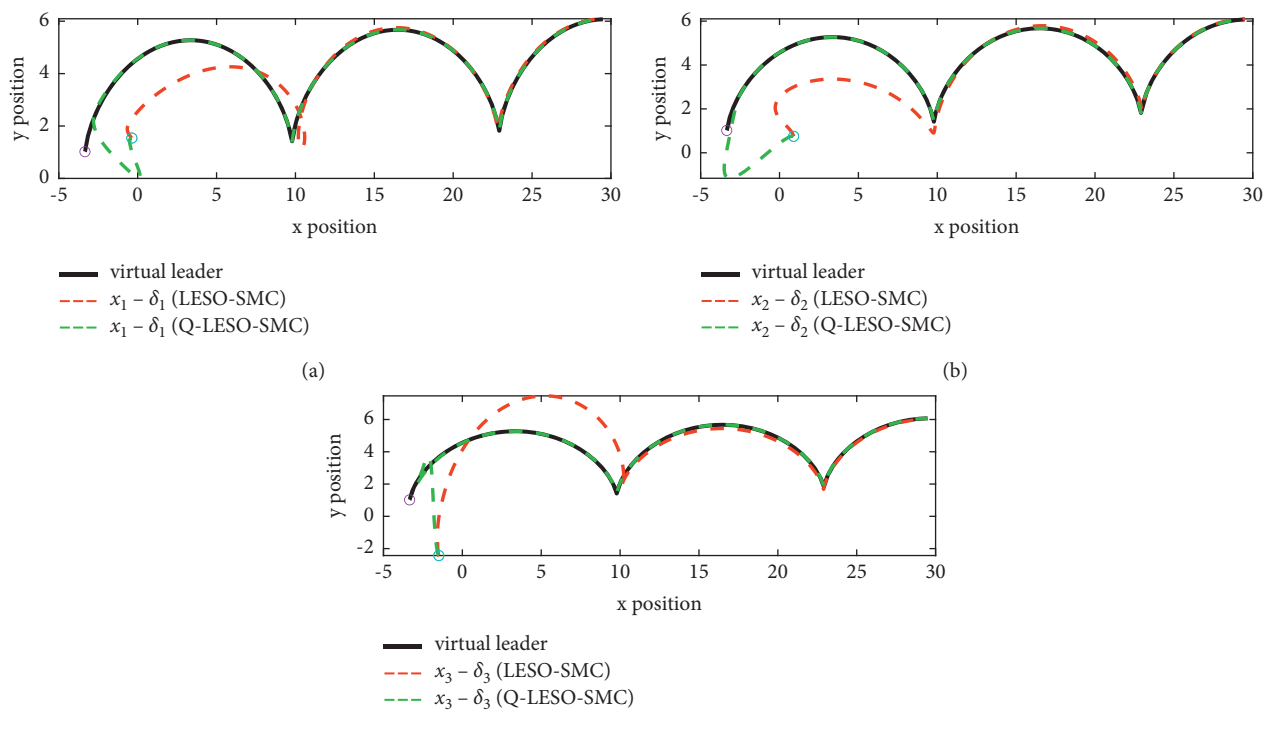


FIGURE 16: The tracking trajectories of the three followers and the virtual leader in Case B(a) The tracking trajectories of the 1st follower for the virtual leader. (b) The tracking trajectories of the 2nd follower for the virtual leader. (c) The tracking trajectories of the 3rd follower for the virtual leader.

formation is achieved with the virtual leader at the center.

Figures 14–15 show the comparison of formation tracking error between Q-LESO-SMC and LESO-SMC with sinusoidal disturbances. As can be seen, the followers track the virtual leader faster with the Q-LESO-SMC-based formation controller. The error steady-state part is locally enlarged, and comparing the magnitude of the steady-state error before and after optimization shows that the error amplitude of Q-LESO-SMC is smaller than that of LESO-SMC.

Similar to Figure 9 in Case A, the position trajectories for three OMRs are shown in Figure 16. It can be seen that 3

followers quickly track the virtual leader under the proposed Q-ESO-SMC method, and the tracking error is smaller than that under the LESO-SMC method.

Remark 5. In Case B, both LESO and Q-LESO methods can achieve well formation control performance. The comparison results show that the Q-LESO exhibited the advantages of faster convergence rate, smaller tracking error, and better disturbance suppression performance. In addition, to get good performance, the LESO needs larger observer gains, which may go beyond bandwidths of practical systems and make the required control energy infeasible. However, higher observer gains may lead to the bigger overshoot (the so-called peaking phenomenon); see Figures 10–12. Therefore, it is a trade-off between cost and performance.

5. Conclusions

In this paper, considering the uncertainties and external disturbances in the formation process, an LESO-SMC with Q-learning adaptive optimization are proposed to achieve formation tracking of multiple OMRs. The simulation results show that the proposed control method has the advantages of faster convergence rate, higher tracking accuracy, and good steady-state performance. However, the LESO brings large overshoot with increasing bandwidth. We will further investigate this problem in the future.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the Jiangsu Natural Science Foundation of PR China (Grant no. BK20171019), the National Natural Science Foundation of China (Grant no. 61703175), and the Natural Science Foundation of Huizhou University (Grant no. hzu201806).

References

[1] W. Ren and R. Beard, *Distributed Consensus in Multi-vehicle Cooperative Control*, Springer, New York, NY, USA, 2008.

[2] B. Tucker and R. Arkin, "Behavior-based formation control for multirobot teams," *IEEE Transactions on Robotics and Automation*, vol. 14, no. 6, pp. 926–939, 1998.

[3] C. Bai, P. Yan, W. Pan, and J. Guo, "Learning-based multi-robot formation control with obstacle avoidance," *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[4] J. H. Acosta, J. P. Jiménez, C. C. Hernádez, and R. M. Clark, "Leader-follower formation control of wheeled mobile robots without attitude measurements," *Applied Sciences*, vol. 11, no. 22, p. 5639, 2021.

[5] B. Park and S. Yoo, "Connectivity-maintaining and collision-avoiding performance function approach for robust leader–follower formation control of multiple uncertain underactuated surface vessels," *Automatica*, vol. 127, Article ID 109501, 2021.

[6] J. Lin, Z. Miao, H. Zhong, W. Peng, Y. Wang, and R. Fierro, "Adaptive image-based leader–follower formation control of mobile robots with visibility constraints," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 7, pp. 6010–6019, 2020.

[7] L. Cao, H. Li, and N. Wang, "Observer-based event-triggered adaptive decentralized fuzzy control for nonlinear large-scale systems," *IEEE Transactions on Fuzzy Systems*, vol. 27, no. 6, pp. 1201–1214, 2018.

[8] E. Almasri and M. K. Uyguroğlu, "Modeling and trajectory planning optimization for the symmetrical multiwheeled omnidirectional mobile robot," *Symmetry*, vol. 13, no. 6, p. 1033, 2021.

[9] N. Lashkari, M. Biglarbegian, and S. X. Yang, "Development of a novel robust control method for formation of heterogeneous multiple mobile robots with autonomous docking capability," *IEEE Transactions on Automation Science and Engineering*, vol. 17, no. 4, pp. 1759–1776, 2020.

[10] X. Yu, J. Ma, N. Ding, and A. Zhang, "Cooperative target enclosing control of multiple mobile robots subject to input disturbances," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 6, pp. 3440–3449, 2021.

[11] R. Siegwart, I. Nourbakhs, and D. Scaramuzza, *Introduction to Autonomous mobile Robots*, MIT press, Cambridge, MA, USA, 2011.

[12] J. Ghommam, M. Saad, F. Mnif, and M. Faisal, "Finite-time circular formation around a moving target with multiple underactuated ODIN vehicles," *Mathematics and Computers in Simulation*, vol. 180, pp. 230–250, 2021.

[13] Y.-W. Choi, J.-W. Choi, S.-G. Im, D. Qian, and S.-G. Lee, "Multi-robot avoidance control based on omni-directional visual SLAM with a fisheye lens camera," *International Journal of Precision Engineering and Manufacturing*, vol. 19, no. 10, pp. 1467–1476, 2018.

[14] G. Zhang, Y. Wang, J. Wang, J. Chen, and D. Qian, "Disturbance observer–based super-twisting sliding mode control for formation tracking of multi-agent mobile robots," *Measurement and Control*, vol. 53, no. 5-6, pp. 908–921, 2020.

[15] B. M. Yousuf, A. S. Khan, and A. Noor, "Multi-agent tracking of non-holonomic mobile robots via non-singular terminal sliding mode control," *Robotica*, vol. 38, no. 11, pp. 1984–2000, 2020.

[16] R. Zhou and T. Wen, "Analysis and tuning of general linear active disturbance rejection controllers," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 7, pp. 5497–5507, 2018.

[17] J. Baek, M. Jin, and S. Han, "A new adaptive sliding-mode control scheme for application to robot manipulators," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 6, pp. 3628–3637, 2016.

[18] G. Bartolini, A. Ferrara, and E. Usai, "Chattering avoidance by second-order sliding mode control," *IEEE Transactions on Automatic Control*, vol. 43, no. 2, pp. 241–246, 1998.

[19] B. He, G. Liu, H. Chen, and X. Hu, "Extended state observer–based sliding mode learning control for mechanical system," *Measurement and Control*, vol. 53, no. 7-8, pp. 1395–1403, 2020.

[20] Y. Zhao, Y. Zhang, and J. Lee, "Lyapunov and sliding mode based leader-follower formation control for multiple mobile robots with an augmented distance-angle strategy," *International Journal of Control, Automation and Systems*, vol. 17, no. 5, pp. 1314–1321, 2019.

[21] Y. Hui and S. Wang, "Trajectory tracking control for non-holonomic wheeled mobile robots with external disturbances and parameter uncertainties," *International Journal of Control, Automation and Systems*, vol. 18, no. 12, pp. 3015–3022, 2020.

[22] I. Jmel, H. Dimassi, S. H. Said, and F. M'Sahil, "Adaptive observer-based sliding mode control for a two-wheeled self-balancing robot under terrain inclination and disturbances," *Mathematical Problems in Engineering*, vol. 2021, 15 pages, 2021.

[23] J. Han, "From PID to active disturbance rejection control," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 3, pp. 900–906, 2009.

[24] Y. Guo, L. Yu, and J. Xu, "Robust finite-time trajectory tracking control of wheeled mobile robots with parametric uncertainties and disturbances," *Journal of Systems Science and Complexity*, vol. 32, no. 5, pp. 1358–1374, 2019.

[25] J. Chen, N. Du, and Y. Han, "Decoupling attitude control of a hypersonic glide vehicle based on a nonlinear extended state observer," *International Journal of Aerospace Engineering*, vol. 2020, 11 pages, 2020.

[26] Y. Guo, L. Yu, and J. M. Xu, "Anti-windup adaptive sliding mode trajectory tracking control of wheeled mobile robot based on extended state observer," *Journal of Systems Science and Mathematical Sciences*, vol. 37, no. 5, pp. 1179–1193, 2017.

[27] Z. Gao, "Scaling and bandwidth-parameterization based controller tuning," in *Proceedings of the 2003 American Control Conference*, pp. 4989–4996, Denver, CO, USA, June 2003.

[28] Y. Huang and W. Xue, "Active disturbance rejection control: methodology and theoretical analysis," *ISA Transactions*, vol. 53, no. 4, pp. 963–976, 2014.

[29] Z. Zheng and L. Sun, "Path following control for marine surface vessel with uncertainties and input saturation," *Neurocomputing*, vol. 177, pp. 158–167, 2016.

[30] K. Chang, Y. Xia, K. Huang, and D. Ma, "Obstacle avoidance and active disturbance rejection control for a quadrotor," *Neurocomputing*, vol. 190, pp. 60–69, 2016.

[31] H. Yang, S. Wang, Z. Zuo, and P. Li, "Trajectory tracking for a wheeled mobile robot with an omnidirectional wheel on uneven ground," *IET Control Theory & Applications*, vol. 14, no. 7, pp. 921–929, 2020.

[32] S. Richard and G. Andrew, *Reinforcement Learning: An Introduction*, MIT press, Cambridge, MA, USA, 2018.

[33] A. E. Hakim, H. Hindersah, and E. Rijanto, "Application of reinforcement learning on self-tuning PID controller for soccer robot multi-agent system," in *Proceedings of the 2013 Joint International Conference on Rural Information and Communication Technology and Electric-Vehicle Technology rICT and ICev-T*, pp. 1–6, Bandung, Indonesia, November 2013.

[34] V. Aghaei, A. Onat, and I. Eksin, "Fuzzy PID controller design using Q-learning algorithm with a manipulated reward function," in *Proceedings of the European Control Conference (ECC)*, pp. 2502–2507, IEEE, Linz, Austria, July, 2015.

[35] Z. Li, W. Liu, L. Li, L. Guo, and W. Zhang, "Modeling and adaptive controlling of cable-drogue docking system for autonomous underwater vehicles," *International Journal of Adaptive Control and Signal Processing*, 2021.

[36] L. Beckenbach, P. Osinenko, and S. Streif, "A Q-learning predictive control scheme with guaranteed stability," *European Journal of Control*, vol. 56, pp. 167–178, 2020.

[37] Z. Chen, B. Qin, M. Sun, and Q. Sun, "Q-learning-based parameters adaptive algorithm for active disturbance rejection control and its application to ship course control," *Neurocomputing*, vol. 408, pp. 51–63, 2020.

[38] J. Hu and Y. S. Lin, "Consensus control for multi-agent systems with double-integrator dynamics and time delays," *IET Control Theory & Applications*, vol. 4, no. 1, pp. 109–118, 2010.

[39] Z. Qing, L. Gao, and Z. Gao, "On stability analysis of active disturbance rejection control for nonlinear time-varying plants with unknown dynamics," in *Proceedings of the 46th IEEE conference on decision and control*, pp. 3501–3506, New Orleans, LA, USA, December, 2007.

[40] Z. Yuan, Y. Tian, Y. Yin, S. Wang, J. Liu, and L. Wu, "Trajectory tracking control of a four mecanum wheeled mobile platform: an extended state observer-based sliding mode approach," *IET Control Theory & Applications*, vol. 14, no. 3, pp. 415–426, 2019.

[41] J. Rath, M. Defoort, H. Karimi, and K. Veluvolu, "Output feedback active suspension control with higher order terminal sliding mode," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 2, pp. 1392–1403, 2016.

Hindawi
Complexity
Volume 2021, Article ID 5079147, 18 pages
https://doi.org/10.1155/2021/5079147

Research Article

Dynamics of the Exponential Population Growth System with Mixed Fractional Brownian Motion

Weijun Ma,[1] Wei Liu,[2] Quanxin Zhu ,[3] and Kaibo Shi [4]

[1]*School of Information Engineering, Ningxia University, Yinchuan, Ningxia 750021, China*
[2]*School of Mathematics and Information Science, North Minzu University, Yinchuan, Ningxia 750021, China*
[3]*School of Mathematics and Statistics, Hunan Normal University, Changsha, Hunan 410081, China*
[4]*School of Electronic Information and Electrical Engineering, Chengdu University, Chengdu 610106, China*

Correspondence should be addressed to Quanxin Zhu; zqx22@hunnu.edu.cn

Received 5 September 2021; Accepted 4 October 2021; Published 30 December 2021

Academic Editor: Hou-Sheng Su

This paper examines the dynamics of the exponential population growth system with mixed fractional Brownian motion. First, we establish some useful lemmas that provide powerful tools for studying the stochastic differential equations with mixed fractional Brownian motion. We offer some explicit expressions and numerical characteristics such as mathematical expectation and variance of the solutions of the exponential population growth system with mixed fractional Brownian motion. Second, we propose two sufficient and necessary conditions for the almost sure exponential stability and the kth moment exponential stability of the solution of the constant coefficient exponential population growth system with mixed fractional Brownian motion. Furthermore, we conduct some large deviation analysis of this mixed fractional population growth system. To the best of the authors' knowledge, this is the first paper to investigate how the Hurst index affects the exponential stability and large deviations in the biological population system. It is interesting that the phenomenon of large deviations always occurs for addressed system when $(1/2) < H < 1$. Moreover, several numerical simulations are reported to show the effectiveness of the proposed approach.

1. Introduction

Many scholars recently have paid considerable attention to stochastic differential equations (SDEs), as they can be applied in many fields such as mathematics, physics, mechanics, biology, economics, complex networks, control engineering, multiagent systems, and financial markets [1–8]. However, these applications are largely dependent on the stability of the systems, namely, the long-time asymptotic behavior of the solutions to SDEs. In particular, in biology, engineering, complex networks, and control systems, it is most important to guarantee that the systems are stable, thus highlighting the need to investigate the stability of SDEs.

As is well known, there is a great amount of literature on stability analysis, for instance, see [2, 3, 9–17] and the references therein. It should be pointed out that the works in [2, 3, 9–14, 18–22] only considered the case of SDEs with Markovian noise such as Brownian motion, telegraph noise (or burst noise), Poisson noise, and Lévy noise. However, the stochastic process may fulfil the long-range dependence in some important fields like economics [17, 23–28], neural networks [5, 29], telecommunication networks [23, 26, 30], biology population [31, 32], and so forth.

The fractional Brownian motion (fBm) with the Hurst index $H \in (0, 1)$ is a Gaussian self-similar process with stationary increments. The concept of fBm was introduced by Kolmogorov [33] and Hurst [34] and then investigated in [23], where an integral is defined as the ordinary Brownian motion in the pointwise sense. fBm is one of the most important driving noises for stochastic systems, mainly as a result of its important properties, for example, long-term dependence and self-similarity features. These important properties make fBm have powerful memory effect and great potential applications [5, 16, 17, 24, 25, 27, 29, 31, 32, 35–38]. Unfortunately, because fBm is not a semimartingale or

Markov process, the theories of these processes cannot be applied to investigate fBm.

Population systems are often subject to environmental noise [2, 15, 18, 21, 39–41]. In particular, environmental Brownian noise can suppress explosions in the generalized Lotka–Volterra model investigated in [39]. Additionally, a predator-prey model with telegraph noise was considered in [40]. Shaikhet investigated the stability of a stochastic glassy-winged sharpshooter population [15], while Khodabin et al. [41] studied the interpolation solution of the population systems with Brownian motion. It should be noted that references [2, 15, 18, 39–41] only considered the case of population systems driven by Markovian noise. However, with regard to the analysis of the dynamics of population systems with non-Markovian processes, there has been little work in the literature [31, 32]. Therefore, it is significant to further reveal the influence of the non-Markovian process on the dynamics of biological population systems.

In this paper, we consider the dynamics of the exponential population growth system with mixed fractional Brownian motion (mfBm), which is a linear combination of independent Brownian motion and fBm. The mfBm is a non-Markovian, long-range dependent, mixed-self-similar, and correlated process [37, 42]. To the best of our knowledge, except for [31, 32], the extant literature heavily focuses on the dynamics of population systems with Brownian motion.

The almost sure exponential stability and the kth moment exponential stability are the most important stability problems. The relationship between them can be described by the theory of large deviations [11]. Several sufficient and necessary conditions for two types of exponential stability of the differential equations with Brownian motion were proposed [12]. However, the aforementioned two types of exponential stability of systems with mfBm have not been studied. Do the Hurst index and mfBm affect the exponential stability of systems? What are the large deviations, which are the rare events in biological population dynamics with mfBm? In this paper, we aim to give a positive answer to the exponential population growth system with mfBm.

In this paper, first, we will establish some useful lemmas that provide powerful tools for studying the stochastic differential equations with mfBm. We offer some explicit expressions and numerical characteristics such as mathematical expectation and variance of the solutions of the exponential population growth system with mfBm. Second, two sufficient and necessary conditions for the almost sure exponential stability and the kth moment exponential stability of the solution of constant coefficient exponential population growth system with mfBm are given. In view of the exponential stability of the system, we investigate the phenomenon of large deviations. In addition, it is indicated that the almost sure exponential stability and instability are clearly different when the Hurst index takes different values.

The rest of the paper is organized as follows. In Section 2, we give some notations and lemmas. In Sections 3 and 4, we present some explicit expressions, mathematical expectations, and variances of the solutions of the exponential population growth system with mfBm. We study the exponential stability in Section 5. In Section 6, we present the phenomenon of large deviations in biological population system. Finally, Section 7 summarizes the conclusions.

2. Preliminaries

Let $\mathbb{R}^n$ be the n-dimensional Euclidean space, $|x|$ be the Euclidean norm of a vector x, $\mathbb{R} = (-\infty, +\infty)$ be the set of real numbers, and $\mathbb{R}_+ = [0, +\infty)$ be the set of positive real numbers. $S(\mathbb{R}_+)$ is the Schwartz space of rapidly decreasing smooth functions on $\mathbb{R}_+$, and $S(\mathbb{R}_+)$ is its dual space of tempered distributions.

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space and $L^2(\mathbb{P})$ be the shorthand notation for $L^2(\Omega, \mathscr{F}, \mathbb{P})$. $B^H(t)$ is a fBm with Hurst index $H \in (0,1)$, while $B(t)$ is a Brownian motion, with both being defined on the complete probability space. Assume that $B(t)$ and $B^H(t)$ are independent.

Let us define an operator M_H for $0 < H < 1$ on $\mathcal{S}(\mathbb{R}_+)$ by

$$\widehat{M_H f}(y) = C_H |y|^{(1/2)-H} \widehat{f}(y), \quad f \in S(\mathbb{R}_+), \tag{1}$$

where $\widehat{f}$ represents the Fourier transform of f and

$$C_H = \left[2\Gamma\left(H - \frac{1}{2}\right)\cos\left(\frac{\pi}{2}\left(H - \frac{1}{2}\right)\right)\right]^{-1} [\Gamma(2H+1)\sin(\pi H)]^{(1/2)}, \tag{2}$$

with Γ denoting the gamma function. $L_H^2(\mathbb{R}_+)$ is the closure of $\mathcal{S}(\mathbb{R}_+)$ with norm

$$\|f\|^2_{L_H^2(\mathbb{R}_+)} = \int_0^{+\infty} (Mf(x))^2 dx, \quad f \in S(\mathbb{R}_+). \tag{3}$$

We take $Mf(x) = f(x)$ for $H = (1/2)$, the identity map.

Definition 1 (see [25, 37]). Let H be a constant belonging to $(0,1)$. A fBm $B^H = \{B^H(t), t \in \mathbb{R}\}$ of Hurst index H is a centered Gaussian process with continuous sample paths and covariance

$$R_H(t,s) = \mathbb{E}\left[B^H(t)B^H(s)\right] = \frac{1}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right);\ t,\ s \in \mathbb{R}. \tag{4}$$

When $H = (1/2)$, the fBm is a standard Brownian motion.

We recall an Itô formula for the stochastic differential equation with fBm

$$dx(t) = \alpha(t)dt + \gamma(t)dB^H(t), \tag{5}$$

on $t \in \mathbb{R}_+$ with the initial value $x(0) = x_0 \in \mathbb{R}$, where $\alpha(t)$ and $\gamma(t)$ are deterministic continuous functions with $\alpha(t) \in L^1(\mathbb{R}_+; \mathbb{R})$, $\gamma(t) \in L_H^2(\mathbb{R}_+; \mathbb{R})$.

We now present a useful lemma of the fBm.

Lemma 1 (see [25, 35, 36]). *Let $Q(t, x(t)) \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}; \mathbb{R})$ and the stochastic process $x(t)$ be described by (5). Assuming that the random variables*

$$Q(t,x(t)),$$
$$\int_0^t \frac{\partial Q}{\partial x}(s,x(s))dx(s), \tag{6}$$
$$\int_0^t s^{2H-1}\gamma^2(s)\frac{\partial^2 Q}{\partial x^2}(s,x(s))ds$$

all belong to $L^2(\mathbb{P})$ *for all* $t \in \mathbb{R}_+$*; then,*

$$Q(t,x(t)) = Q(0,x_0) + \int_0^t \frac{\partial Q}{\partial s}(s,x(s))ds + \int_0^t \frac{\partial Q}{\partial x}(s,x(s))dx(s) + H\int_0^t s^{2H-1}\gamma^2(s)\frac{\partial^2 Q}{\partial x^2}(s,x(s))ds. \tag{7}$$

We next state the mfBm-Itô lemma, which plays a critical role in what follows.

Consider a stochastic differential equation with mfBm

$$\begin{cases} dx(t) = \alpha(t)dt + \beta(t)dB(t) + \gamma(t)dB^H(t), & t \in \mathbb{R}_+, \\ x(0) = x_0, \end{cases} \tag{8}$$

where $\alpha(t) \in L^1(\mathbb{R}_+;\mathbb{R})$, $\beta(t) \in L^2(\mathbb{R}_+;\mathbb{R})$, *and* $\gamma(t) \in L_H^2(\mathbb{R}_+;\mathbb{R})$.

Lemma 2. *Let* $Q(t,x(t)) \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R};\mathbb{R})$ *and the stochastic process* $x(t)$ *be described by (8). Assume that the random variables*

$$Q(t,x(t)),$$
$$\int_0^t \frac{\partial Q}{\partial x}(s,x(s))dx(s),$$
$$\int_0^t \beta^2(s)\frac{\partial^2 Q}{\partial x^2}(s,x(s))ds, \tag{9}$$
$$\int_0^t s^{2H-1}\gamma^2(s)\frac{\partial^2 Q}{\partial x^2}(s,x(s))ds$$

all belong to $L^2(\mathbb{P})$ *for all* $t \in \mathbb{R}_+$*; then,*

$$\begin{aligned} Q(t,x(t)) = Q(0,x_0) &+ \int_0^t \left(\frac{\partial Q}{\partial s}(s,x(s)) + \alpha(s)\frac{\partial Q}{\partial x}(s,x(s))\right)ds \\ &+ \int_0^t \left(\frac{1}{2}\beta^2(s) + Hs^{2H-1}\gamma^2(s)\right)\frac{\partial^2 Q}{\partial x^2}(s,x(s))ds \\ &+ \int_0^t \beta(s)\frac{\partial Q}{\partial x}(s,x(s))dB(s) + \int_0^t \gamma(s)\frac{\partial Q}{\partial x}(s,x(s))dB^H(s). \end{aligned} \tag{10}$$

Proof. Inspired by references [22, 43, 44], we here employ two methods to prove Lemma 2.

Method 1. Using fBm-Itô Lemma 1 to $Q(t,x(t))$, we have

$$\begin{aligned} dQ(t,x(t)) &= \left(\frac{\partial Q}{\partial t} + \frac{1}{2}\beta^2(t)\frac{\partial^2 Q}{\partial x^2} + Ht^{2H-1}\gamma^2(t)\frac{\partial^2 Q}{\partial x^2}\right)dt + \frac{\partial Q}{\partial x}dx(t), \\ &= \left(\frac{\partial Q}{\partial t} + \frac{1}{2}\beta^2(t)\frac{\partial^2 Q}{\partial x^2} + Ht^{2H-1}\gamma^2(t)\frac{\partial^2 Q}{\partial x^2} + \alpha(t)\frac{\partial Q}{\partial x}\right)dt + \beta(t)\frac{\partial Q}{\partial x}dB(t) + \gamma(t)\frac{\partial Q}{\partial x}dB^H(t). \end{aligned} \tag{11}$$

Then, integrating both sides of the above expression (11) from 0 to t gives (10).

Method 2. Using the well-known Taylor expansion formula leads to

$$dQ(t,x(t)) = \frac{\partial Q}{\partial t}dt + \frac{\partial Q}{\partial x}dx(t) + \frac{1}{2}\frac{\partial^2 Q}{\partial x^2}(dx(t))^2 + o(dt). \tag{12}$$

Because $\mathbb{E}[B(t)]^2 = t$ and $\mathbb{E}[B^H(t)]^2 = t^{2H}$, we can derive that

$$\begin{aligned}
&[\mathrm{d}B(t)]^2 = \mathrm{d}t,\\
&\left[\mathrm{d}B^H(t)\right]^2 \approx \mathrm{Var}\left(\mathrm{d}B^H(t)\right)\\
&\qquad = \mathbb{E}\left(\mathrm{d}B^H(t)\right)^2 = \mathrm{d}t^{2H} = 2Ht^{2H-1}\mathrm{d}t.
\end{aligned} \tag{13}$$

Moreover, the Itô formulas and independence of Brownian motion and fBm show that

$$\mathrm{d}t \cdot \mathrm{d}t = \mathrm{d}t \cdot \mathrm{d}B(t) = \mathrm{d}B(t) \cdot \mathrm{d}B^H(t) = \mathrm{d}t \cdot \mathrm{d}B^H(t) = 0. \tag{14}$$

Therefore,

$$\begin{aligned}
(\mathrm{d}x(t))^2 &= \left(\alpha(t)\mathrm{d}t + \beta(t)\mathrm{d}B(t) + \gamma(t)\mathrm{d}B^H(t)\right)^2\\
&= \alpha^2(t)(\mathrm{d}t)^2 + \beta^2(t)(\mathrm{d}B(t))^2 + \gamma^2(t)\left(\mathrm{d}B^H(t)\right)^2 + 2\alpha(t)\beta(t)\mathrm{d}t\mathrm{d}B(t)\\
&\quad + 2\alpha(t)\gamma(t)\mathrm{d}t\mathrm{d}B^H(t) + 2\beta(t)\gamma(t)\mathrm{d}B(t)\mathrm{d}B^H(t)\\
&= \beta^2(t)\mathrm{d}t + 2Ht^{2H-1}\gamma^2(t)\mathrm{d}t + o(\mathrm{d}t),
\end{aligned} \tag{15}$$

where

$$\begin{aligned}
o(\mathrm{d}t) &= \alpha^2(t)(\mathrm{d}t)^2 + 2\alpha(t)\beta(t)\mathrm{d}t\mathrm{d}B(t)\\
&\quad + 2\alpha(t)\gamma(t)\mathrm{d}t\mathrm{d}B^H(t) + 2\beta(t)\gamma(t)\mathrm{d}B(t)\mathrm{d}B^H(t).
\end{aligned} \tag{16}$$

By substituting (15) into (12) and letting $\mathrm{d}t \longrightarrow 0$ yields

$$\begin{aligned}
\mathrm{d}Q(t,x(t)) &= \frac{\partial Q}{\partial t}\mathrm{d}t + \frac{1}{2}\frac{\partial^2 Q}{\partial x^2}\left(\beta^2(t)\mathrm{d}t + 2Ht^{2H-1}\gamma^2(t)\mathrm{d}t\right) + \frac{\partial Q}{\partial x}\left(\alpha(t)\mathrm{d}t + \beta(t)\mathrm{d}B(t) + \gamma(t)\mathrm{d}B^H(t)\right)\\
&= \frac{\partial Q}{\partial t}\mathrm{d}t + \alpha(t)\frac{\partial Q}{\partial x}\mathrm{d}t + \beta(t)\frac{\partial Q}{\partial x}\mathrm{d}B(t) + \gamma(t)\frac{\partial Q}{\partial x}\mathrm{d}B^H(t) + \frac{1}{2}\beta^2(t)\frac{\partial^2 Q}{\partial x^2}\mathrm{d}t + Ht^{2H-1}\gamma^2(t)\frac{\partial^2 Q}{\partial x^2}\mathrm{d}t.
\end{aligned} \tag{17}$$

Hence, integrating both sides of the above expression from 0 to t yields (10). □

Lemma 3. *Assuming that the conditions of Lemma 2 hold, if $x(t)$ is a solution of system (8), then*

$$\begin{aligned}
\mathrm{d}Q(t,x(t)) &= \frac{\partial Q}{\partial t}(t,x(t))\mathrm{d}t + \frac{\partial Q}{\partial x}(t,x(t))\mathrm{d}x(t)\\
&\quad + \frac{1}{2}\frac{\partial^2 Q}{\partial x^2}(t,x(t))(\mathrm{d}x(t))^2,
\end{aligned} \tag{18}$$

where $(dx(t))^2 = (dx(t)) \cdot (dx(t))$ is computed by the rules

$$\begin{aligned}
&\mathrm{d}t \cdot \mathrm{d}t = \mathrm{d}t \cdot \mathrm{d}B(t) = \mathrm{d}B(t) \cdot \mathrm{d}B^H(t) = \mathrm{d}t \cdot \mathrm{d}B^H(t) = 0,\\
&\mathrm{d}B(t) \cdot \mathrm{d}B(t) = \mathrm{d}t,\\
&\mathrm{d}B^H(t) \cdot \mathrm{d}B^H(t) = 2Ht^{2H-1}\mathrm{d}t.
\end{aligned} \tag{19}$$

Remark 1. Lemmas 2 and 3 are called the Itô lemma of mfBm (mfBm-Itô lemma). It follows from (18) that the forms of the Itô formulas for Brownian motion and mfBm are highly unified. However, the computation rules of $(\mathrm{d}x(t))^2$ are slightly different. Moreover, when $H = (1/2)$, system (8) is driven by two independent Brownian motions. Therefore, Lemmas 2 and 3 also hold.

Remark 2. Several versions of the Itô lemma for fBm can be found in the literature (see [26–28, 36, 45–49]). Therefore, Lemma 2 or 3 is an extension of the fBm-Itô lemma with generalizability of those known results.

Based on Lemma 2, we will give some useful lemmas for mfBm.

Lemma 4. *Assuming that $u(t)$ and $v(t)$ are Borel measurable bounded scalar functions defined on $\mathbb{R}_+$, then*

$$\begin{aligned}
&\mathbb{E}\left[\exp\left\{\int_{t_0}^{t}\left(u(s)\mathrm{d}B(s) + v(s)\mathrm{d}B^H(s)\right)\right\}\right]\\
&= \exp\left\{\int_{t_0}^{t}\left(\frac{1}{2}u^2(s) + Hs^{2H-1}v^2(s)\right)\mathrm{d}s\right\}, \quad 0 \le t_0 \le t.
\end{aligned} \tag{20}$$

Proof. Let $z(t) = \exp\left\{\int_{t_0}^{t}(u(s)\mathrm{d}B(s) + v(s)\mathrm{d}B^H(s))\right\}$, and apply Lemma 2.

$$\begin{aligned} \mathrm{d}z(t) &= \exp\left\{\int_{t_0}^{t}\left(u(s)\mathrm{d}B(s)+v(s)\mathrm{d}B^H(s)\right)\right\}\left(u(t)\mathrm{d}B(t)+v(t)\mathrm{d}B^H(t)\right) \\ &+\frac{1}{2}\exp\left\{\int_{t_0}^{t}\left(u(s)\mathrm{d}B(s)+v(s)\mathrm{d}B^H(s)\right)\right\}u^2(t)\mathrm{d}t \\ &+H\ \exp\left\{\int_{t_0}^{t}\left(u(s)\mathrm{d}B(s)+v(s)\mathrm{d}B^H(s)\right)\right\}t^{2H-1}v^2(t)\mathrm{d}t. \end{aligned} \tag{21}$$

It is easy to obtain

$$\begin{aligned} z(t) &= z(t_0)+\int_{t_0}^{t}\exp\left\{\int_{t_0}^{s}\left(u(r)\mathrm{d}B(r)+v(r)\mathrm{d}B^H(r)\right)\right\}\left(u(s)\mathrm{d}B(s)+v(s)\mathrm{d}B^H(s)\right) \\ &+\frac{1}{2}\int_{t_0}^{t}\exp\left\{\int_{t_0}^{s}\left(u(r)\mathrm{d}B(r)+v(r)\mathrm{d}B^H(r)\right)\right\}u^2(s)\mathrm{d}s \\ &+H\int_{t_0}^{t}\exp\left\{\int_{t_0}^{s}\left(u(r)\mathrm{d}B(r)+v(r)\mathrm{d}B^H(r)\right)\right\}s^{2H-1}v^2(s)\mathrm{d}s. \end{aligned} \tag{22}$$

Taking mathematical expectation on both sides of (22) gives

$$\mathbb{E}[z(t)] = \mathbb{E}[z(t_0)]+\frac{1}{2}\int_{t_0}^{t}\mathbb{E}[z(s)]u^2(s)\mathrm{d}s+H\int_{t_0}^{t}\mathbb{E}[z(s)]s^{2H-1}v^2(s)\mathrm{d}s, \tag{23}$$

that is,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[z(t)] = \left(\frac{1}{2}u^2(t)+Ht^{2H-1}v^2(t)\right)\mathbb{E}[z(t)],\ \mathbb{E}[z(t_0)] = 1. \tag{24}$$

Consequently,

$$\mathbb{E}[z(t)] = \exp\left\{\int_{t_0}^{t}\left(\frac{1}{2}u^2(s)+Hs^{2H-1}v^2(s)\right)\mathrm{d}s\right\}. \tag{25}$$

□

Lemma 5. *Assuming that $k>0$, $u(t)$ and $v(t)$ are Borel measurable bounded scalar functions defined on $\mathbb{R}_+$, and*

$$z(t) = \exp\left\{\int_{t_0}^{t}\left(u(s)\mathrm{d}B(s)+v(s)\mathrm{d}B^H(s)\right)\right\}, \tag{26}$$

then

$$\mathbb{E}\left[z^k(t)\right] = \exp\left\{k^2\int_{t_0}^{t}\left(\frac{1}{2}u^2(s)+Hs^{2H-1}v^2(s)\right)\mathrm{d}s\right\},\quad 0\le t_0\le t. \tag{27}$$

Proof. By Lemma 2, we gain

$$\begin{aligned} z^k(t) &= z^k(t_0)+k\int_{t_0}^{t}u(s)z^k(s)\mathrm{d}B(s) \\ &+k\int_{t_0}^{t}v(s)z^k(s)\mathrm{d}B^H(s) \\ &+k^2\int_{t_0}^{t}\left(\frac{1}{2}u^2(s)+Hs^{2H-1}v^2(s)\right)z^k(s)\mathrm{d}s. \end{aligned} \tag{28}$$

Taking mathematical expectation on both sides of (28) gives

$$\begin{aligned} \mathbb{E}\left[z^k(t)\right] &= \mathbb{E}\left[z^k(t_0)\right] \\ &+k^2\int_{t_0}^{t}\left(\frac{1}{2}u^2(s)+Hs^{2H-1}v^2(s)\right)\mathbb{E}\left[z^k(s)\right]\mathrm{d}s, \end{aligned} \tag{29}$$

that is,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[z^k(t)\right] = k^2\left(\frac{1}{2}u^2(t)+Ht^{2H-1}v^2(t)\right)\mathbb{E}\left[z^k(t)\right],\mathbb{E}\left[z^k(t_0)\right] = 1. \tag{30}$$

Consequently,

$$\mathbb{E}\left[z^k(t)\right] = \exp\left\{k^2 \int_{t_0}^{t} \left(\frac{1}{2}u^2(s) + Hs^{2H-1}v^2(s)\right)\mathrm{d}s\right\}. \tag{31}$$

□

Lemma 6. *Assuming that $k > 0$, $u(t)$ and $v(t)$ are Borel measurable bounded scalar functions defined on $\mathbb{R}_+$, and*

$$z(t) = \exp\left(u(t)B(t) + v(t)B^H(t)\right), \tag{32}$$

then

$$\mathbb{E}\left[z^k(t)\right] = \exp\left[k^2\left(\frac{1}{2}u^2(t) + Ht^{2H-1}v^2(t)\right)\right]. \tag{33}$$

Proof. Using the mfBm-Itô Lemma 2 implies

$$\begin{aligned} z^k(t) = z^k(0) &+ k\int_0^t u(s)z^k(s)\mathrm{d}B(s) \\ &+ k\int_0^t v(s)z^k(s)\mathrm{d}B^H(s) \\ &+ k^2\int_0^t \left(\frac{1}{2}u^2(s) + Hs^{2H-1}v^2(s)\right)z^k(s)\mathrm{d}s. \end{aligned} \tag{34}$$

Taking expectation on both sides of (34) yields

$$\mathbb{E}\left[z^k(t)\right] = \mathbb{E}\left[z^k(0)\right] + k^2\int_0^t \left(\frac{1}{2}u^2(s) + Hs^{2H-1}v^2(s)\right)\mathbb{E}\left[z^k(s)\right]\mathrm{d}s, \tag{35}$$

that is,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[z^k(t)\right] = k^2\left(\frac{1}{2}u^2(t) + Ht^{2H-1}v^2(t)\right)\mathbb{E}\left[z^k(t)\right],\ \mathbb{E}\left[z^k(0)\right] = 1. \tag{36}$$

Consequently,

$$\mathbb{E}\left[z^k(t)\right] = \exp\left[k^2\left(\frac{1}{2}u^2(t) + Ht^{2H-1}v^2(t)\right)\right]. \tag{37}$$

Moreover, when $u(t) = u$ and $v(t) = v$ (u and v are constants, the same as below), one finds that

$$\mathbb{E}\left[z^k(t)\right] = \exp\left[k^2\left(\frac{1}{2}u^2t + Ht^{2H}v^2\right)\right]. \tag{38}$$

□

Remark 3. If $u(t) = 0$ or $v(t) = 0$, then Lemmas 4–6 are cases of the corresponding fBm or classical Brownian motion theory, respectively. Therefore, Lemmas 4–6 are generalizations of the classical Brownian motion and fBm theories.

Lemma 7 (law of the iterated logarithm for fBm [17, 25, 50]). *There exists a suitable constant $C_H > 0$ such that*

$$\limsup_{t\longrightarrow+\infty} \frac{B^H(t)}{t^H\sqrt{\log\log t}} = C_H, \tag{39}$$

and then

$$\limsup_{t\longrightarrow+\infty} \frac{B^H(t)}{t^H\sqrt{\log\log t}} = C_H + \varepsilon, \tag{40}$$

which indicates that

$$\limsup_{t\longrightarrow+\infty} \frac{B^H(t)}{t} = 0, \tag{41}$$

where $\epsilon > 0$ is a constant.

Remark 4. The conclusion of Lemma 7 also holds when $H = (1/2)$ and $C_{(1/2)} = \sqrt{2}$ (see [2], Theorem 4.2 on page 16), namely, the case of standard Brownian motion.

Lemma 8. *Assuming that $u(t)$ and $v(t)$ are Borel measurable bounded scalar functions defined on $\mathbb{R}_+$ and*

$$x(t) = \int_0^t u(s)\mathrm{d}B(s) + \int_0^t v(s)\mathrm{d}B^H(s), \quad \frac{1}{2} < H < 1, \tag{42}$$

then

$$x(t) \sim \mathbb{N}\left(0, \int_0^t u^2(s)\mathrm{d}s + H\int_0^t v^2(s)(t-s)^{2H-1}\mathrm{d}s\right). \tag{43}$$

Proof. We recall the basic property of fBm (see, e.g., [5], Remark 2.2, and [25], Definition 1.1.1 on page 5):

$$\mathbb{E}[x(t)] = \mathbb{E}\left[\int_0^t u(s)\mathrm{d}B(s)\right] + \mathbb{E}\left[\int_0^t v(s)\mathrm{d}B^H(s)\right] = 0. \tag{44}$$

It follows from the independence and isometry of Brownian motion and fBm that

$$\begin{aligned}\mathrm{Var}[x(t)] &= \mathbb{E}[x(t) - \mathbb{E}(x(t))]^2 = \mathbb{E}\left[\int_0^t u(s)\mathrm{d}B(s) + \int_0^t v(s)\mathrm{d}B^H(s)\right]^2 \\ &= \mathbb{E}\left[\int_0^t u(s)\mathrm{d}B(s)\right]^2 + \mathbb{E}\left[\int_0^t v(s)\mathrm{d}B^H(s)\right]^2 \\ &= \int_0^t u^2(s)\mathrm{d}s + H(2H-1)\int_0^t\int_0^t v(\xi)v(\eta)|\xi-\eta|^{2H-2}\mathrm{d}\xi\mathrm{d}\eta \\ &= \int_0^t u^2(s)\mathrm{d}s + H(2H-1)\int_0^t\int_0^\xi v^2(s)(\xi-s)^{2H-2}\mathrm{d}s\mathrm{d}\xi \\ &= \int_0^t u^2(s)\mathrm{d}s + H(2H-1)\int_0^t\int_s^t v^2(s)(\xi-s)^{2H-2}\mathrm{d}\xi\mathrm{d}s \\ &= \int_0^t u^2(s)\mathrm{d}s + H\int_0^t v^2(s)(t-s)^{2H-1}\mathrm{d}s.\end{aligned} \tag{45}$$

□

Remark 5. Notably, Lemma 8 is not suitable for $0 < H < (1/2)$, since the kernel $|\xi - \eta|^{2H-2}$ cannot be integrated over the diagonal.

3. The Exponential Population Growth System with mfBm

In this section, we discuss the exponential population growth system with mfBm.

We consider a simple population growth system

$$\begin{cases} \dfrac{\mathrm{d}P(t)}{\mathrm{d}t} = \mu(t)P(t), \\ \\ P(0) = P_0, \end{cases} \tag{46}$$

where $P(t)$ and $\mu(t)$ are the size and relative growth rate of the population at time t, respectively. Let $\mu(t) = \theta(t)$ be a non-random function. Thus, we get that

$$P(t) = P_0 \exp\left\{\int_0^t \theta(s)\mathrm{d}s\right\}. \tag{47}$$

In a special case, if $\theta(t) = \theta$, we obtain

$$P(t) = P_0 \exp(\theta t). \tag{48}$$

It might happen that $\mu(t)$ is not completely known but subject to environmental noise. In other words,

$$\mu(t) = \theta(t) + \text{“noise”}. \tag{49}$$

Let “noise” $= u(t)W(t) + v(t)W^H(t)$; then,

$$\mu(t) = \theta(t) + u(t)W(t) + v(t)W^H(t), \tag{50}$$

where $W(t) = (\mathrm{d}B(t)/\mathrm{d}t)$ and $W^H(t) = (\mathrm{d}B^H(t)/\mathrm{d}t)$ are one-dimensional Gaussian white noise and fractional Gaussian noise, respectively. $B(t)$ and $B^H(t)$ are one-dimensional Brownian motion and fBm, respectively. $u(t)$ and $v(t)$ denote the intensities of the noise at t. Thus, this exponential population growth system with mfBm may be described as follows:

$$\begin{cases} \mathrm{d}P(t) = \theta(t)P(t)\mathrm{d}t + u(t)P(t)\mathrm{d}B(t) + v(t)P(t)\mathrm{d}B^H(t), \\ P(0) = P_0. \end{cases} \tag{51}$$

Remark 6. The exponential population growth system with mfBm (51) is a non-Markovian process because the stochastic perturbation is a mfBm. Notably system (51) reduces to the fractional exponential population growth system, if $u(t) = 0, v(t) \neq 0$; it becomes a stochastic exponential population growth system, if $u(t) \neq 0, v(t) = 0$; and it reduces to a deterministic exponential population growth system, if $u(t) = 0, v(t) = 0$. Thus, the exponential population growth system with mfBm (51) includes fractional, stochastic, and deterministic systems as special cases.

Theorem 1. *The explicit solution of system (51) is given by*

$$P(t) = P_0 \exp\left(\int_0^t \left(\theta(s) - \frac{1}{2}u^2(s) - Hs^{2H-1}v^2(s)\right)\mathrm{d}s + \int_0^t u(s)\mathrm{d}B(s) + \int_0^t v(s)\mathrm{d}B^H(s)\right). \tag{52}$$

Proof. With the help of (51), we deduce

$$\int_0^t \frac{\mathrm{d}P(s)}{P(s)} = \int_0^t \theta(s)\mathrm{d}s + \int_0^t u(s)\mathrm{d}B(s) + \int_0^t v(s)\mathrm{d}B^H(s). \tag{53}$$

Using the mfBm-Itô Lemma 2 to $\ln x$, we obtain

$$\mathrm{d}(\ln P(t)) = \frac{\mathrm{d}P(t)}{P(t)} - \left(\frac{1}{2}u^2(t) + Ht^{2H-1}v^2(t)\right)\mathrm{d}t, \tag{54}$$

or

$$\int_0^t \frac{\mathrm{d}P(s)}{P(s)} = \ln\frac{P(t)}{P_0} + \int_0^t \left(\frac{1}{2}u^2(s) + Hs^{2H-1}v^2(s)\right)\mathrm{d}s. \tag{55}$$

By (53) and (55), we know that

$$\ln\frac{P(t)}{P_0} = \int_0^t \left(\theta(s) - \frac{1}{2}u^2(s) - Hs^{2H-1}v^2(s)\right)\mathrm{d}s + \int_0^t u(s)\mathrm{d}B(s) + \int_0^t v(s)\mathrm{d}B^H(s). \tag{56}$$

Consequently,

$$P(t) = P_0 \exp\left(\int_0^t \left(\theta(s) - \frac{1}{2}u^2(s) - Hs^{2H-1}v^2(s)\right)\mathrm{d}s + \int_0^t u(s)\mathrm{d}B(s) + \int_0^t v(s)\mathrm{d}B^H(s)\right). \tag{57}$$

Moreover, when $\theta(t) = \theta$, $u(t) = u$ and $v(t) = v$, one finds that

$$P(t) = P_0 \exp\left(\left(\theta - \frac{1}{2}u^2 - \frac{1}{2}t^{2H-1}v^2\right)t + uB(t) + vB^H(t)\right). \tag{58}$$

□

Theorem 2. *Assuming that P_0, $B(t)$ and $B^H(t)$ are independent random variables in (51), then*

$$\mathbb{E}[P(t)] = \mathbb{E}[P_0]\exp\left(\int_0^t \theta(s)\mathrm{d}s\right), \tag{59}$$

and

$$\mathrm{Var}[P(t)] = \exp\left(2\int_0^t \theta(s)\mathrm{d}s\right)\left\{\left(\mathrm{Var}[P_0] + \mathbb{E}^2[P_0]\right)\exp\left(\int_0^t \left(u^2(s) + 2Hs^{2H-1}v^2(s)\right)\mathrm{d}s\right) - \mathbb{E}^2[P_0]\right\}. \tag{60}$$

Proof. From (52), one obtains

$$\mathbb{E}[P(t)] = \mathbb{E}[P_0]\exp\left(\int_0^t \left(\theta(s) - \frac{1}{2}u^2(s) - Hs^{2H-1}v^2(s)\right)\mathrm{d}s\right) \times \mathbb{E}\left[\exp\left(\int_0^t \left(u(s)\mathrm{d}B(s) + v(s)\mathrm{d}B^H(s)\right)\right)\right]. \tag{61}$$

Using Lemma 4, one knows that

$$\mathbb{E}\left[\exp\left(\int_0^t \left(u(s)\mathrm{d}B(s) + v(s)\mathrm{d}B^H(s)\right)\right)\right] = \exp\left[\int_0^t \left(\frac{1}{2}u^2(s) + Hs^{2H-1}v^2(s)\right)\mathrm{d}s\right]. \tag{62}$$

Substituting (62) into (61) results in

$$\mathbb{E}[P(t)] = \mathbb{E}[P_0]\exp\left(\int_0^t \theta(s)\mathrm{d}s\right). \tag{63}$$

Furthermore,

$$\mathbb{E}\left[P^2(t)\right] = \mathbb{E}\left[P_0^2\right]\exp\left(2\int_0^t \left(\theta(s) - \frac{1}{2}u^2(s) - Hs^{2H-1}v^2(s)\right)\mathrm{d}s\right) \times \mathbb{E}\left[\exp\left(2\int_0^t \left(u(s)\mathrm{d}B(s) + v(s)\mathrm{d}B^H(s)\right)\right)\right]. \tag{64}$$

According to Lemma 5, one has

$$\mathbb{E}\left[\exp\left(2\int_0^t \left(u(s)\mathrm{d}B(s) + v(s)\mathrm{d}B^H(s)\right)\right)\right] = \exp\left[4\int_0^t \left(\frac{1}{2}u^2(s) + Hs^{2H-1}v^2(s)\right)\mathrm{d}s\right], \tag{65}$$

which together with (64) leads to

$$\mathbb{E}\left[P^2(t)\right] = \mathbb{E}\left[P_0^2\right]\exp\left(2\int_0^t\left(\theta(s) + \frac{1}{2}u^2(s) + Hs^{2H-1}v^2(s)\right)\mathrm{d}s\right). \tag{66}$$

Consequently,

$$\begin{aligned}\mathrm{Var}[P(t)] &= \exp\left(2\int_0^t\theta(s)\mathrm{d}s\right)\left\{\mathbb{E}\left[P_0^2\right]\exp\left(\int_0^t\left(u^2(s) - 2Hs^{2H-1}v^2(s)\right)\mathrm{d}s\right) - \mathbb{E}^2[P_0]\right\}\\ &= \exp\left(2\int_0^t\theta(s)\mathrm{d}s\right)\left\{\left(\mathrm{Var}[P_0] + \mathbb{E}^2[P_0]\right)\exp\left(\int_0^t\left(u^2(s) - 2Hs^{2H-1}v^2(s)\right)\mathrm{d}s\right) - \mathbb{E}^2[P_0]\right\}.\end{aligned} \tag{67}$$

□

Corollary 1. *Assuming that P_0 is a non-random variable, then*

$$\mathbb{E}[P(t)] = [P_0]\exp\left(\int_0^t\theta(s)\mathrm{d}s\right). \tag{68}$$

In particular, if $\theta(t) = \theta$, we get

$$\mathbb{E}[P(t)] = P_0\exp(\theta t), \tag{69}$$

which is the same as (48).

Corollary 2. *Assuming that P_0 is a non-random variable, $\theta(t) = \theta$, $u(t) = u$, and $v(t) = v$, then*

$$\mathrm{Var}[P(t)] = P_0^2\exp(2\theta t)\left(\exp\left(u^2 + t^{2H-1}v^2\right)t - 1\right). \tag{70}$$

4. The Generalized Exponential Population Growth System with mfBm

The symbols $W(t) = (W_1(t), W_2(t), \ldots, W_n(t))$, $B(t) = (B_1(t), B_2(t), \ldots, B_n(t))$, $W^H(t) = (W_1^H(t), W_2^H(t), \ldots, W_n^H(t))$, and $B^H(t) = (B_1^H(t), B_2^H(t), \ldots, B_n^H(t))$ denote, respectively, the n-dimensional Gaussian white noise, Brownian motion, fractional Gaussian noise, and fBm.

By considering

$$\mu(t) = u(t) + \frac{1}{2}\sum_{i=1}^n u_i^2(t) + \sum_{i=1}^n v_i(t)W_i(t) + \sum_{i=1}^n w_i(t)W_i^H(t), \tag{71}$$

where $v_i(t)$ and $w_i(t)\,(i = 1, \ldots, n)$ are the intensities of the noise at time t and $u_i(t)$ denote the error of the growth rate subject to randomly fluctuating environment source i at time t, then we obtain the generalized exponential population growth system with mfBm:

$$\begin{cases}\mathrm{d}P(t) = \left(\left(u(t) + \frac{1}{2}\sum_{i=1}^n u_i^2(t)\right)\mathrm{d}t + \sum_{i=1}^n v_i(t)\mathrm{d}B_i(t) + \sum_{i=1}^n w_i(t)\mathrm{d}B_i^H(t)\right)P(t),\\ P(0) = P_0.\end{cases} \tag{72}$$

Theorem 3. *The explicit solution of system (72) is given by*

$$P(t) = P_0\exp\left(\int_0^t\left(u(s) + \frac{1}{2}\sum_{i=1}^n\left(u_i^2(s) - v_i^2(s)\right) - Hs^{2H-1}\sum_{i=1}^n w_i^2(s)\right)ds + \sum_{i=1}^n\int_0^t v_i(s)dB_i(s) + \sum_{i=1}^n\int_0^t w_i(s)dB_i^H(s)\right). \tag{73}$$

Proof. It can be obtained from (72) that

$$\int_0^t \frac{\mathrm{d}P(s)}{P(s)} = \int_0^t \left(u(s) + \frac{1}{2}\sum_{i=1}^{n} u_i^2(s) \right) \mathrm{d}s + \sum_{i=1}^{n} \int_0^t v_i(s)\mathrm{d}B_i(s) + \sum_{i=1}^{n} \int_0^t w_i(s)\mathrm{d}B_i^H(s). \tag{74}$$

Applying mfBm-Itô Lemma 3 to $\ln x$, we get that

$$\mathrm{d}(\ln P(t)) = \frac{\mathrm{d}P(t)}{P(t)}(t) - \left(\frac{1}{2}\sum_{i=1}^{n} v_i^2(t) + Ht^{2H-1}\sum_{i=1}^{n} w_i^2(t) \right) \mathrm{d}t, \tag{75}$$

or

$$\int_0^t \frac{\mathrm{d}P(s)}{P(s)} = \ln\frac{P(s)}{P_0} + \int_0^t \left(\frac{1}{2}\sum_{i=1}^{n} v_i^2(s) + Hs^{2H-1}\sum_{i=1}^{n} w_i^2(s) \right) \mathrm{d}s. \tag{76}$$

According to (74) and (76), we show that

$$\ln\frac{P(t)}{P_0} = \int_0^t \left(u(s) + \frac{1}{2}\sum_{i=1}^{n} \left(u_i^2(s) - v_i^2(s) \right) - Hs^{2H-1}\sum_{i=1}^{n} w_i^2(s) \right) \mathrm{d}s + \sum_{i=1}^{n} \int_0^t v_i(s)\mathrm{d}B_i(s) + \sum_{i=1}^{n} \int_0^t w_i(s)\mathrm{d}B_i^H(s). \tag{77}$$

Consequently,

$$P(t) = P_0 \exp\left(\int_0^t \left(u(s) + \frac{1}{2}\sum_{i=1}^{n} \left(u_i^2(s) - v_i^2(s) \right) - Hs^{2H-1}\sum_{i=1}^{n} w_i^2(s) \right) \mathrm{d}s + \sum_{i=1}^{n} \int_0^t v_i(s)\mathrm{d}B_i(s) + \sum_{i=1}^{n} \int_0^t w_i(s)\mathrm{d}B_i^H(s). \right) \tag{78}$$

Moreover, when $u(t) = u$, $u_i(t) = u_i$, $v_i(t) = v_i$, and $w_i(t) = w_i (i = 1, \ldots, n)$, one finds that

$$P(t) = P_0 \exp\left[\left(u + \frac{1}{2}\sum_{i=1}^{n} \left(u_i^2 - v_i^2 \right) - \frac{1}{2}t^{2H-1}\sum_{i=1}^{n} w_i^2 \right) t + \sum_{i=1}^{n} v_i B_i(t) + \sum_{i=1}^{n} w_i B_i^H(t) \right]. \tag{79}$$

□

Theorem 4. *Assuming that P_0, $B_i(t)$, and $B_i^H(t) (i = 1, \ldots, n)$ are independent random variables in system (72), then and*

$$\mathbb{E}[P(t)] = \mathbb{E}[P_0] \exp\left(\int_0^t \left(u(s) + \frac{1}{2}\sum_{i=1}^{n} u_i^2(s) \right) \mathrm{d}s \right), \tag{80}$$

$$\mathrm{Var}[P(t)] = \exp\left(2\int_0^t \left(u(s) + \frac{1}{2}\sum_{i=1}^{n} u_i^2(s) \right) \mathrm{d}s \right) \times \left\{ \left(\mathrm{Var}[P_0] + \mathbb{E}^2[P_0] \right) \exp\left(\sum_{i=1}^{n} \int_0^t \left(v_i^2(s) + 2Hs^{2H-1} w_i^2(s) \right) \mathrm{d}s \right) - \mathbb{E}^2[P_0] \right\}. \tag{81}$$

Proof. From (73),

$$\mathbb{E}[P(t)] = \mathbb{E}[P_0]\exp\left(\int_0^t\left[u(s)+\frac{1}{2}\sum_{i=1}^{n}u_i^2(s)-v_i^2(s)-Hs^{2H-1}\sum_{i=1}^{n}w_i^2(s)\right]ds\right) \times \mathbb{E}\left[\exp\left(\sum_{i=1}^{n}\int_0^t v_i(s)dB_i(s)+\sum_{i=1}^{n}\int_0^t w_i(s)dB_i^H(s)\right)\right]. \tag{82}$$

Employing Lemma 4 and independent of the Brownian motion and fBm, one sees that

$$\mathbb{E}\left[\exp\left(\sum_{i=1}^{n}\int_0^t v_i(s)\mathrm{d}B_i(s)+\sum_{i=1}^{n}\int_0^t w_i(s)\mathrm{d}B_i^H(s)\right)\right] = \exp\left[\sum_{i=1}^{n}\int_0^t\left(\frac{1}{2}v_i^2(s)+Hs^{2H-1}w_i^2(s)\right)\mathrm{d}s\right]. \tag{83}$$

Substituting (83) into (82) leads to

$$\mathbb{E}[P(t)] = \mathbb{E}[P_0]\exp\left(\int_0^t\left(u(s)+\frac{1}{2}\sum_{i=1}^{n}u_i^2(s)\right)\mathrm{d}s\right). \tag{84}$$

Furthermore,

$$\mathbb{E}\left[P^2(t)\right] = \mathbb{E}\left[P_0^2\right]\exp\left[2\int_0^t\left(u(s)+\frac{1}{2}\sum_{i=1}^{n}\left(u_i^2(s)-v_i^2(s)\right)-Hs^{2H-1}\sum_{i=1}^{n}w_i^2(s)\right)\mathrm{d}s\right]\times\mathbb{E}\left[\exp\left(2\left(\sum_{i=1}^{n}\int_0^t v_i(s)\mathrm{d}B_i(s)+\sum_{i=1}^{n}\int_0^t w_i(s)\mathrm{d}B_i^H(s)\right)\right)\right]. \tag{85}$$

Applying Lemma 5, one derives that

$$\mathbb{E}\left[\exp\left(2\left(\sum_{i=1}^{n}\int_0^t v_i(s)\mathrm{d}B_i(s)+\sum_{i=1}^{n}\int_0^t w_i(s)\mathrm{d}B_i^H(s)\right)\right)\right] = \exp\left(4\left(\sum_{i=1}^{n}\int_0^t\left(\frac{1}{2}v_i^2(s)+Hs^{2H-1}w_i^2(s)\right)\mathrm{d}s\right)\right). \tag{86}$$

Hence,

$$\mathbb{E}\left[P^2(t)\right] = \mathbb{E}\left[P_0^2\right]\exp\left[2\int_0^t\left(u(s)+\frac{1}{2}\sum_{i=1}^{n}\left(u_i^2(s)+v_i^2(s)\right)+Hs^{2H-1}\sum_{i=1}^{n}w_i^2(s)\right)\mathrm{d}s\right]. \tag{87}$$

Therefore,

$$\begin{aligned}\mathrm{Var}[P(t)] &= \exp\left(2\int_0^t\left(u(s)+\frac{1}{2}\sum_{i=1}^{n}u_i^2(s)\right)\mathrm{d}s\right)\times\left\{\mathbb{E}\left[P_0^2\right]\exp\left(\sum_{i=1}^{n}\int_0^t\left(v_i^2(s)+2Hs^{2H-1}w_i^2(s)\right)\mathrm{d}s\right)-\mathbb{E}^2[P_0]\right\}\\ &= \exp\left(2\int_0^t\left(u(s)+\frac{1}{2}\sum_{i=1}^{n}u_i^2(s)\right)\mathrm{d}s\right)\times\left\{\left(\mathrm{Var}[P_0]+\mathbb{E}^2[P_0]\right)\exp\left(\sum_{i=1}^{n}\int_0^t\left(v_i^2(s)+2Hs^{2H-1}w_i^2(s)\right)\mathrm{d}s\right)-\mathbb{E}^2[P_0]\right\}.\end{aligned} \tag{88}$$

□

Corollary 3. *Assuming that P_0 is a non-random variable, $u(t) = u$, $u_i(t) = u_i$, $v_i(t) = v_i$, and $w_i(t) = w_i (i = 1, \ldots, n)$, then*

$$\mathbb{E}[P(t)] = P_0 \exp\left(\left(u + \frac{1}{2}\sum_{i=1}^{n} u_i^2\right)t\right), \tag{89}$$

and

$$\mathrm{Var}[P(t)] = P_0^2 \exp\left(2\left(u + \frac{1}{2}\sum_{i=1}^{n} u_i^2\right)t\right)\left[\exp\sum_{i=1}^{n}\left(v_i^2 + t^{2H-1}w_i^2\right)t\right) - 1]. \tag{90}$$

5. Exponential Stability

For the sake of simplicity, in what follows, we discuss a constant coefficient exponential population growth system with mfBm:

$$\begin{cases} \mathrm{d}P(t) = \theta P(t)\mathrm{d}t + uP(t)\mathrm{d}B(t) + vP(t)\mathrm{d}B^H(t), \\ P(0) = P_0, \end{cases} \tag{91}$$

where the parameters are defined as before.

$P(t; t_0, P_0)$ represents the solution of system (91) at time t, with an initial value P_0 at time t_0. We then give the definition of sample (or simply) Lyapunov exponent and kth moment Lyapunov exponent for the solution of system (91).

Definition 2 (see [2, 14]). The two limits,

$$L \coloneqq \limsup_{t\longrightarrow+\infty} \frac{1}{t}\log\left(\left|P(t; t_0, P_0)\right|\right), \tag{92}$$

and

$$L(k) \coloneqq \limsup_{t\longrightarrow+\infty} \frac{1}{t}\log\mathbb{E}\left(\left|P(t; t_0, P_0)\right|\right)^k, \tag{93}$$

are called the sample Lyapunov exponent and the kth moment Lyapunov exponent, respectively.

5.1. The Almost Sure Exponential Stability. The following theorem presents a necessary and sufficient condition for system (91) to be almost surely exponentially stable.

Theorem 5

(i) When $0 < H < (1/2)$, the trivial solution of system (91) is almost surely exponentially stable if $\theta \le 0$ or $\theta < (1/2)u^2$ but unstable if $\theta > (1/2)u^2$.

(ii) When $H = (1/2)$, the trivial solution of system (91) is almost surely exponentially stable if $\theta \le 0$ or $\theta < (1/2)(u^2 + v^2)$ but unstable if $\theta > (1/2)(u^2 + v^2)$.

(iii) When $(1/2) < H < 1$, the trivial solution of system (91) is almost surely exponentially stable for all parameters θ, u and v.

Proof. From (41) and (58), we get

$$\begin{aligned} L &= \limsup_{t\longrightarrow+\infty} \frac{1}{t}\log\left[\left|P(t; 0, P_0)\right|\right] = \limsup_{t\longrightarrow+\infty} \frac{1}{t}\log\left[\left|P_0\right|\exp\left(\left(\theta - \frac{1}{2}u^2 - \frac{1}{2}t^{2H-1}v^2\right)t + uB(t) + vB^H(t)\right)\right] \\ &= \lim_{t\longrightarrow+\infty}\left(\theta - \frac{1}{2}u^2 - \frac{1}{2}t^{2H-1}v^2\right). \end{aligned} \tag{94}$$

Then, we obtain

$$L = \begin{cases} \theta - \dfrac{1}{2}u^2, & 0 < H < \dfrac{1}{2}, \\ \theta - \dfrac{1}{2}\left(u^2 + v^2\right), & H = \dfrac{1}{2}, \\ -\infty, & \dfrac{1}{2} < H < 1. \end{cases} \tag{95}$$

Consequently, the trivial solution is almost surely exponentially stable if and only if the sample Lyapunov exponent is negative. Hence, Theorem 5 must hold. □

Remark 7. Theorem 5 shows that the trivial solution of system (91) is almost surely exponentially stable for the arbitrary Hurst index $0 < H < 1$ when $\theta \le 0$. This means that the mfBm cannot destabilize the exponential population growth system (Figures 1–6).

5.2. kth Moment Exponential Stability. The following theorem gives a necessary and sufficient condition for the kth moment exponential stability of system (91).

Theorem 6

(i) When $0 < H < (1/2)$, the trivial solution of system (91) is kth moment exponentially stable if one of the following conditions holds: (1) $0 < k < 1$ and $\theta \le 0$; (2) $0 < k < 1$ and $\theta < (1/2)(1-k)u^2$; (3) $k = 1$ and $\theta < 0$; or (4) $k > 1$, $\theta < 0$, and $(1/2)(k-1)u^2 < -\theta$. Moreover, it is kth moment exponentially unstable if one of the following conditions is satisfied: (1) $0 < k < 1$ and $\theta > (1/2)(1-k)u^2$; (2) $k = 1$ and $\theta > 0$; or (3) $k > 1$ and $\theta \ge 0$.

(ii) When $H = (1/2)$, the trivial solution of system (91) is kth moment exponentially stable if one of the following conditions holds: (1) $0 < k < 1$ and $\theta \le 0$; (2) $0 < k < 1$ and $\theta < (1/2)(1-k)(u^2+v^2)$; (3) $k = 1$ and $\theta < 0$; or (4) $k > 1$, $\theta < 0$, and $(1/2)(k-1)(u^2+v^2) < -\theta$. Moreover, it is kth moment exponentially unstable if one of the following conditions is satisfied: (1) $0 < k < 1$ and $\theta > (1/2)(1-k)(u^2+v^2)$; (2) $k = 1$ and $\theta > 0$; or (3) $k > 1$ and $\theta \ge 0$.

(iii) When $(1/2) < H < 1$, the trivial solution of system (91) is kth moment exponentially stable if $0 < k < 1$ or $k = 1$ and $\theta < 0$, but it is kth moment exponentially unstable if $k > 1$ or $k = 1$ and $\theta > 0$.

Proof. According to Lemma 6 and (58), one deduces

$$L(k) = \limsup_{t\longrightarrow+\infty} \frac{1}{t}\log\mathbb{E}\left[\left|P(t;0,P_0)\right|^k\right] = \limsup_{t\longrightarrow+\infty} \frac{1}{t}\log\left\{\left|P_0\right|^k\mathbb{E}\left[\exp\left(k\left(\theta - \frac{1}{2}u^2 - \frac{1}{2}t^{2H-1}v^2\right)t + kuB(t) + kvB^H(t)\right)\right]\right\} \tag{96}$$
$$= \lim_{t\longrightarrow+\infty}\left(k\theta + \frac{1}{2}k(k-1)u^2 + \frac{1}{2}k(k-1)v^2t^{2H-1}\right).$$

Therefore, one gets

$$L(k) = \begin{cases} k\theta + \frac{1}{2}k(k-1)u^2, & 0 < H < \frac{1}{2}, \\ k\theta + \frac{1}{2}k(k-1)\left(u^2+v^2\right), & H = \frac{1}{2}, \\ -\infty, & \frac{1}{2} < H < 1 \text{ and } 0 < k < 1, \\ \theta, & \frac{1}{2} < H < 1 \text{ and } k = 1, \\ +\infty, & \frac{1}{2} < H < 1 \text{ and } k > 1. \end{cases} \tag{97}$$

The trivial solution is kth moment exponentially stable if and only if the kth moment Lyapunp>ov exponent is negative. Consequently, the required assertion (iii) follows immediately from (97). When $H = (1/2)$, by considering $L(k) = k\theta + (1/2)k(k-1)(u^2+v^2) < 0$, we can get the following four cases: (1) $0 < k < 1$ and $\theta \le 0$; (2) $0 < k < 1$ and $\theta < (1/2)(1-k)(u^2+v^2)$; (3) $k = 1$ and $\theta < 0$; and (4) $k > 1$, $\theta < 0$ and $(1/2)(k-1)(u^2+v^2) < -\theta$. Moreover, by considering $L(k) = k\theta + (1/2)k(k-1)(u^2+v^2) > 0$, we can obtain the following three cases: (1) $0 < k < 1$ and $\theta > (1/2)(1-k)(u^2+v^2)$; (2) $k = 1$ and $\theta > 0$; and (3) $k > 1$ and $\theta \ge 0$. Hence (ii) must hold. Similar to the proof in (ii), one can easily deduce (i). □

6. The Phenomenon of Large Deviations in Biological Population System

Large deviations [51] involve the probability of rare events in random processes. Such events can have dramatic consequences despite their scarcity, especially if one takes into account the behavior of systems in geology [52], population [53], and climate science [54] on long time scales. These large deviations are considered to be the link between the almost sure exponential stability and the kth moment exponential stability. In general, for sufficiently large $k > 0$, even for stable systems, the kth moment Lyapunov exponent is often positive due to the large deviations. Therefore, it is important to explore the phenomenon of large deviations in biological population systems such that $L < 0$ and $L(k) > 0$ for sufficiently large $k > 0$. When $L < 0$, for $P_0 \neq 0$, $|P(t;t_0,P_0)| \longrightarrow 0$ $(t \longrightarrow +\infty)$ almost surely, and the random variable $P_{\max} = \sup\{|P(t;t_0,P_0)|: t \ge t_0\}$ is almost surely finite. In fact, if $P_{\max} \ge |P(t;t_0,P_0)| > 0$, then $\mathbb{E}[P_{\max}^k] < 0$ for all $k \le 0$, and $\mathbb{E}[P_{\max}^k]$ is big for sufficiently large $k > 0$, resulting in the kth moment exponential instability, namely, the kth moment Lyapunov exponent $L(k) > 0$.

We now show the large deviations of system (91).

Theorem 7

(i) When $0 < H < (1/2)$, one of the following conditions fulfils the conditions $L = \theta - (u^2/2) < 0$ and $L(k) = k\theta + (1/2)k(k-1)u^2 > 0$:

(1) $0 < k < 1$ and $(1/2)(1-k)u^2 < \theta < (1/2)u^2$
(2) $k = 1$ and $0 < \theta < (1/2)u^2$

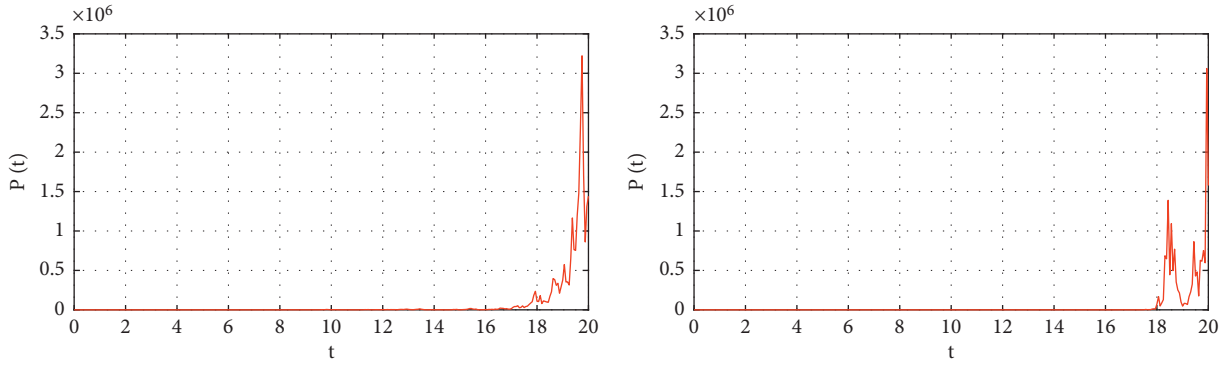


FIGURE 1: The trajectories of system (91) with $P_0 = 1, H = 0.3$. (a) $\theta = 1, u = 1, v = 1$. (b) $\theta = 1, u = 1, v = 2$. The trajectories of system (91) are almost surely exponentially unstable for both (a) and (b) in the sense of Theorem 5 (i).

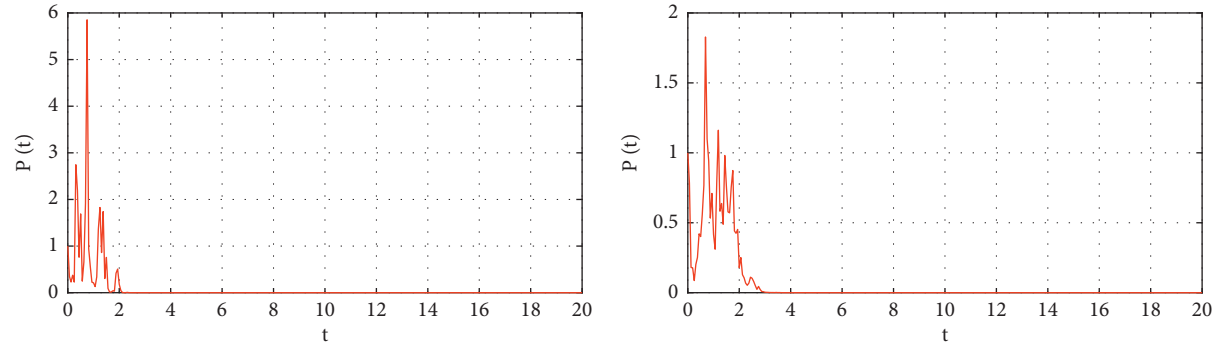


FIGURE 2: The trajectories of system (91) with $P_0 = 1, H = 0.3$. (a) $\theta = 1, u = 3, v = 2$. (b) $\theta = -1, u = 1, v = 1$. The trajectories of system (91) are almost surely exponentially stable for both (a) and (b) in the sense of Theorem 5 (i).

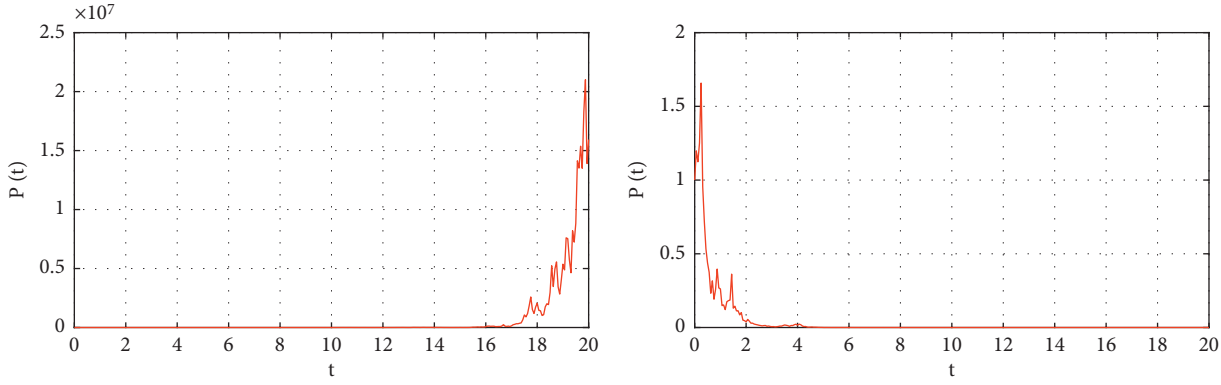


FIGURE 3: The trajectories of system (91) with $P_0 = 1, H = 0.5$. (a) $\theta = 2, u = 1, v = 1$. (b) $\theta = -1, u = 1, v = 1$. The trajectories of system (91) are almost surely exponentially unstable for (a) but stable for (b) in the sense of Theorem 5 (ii).

(3) $k > 1$ *and* $0 \le \theta < (1/2)u^2$

Thus, it holds for some sufficiently large k *when* $0 \le \theta < (1/2)u^2$. *Therefore, the trivial solution of system (91) has large deviations for some special parameters.*

(ii) When $H = (1/2)$, *one of the following conditions fulfils the conditions* $L = \theta - (1/2)(u^2 + v^2) < 0$ *and* $L(k) = k\theta + (1/2)k(k-1)(u^2 + v^2) > 0$:

(1) $0 < k < 1$ *and* $(1/2)(1-k)(u^2 + v^2) < \theta < (1/2)(u^2 + v^2)$
(2) $k = 1$ *and* $0 < \theta < (1/2)(u^2 + v^2)$
(3) $k > 1$ *and.* $0 \le \theta < (1/2)(u^2 + v^2)$

Thus, it holds for some sufficiently large k *when* $0 \le \theta < (1/2)(u^2 + v^2)$. *Therefore, the trivial solution of system (91) has large deviations for some special parameters.*

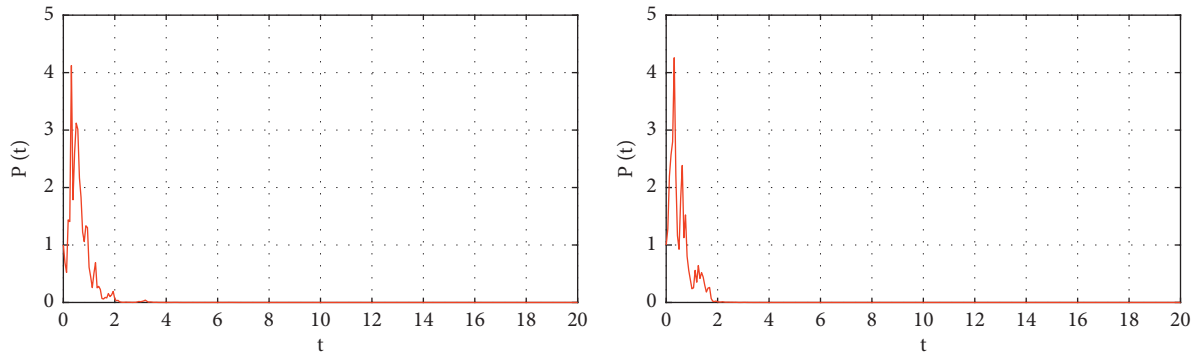


FIGURE 4: The trajectories of system (91) with $P_0 = 1, H = 0.5$. (a) $\theta = 1, u = 2, v = 1$. (b) $\theta = 1, u = 1, v = 2$. The trajectories of system (91) are almost surely exponentially stable for both (a) and (b) in the sense of Theorem 5 (ii).

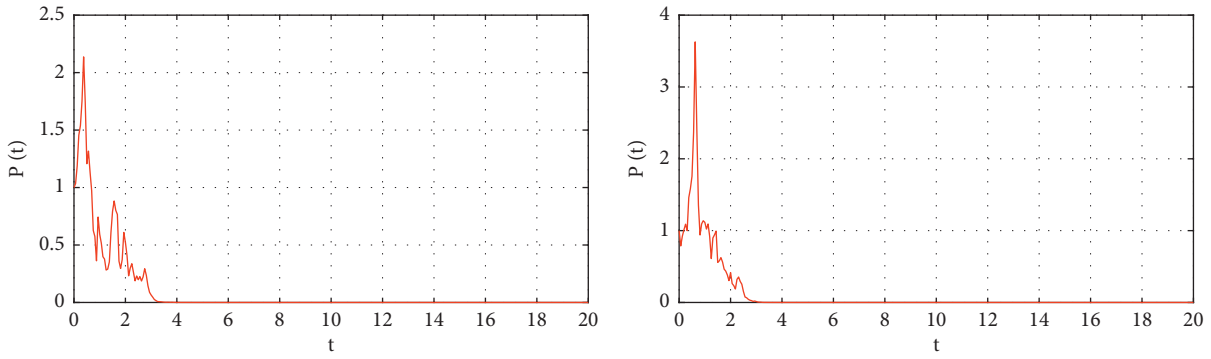


FIGURE 5: The trajectories of system (91) with $P_0 = 1, H = 0.8$. (a) $\theta = 1, u = 1, v = 1$. (b) $\theta = -1, u = 1, v = 1$. The trajectories of system (91) are almost surely exponentially stable for both (a) and (b) in the sense of Theorem 5 (iii).

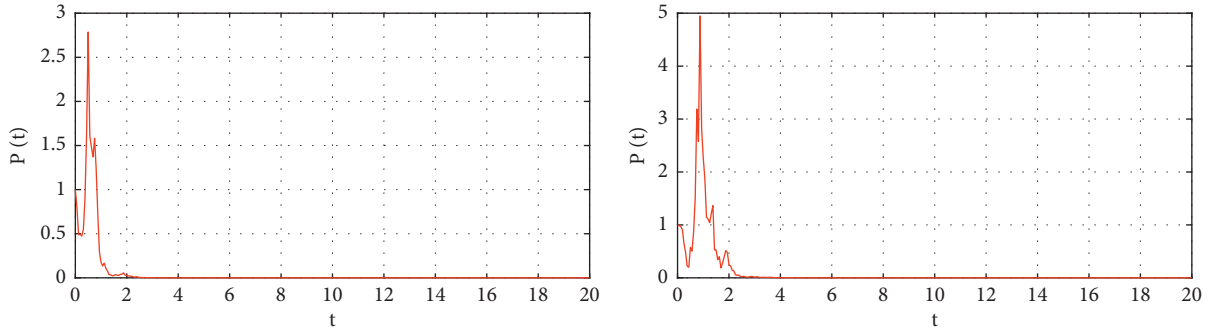


FIGURE 6: The trajectories of system (91) with $P_0 = 1, H = 0.8$. (a) $\theta = 1, u = 1, v = 2$. (b) $\theta = 1, u = 2, v = 1$. The trajectories of system (91) are almost surely exponentially stable for both (a) and (b) in the sense of Theorem 5 (iii).

(iii) When $(1/2) < H < 1$, *we have* $L = -\infty$ *and* $L(k) = -\infty$, θ, $+\infty$ *for* $0 < k < 1$, $k = 1$, $k > 1$, *respectively. Therefore, the conditions* $L < 0$ *and* $L(k) > 0$ *hold for all parameters* θ, u, v, *and* $k > 1$. *Consequently, large deviations occur for all parameters* θ, u, *and* v *provided that* $(1/2) < H < 1$.

Proof. Thus, we can prove this theorem using the expressions of L and $L(k)$. □

Remark 8. For the exponential population growth system with fBm (91) (i.e., $u = 0$), if $0 < H < (1/2)$, then $L = \theta$ and $L(k) = k\theta$, which implies that L and $L(k)$ have the same sign

for all $k > 0$. Thus, the trivial solution of the system is almost surely exponentially stable and kth moment exponentially stable for all $k > 0$, if $\theta < 0$. Thus, there are no large deviations of the trivial solution for the exponential population growth system with fBm, which is different from case (i). However, if $H = (1/2)$ and $(1/2) < H < 1$, the conclusions are similar to that in cases (ii) and (iii), respectively. Therefore, Theorems 5, 6, and 7 are extensions of those reported in [17, 38, 55].

Remark 9. Large deviations are the most concise formulas of the relation between the almost sure exponential stability and the kth moment exponential stability of system (91). This relation is tight for small k; however, for sufficiently large k, $L < 0$ and $L(k) > 0$. In other words, the following phenomenon may happen: for fixed t, $P(t; t_0, P_0) \longrightarrow 0$ $(t \longrightarrow +\infty)$ with sample Lyapunov exponent L with probability one, while for $|P(t; t_0, P_0)|$ is still large with small probability (see [17], Remark 5.2). This very rare event makes $\mathbb{E}[|P(t; t_0, P_0)|^k]$ big for sufficiently large k and gives rise to the kth moment exponential instability in system (91) due to the kth moment Lyapunov exponent $L(k) > 0$, which is a very interesting phenomenon in biological population systems.

7. Conclusion

This paper investigates the dynamics of the exponential population growth system with mfBm. We first present some lemmas of the SDEs with mfBm. Then, we offer some explicit expressions, mathematical expectations, and variances of the solutions of the exponential population growth system with mfBm. Furthermore, we provide two necessary and sufficient conditions that prove the exponential stability in the kth moment and the almost sure stability of the constant coefficient exponential population growth system with mfBm. We also obtain the large deviations for the system, showing that the Hurst index affects the exponential stability and the large deviations. Notably, the solution of the suggested system of large deviations always exists when $(1/2) < H < 1$ due to the long-range dependence, which is an important property of the fBm. Finally, the methods and lemmas presented in this paper can be applied to stochastic networked systems with mfBm, such as multiagent systems, complex networks, and neural networks, to study the dynamics of consensus, synchronization, and stability.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

All authors contributed equally.

Acknowledgments

This study was supported in part by the Ningxia Key Research and Development Program (Introduction of Talents Project) (2018BEB04029 and 2020BEB04039), Ningxia Natural Science Foundation (2019AAC03029), and National Natural Science Foundation of China (62063028, 12001016, 62173139, and 61773217).

References

[1] Z. Zhang, S. Chen, and Y. Zheng, "Fully distributed scaled consensus tracking of high-order multiagent systems with time delays and disturbances," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 1, pp. 305–314, 2022.

[2] X. Mao, *Stochastic Differential Equations and Applications*, Horwood, Chichester, UK, second edition, 2007.

[3] W. Ma, X. Luo, and Q. Zhu, "Practical exponential stability of stochastic age-dependent capital system with Lévy noise," *Systems & Control Letters*, vol. 144, Article ID 104759, 2020.

[4] H. Yang and X. Li, "Explicit approximations for nonlinear switching diffusion systems in finite and infinite horizons," *Journal of Differential Equations*, vol. 265, no. 7, pp. 2921–2967, 2018.

[5] W. Ma, B. Ding, H. Yang, and Q. Zhang, "Mean-square dissipativity of numerical methods for a class of stochastic neural networks with fractional Brownian motion and jumps," *Neurocomputing*, vol. 166, pp. 256–264, 2015.

[6] Y. Zheng, J. Ma, and L. Wang, "Consensus of hybrid multiagent systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 4, pp. 1359–1365, 2017.

[7] H. Su, H. Wu, and J. Lam, "Positive edge-consensus for nodal networks via output feedback," *IEEE Transactions on Automatic Control*, vol. 64, no. 3, pp. 1244–1249, 2019.

[8] H. Su, Y. Sun, and Z. Zeng, "Semiglobal observer-based nonnegative edge consensus of networked systems with actuator saturation," *IEEE Transactions on Cybernetics*, vol. 50, no. 6, pp. 2827–2836, 2020.

[9] W. Cao and Q. Zhu, "Razumikhin-type theorem for pth exponential stability of impulsive stochastic functional differential equations based on vector Lyapunov function p th exponential stability of impulsive stochastic functional differential equations based on vector Lyapunov function," *Nonlinear Analysis: Hybrid Systems*, vol. 39, Article ID 100983, 2021.

[10] X. Yang and Q. Zhu, "Stabilization of stochastic functional differential systems by steepest descent feedback controls," *IET Control Theory & Applications*, vol. 15, no. 6, pp. 805–813, 2021.

[11] L. Arnold, E. Oeljeklaus, and E. Pardoux, "Almost sure and moment stability for linear ito equations," in *Lyapunov Exponents*, L. Arnold and V. Wihstutz, Eds., Springer-Verlag, Berlin, Germany, 1986.

[12] R. Khasminskii, *Stochastic Stability of Differential Equations*, Springer, Berlin, Germany, second edition, 2011.

[13] W. Q. Zhu, "Lyapunov exponent and stochastic stability of quasi-non-integrable Hamiltonian systems," *International Journal of Non-linear Mechanics*, vol. 39, no. 4, pp. 569–579, 2004.

[14] J. Bao, X. Mao, and C. Yuan, "Lyapunov exponents of hybrid stochastic heat equations," *Systems & Control Letters*, vol. 61, no. 1, pp. 165–172, 2012.

[15] L. Shaikhet, "Stability of a positive equilibrium state for a stochastically perturbed mathematical model of glassy-

winged sharpshooter population," *Mathematical Biosciences and Engineering*, vol. 11, no. 5, pp. 1167–1174, 2014.
[16] X. Fan and C. Yuan, "Lyapunov exponents of PDEs driven by fractional noise with Markovian switching," *Statistics & Probability Letters*, vol. 110, pp. 39–50, 2016.
[17] C. Zeng, Y. Chen, and Q. Yang, "Almost sure and moment stability properties of fractional order Black-Scholes model," *Fractional Calculus and Applied Analysis*, vol. 16, no. 2, pp. 317–331, 2013.
[18] M. Liu and K. Wang, "Stochastic Lotka-Volterra systems with Lévy noise," *Journal of Mathematical Analysis and Applications*, vol. 410, no. 2, pp. 750–763, 2014.
[19] Y. Zheng, W. Chen, and L. Wang, "Finite-time consensus for stochastic multi-agent systems," *International Journal of Control*, vol. 84, no. 10, pp. 1644–1652, 2011.
[20] Y. Zheng, Q. Zhao, J. Ma, and L. Wang, "Second-order consensus of hybrid multi-agent systems," *Systems & Control Letters*, vol. 125, pp. 51–58, 2019.
[21] W. Guo, Q. Zhang, and L. Rong, "A stochastic epidemic model with nonmonotone incidence rate: sufficient and necessary conditions for near-optimality," *Information Sciences*, vol. 467, pp. 670–684, 2018.
[22] Z. Yang, "Default probability of American lookback option in a mixed jump-diffusion model," *Physica A: Statistical Mechanics and Its Applications*, vol. 540, Article ID 123242, 2020.
[23] B. B. Mandelbrot and J. W. Van Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Review*, vol. 10, no. 4, pp. 422–437, 1968.
[24] Y. Mishura, *Stochastic Calculus for Fractional Brownian Motion and Related Processes*, Springer, Berlin, Germany, 2008.
[25] F. Biagini, Y. Hu, B. Øksendal, and T. Zhang, *Stochastic Calculus for Fractional Brownian Motion and Applications*, Springer, London, UK, 2008.
[26] T. E. Duncan, Y. Hu, and B. Pasik-Duncan, "Stochastic calculus for fractional brownian motion I. Theory," *SIAM Journal on Control and Optimization*, vol. 38, no. 2, pp. 582–612, 2000.
[27] Y. Hu, B. Øksendal, and D. M. Salopek, "Weighted local time for fractional brownian motion and applications to finance," *Stochastic Analysis and Applications*, vol. 23, pp. 1–30, 2005.
[28] Y. Hu, "Integral transformations and anticipative calculus for fractional Brownian motions," *Memoirs of the American Mathematical Society*, vol. 175, no. 825, 2005.
[29] W. Zhou, X. Zhou, J. Yang, J. Zhou, and D. Tong, "Stability analysis and application for delayed Neural Networks driven by fractional Brownian noise," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 5, pp. 1491–1502, 2018.
[30] M. Grigoriu, *Stochastic Calculus. Applications in Science and Engineering*, Birkhäuser, Boston, MA, USA, 2002.
[31] W.-J. Ma, Q.-M. Zhang, and C.-Z. Han, "Numerical analysis for stochastic age-dependent population equations with fractional Brownian motion," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 4, pp. 1884–1893, 2012.
[32] G. Jumarie, "New stochastic fractional models for Malthusian growth, the Poissonian birth process and optimal management of populations," *Mathematical and Computer Modelling*, vol. 44, no. 3-4, pp. 231–254, 2006.
[33] A. N. Kolmogorov, "Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum," *Academy URSS (N.S.)*, vol. 26, pp. 115–118, 1940, in German.
[34] H. E. Hurst, "Long-term storage capacity in reservoirs," *Transactions of the American Society of Civil Engineers*, vol. 116, pp. 400–410, 1951.
[35] L. Yan, "Maximal inequalities for the iterated fractional integrals," *Statistics & Probability Letters*, vol. 69, no. 1, pp. 69–79, 2004.
[36] C. Bender, "An Itô formula for generalized functionals of a fractional Brownian motion with arbitrary Hurst parameter," *Stochastic Processes and Their Applications*, vol. 104, no. 1, pp. 81–106, 2003.
[37] Y. Mishura and M. Zili, *Stochastic Analysis of Mixed Fractional Gaussian Processes*, Elsevier, Amsterdam, Netherlands, 2018.
[38] Q. Yang, C. Zeng, and C. Wang, "Fractional noise destroys or induces a stochastic bifurcation," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 23, no. 4, Article ID 043120, 2013.
[39] X. Mao, G. Marion, and E. Renshaw, "Environmental Brownian noise suppresses explosions in population dynamics," *Stochastic Processes and Their Applications*, vol. 97, no. 1, pp. 95–110, 2002.
[40] Y. Takeuchi, N. H. Du, N. T. Hieu, and K. Sato, "Evolution of predator-prey systems described by a Lotka-Volterra equation under random environment," *Journal of Mathematical Analysis and Applications*, vol. 323, no. 2, pp. 938–957, 2006.
[41] M. Khodabin, K. Maleknejad, M. Rostami, and M. Nouri, "Interpolation solution in generalized stochastic exponential population growth model," *Applied Mathematical Modelling*, vol. 36, no. 3, pp. 1023–1033, 2012.
[42] P. Cheridito, "Mixed fractional Brownian motion," *Bernoulli*, vol. 7, no. 6, pp. 913–934, 2001.
[43] L. Wang, R. Zhang, L. Yang, Y. Su, and F. Ma, "Pricing geometric Asian rainbow options under fractional Brownian motion," *Physica A: Statistical Mechanics and Its Applications*, vol. 494, pp. 8–16, 2018.
[44] W.-G. Zhang, Z. Li, and Y.-J. Liu, "Analytical pricing of geometric Asian power options on an underlying driven by a mixed fractional Brownian motion," *Physica A: Statistical Mechanics and Its Applications*, vol. 490, pp. 402–418, 2018.
[45] E. Alòs, O. Mazet, and D. Nualart, "Stochastic calculus with respect to Gaussian processes," *Annals of Probability*, vol. 29, pp. 766–801, 2001.
[46] P. Carmona, L. Coutin, and G. Montseny, "Stochastic integration with respect to fractional Brownian motion," *Annales de l'Institut Henri Poincare (B): Probability and Statistics*, vol. 39, no. 1, pp. 27–68, 2003.
[47] L. Coutin, D. Nualart, and C. A. Tudor, "Tanaka formula for the fractional Brownian motion," *Stochastic Processes and Their Applications*, vol. 94, no. 2, pp. 301–315, 2001.
[48] L. Decreusefond and A. S. Üstünel, "Stochastic analysis of the fractional Brownian motion," *Potential Analysis*, vol. 10, pp. 177–214, 1998.
[49] N. Privault, "Skorohod stochastic integration with respect to non-adapted processes on Wiener space," *Stochastics and Stochastics Reports*, vol. 65, no. 1-2, pp. 13–39, 1998.
[50] M. A. Arcones, "On the law of the iterated Logarithm for Gaussian processes," *Journal of Theoretical Probability*, vol. 8, no. 4, pp. 877–903, 1995.
[51] H. Touchette, "The large deviation approach to statistical mechanics," *Physics Reports*, vol. 478, no. 1-3, pp. 1–69, 2009.
[52] P. E. Gretener, "Significance of the rare event in geology," *AAPG Bulletin*, vol. 51, no. 11, pp. 2197–2206, 1967.
[53] G. Ruget, "Large deviations and more or less rare events in population dynamics," in *Rhythms in Biology and Other Fields*

of Application: Deterministic and Stochastic Approaches, pp. 388–400, Springer-Verlag, Berlin, Germany, 1983.

[54] D. R. Easterling, G. A. Meehl, C. Parmesan, S. A. Changnon, T. R. Karl, and L. O. Mearns, "Climate extremes: observations, modeling, and impacts," *Science*, vol. 289, no. 5487, pp. 2068–2074, 2000.

[55] C. Zeng, Y. Chen, and Q. Yang, "Almost sure and moment stability properties of LTI stochastic dynamic systems driven by fractional Brownian motion," in *Proceedings of the 51st IEEE Conference on Decision and Control*, pp. 4733–4736, Maui, HI, USA, December 2012.

Hindawi
Complexity
Volume 2021, Article ID 3866965, 8 pages
https://doi.org/10.1155/2021/3866965

Research Article

Cyclic Mappings and Further Results in B-Metric-Like Spaces

Shengquan Weng,[1] **Quanxin Zhu**,[2] **Baoying Du**,[1] **and Kaibo Shi**[3]

[1]*Faculty of Physics and Mathematics, Yibin University, Yibin 644000, China*
[2]*School of Mathematics and Statistics, Hunan Normal University, Changsha 410081, China*
[3]*School of Electronic Information and Electrical Engineering, Chengdu University, Chengdu 610106, China*

Correspondence should be addressed to Quanxin Zhu; zqx22@126.com

Received 11 September 2021; Accepted 26 November 2021; Published 26 December 2021

Academic Editor: Xiao Ling Wang

Fixed point problem of many mappings has been widely studied in the research work of fixed point theory. The generalized metric space is one of the research objects of fixed point theory. B-metric-like space is one of the generalized metric spaces; in fact, the research work in B-metric-like spaces is attractive. The intention of this paper is to introduce the concept of other cyclic mappings, named as $L\beta$-type cyclic mappings in the setting of B-metric-like space, study the existence and uniqueness of fixed point problem of $L\beta$-type cyclic mapping, and obtain some new results in B-metric-like spaces. Furthermore, the main results in this paper are illustrated by a concrete example. The work of this paper extend and promote the previous results in B-metric-like spaces.

1. Introduction

The fixed point theory is well known, and it is a very important branch of mathematics. In fact, the research of fixed point theory plays an important role in many aspects, especially in nonlinear analysis. In general, nonlinear analysis is a basic theory and method to deal with nonlinear problems from the perspective of mathematics. The nonlinear analysis is based on Banach space, Hilbert space, Sobolev space, paracompact space and unit decomposition, the properties of $\triangle$-Laplacian operator, the regularization theory of elliptic equation, Bochner integrability and vector value distribution, etc. It mainly includes topological degree theory and its application, convex analysis and optimization, monotone operator theory, variation and critical point theory, and branching theory. The category of fixed point theory is extensive. It is well known that, in 1922, Banach [1] published an essay, which became a classic work. Under the frame of metric space, the existence and uniqueness of a fixed point in terms of self-contractive mapping are studied. The Banach contraction principle is a classic result. It is an important tool in metric space theory. There are many extension of metric spaces, such as B-metric space [2], metric-like space [3], extended B-metric space [4], controlled metric-type space [5], double controlled metric-type space [6], double controlled metric-like space [7], probabilistic B-metric space [8], S-metric space [9, 10], and partial B-metric space [11]. In 2013, Alghamdi et al. [12] generalized the notion of three metric spaces, that is, B-metric space, metric-like space, and partial metric space. Based on this research work, they introduced the concept of an other special metric space and named it as B-metric-like space; furthermore, they established the theorems of existence and uniqueness of fixed points in this space. In 2018, Jleli and Samet [13] introduced the concept of F-metric space, which generalized the notion of metric spaces. A natural topology τF was defined in F-metric spaces, and they studied their topological properties. At the same time, they established a new version of Banach contraction principle in the setting of F-metric spaces. In order to illustrate their research, several examples were presented.

Definition 1 (see [13]). Let X be a nonempty set. Suppose that d: $X \times X \longrightarrow [0, +\infty)$ is a given mapping. Assume that there exist $(f, \alpha) \in tFn \times q[0, +\infty)$ such that the following three properties are satisfied:

(d1) $(x, y) \in X \times X$, $d(x, y) = 0 \Longleftrightarrow x = y$.
(d2) $d(x, y) = d(y, x)$ for all $(x, y) \in X \times X$.

(d3) For each $(x, y) \in X \times X$ and for every $N \in \mathbb{N}$, $N \geq 2$, and for each $(u_i)_i^N \subset X$, with $(u_1, u_N) = (x, y)$, we get

$$d(x, y) > 0 \Longrightarrow f(d(x, y)) \leq f\left(\sum_{i=1}^{N-1} D(x_i, x_{i+1})\right) + \alpha. \quad (1)$$

D is said to be an F metric on X, and the pair (X, D) is called an F-metric space.

From Definition 1, assume that F denotes the set of continuous function f, and the function f: $(0, +\infty) \longrightarrow R$ which satisfies some conditions is concretely given as follows:

(F1) f is nondecreasing; in detail, $0 < s < t \Longrightarrow f(s) \leq f(t)$.

(F2) $\lim_{n \longrightarrow \infty} \alpha_n = 0$ if and only if $\lim_{n \longrightarrow \infty} f(\alpha_n) = -\infty$ for every sequence $\{t_n\} \subseteq R^+$.

In fact, for the certain basic properties for the metric space, such as its third properties, it was modified with the set of continuous function F, for example, (d3) in Definition 1. Generalization work of metric spaces is important for fixed point theory. On the other hand, the extension of many mappings also is an important work. A generation of Banach contraction mapping, which appeared in F-metric spaces, was proposed and some coincident fixed point results were established as follows.

Theorem 1. *Assume that (X, D) is an F-metric space, and let F: $X \longrightarrow X$ be a given mapping that satisfies the following conditions:*

(i) The pair (X, D) is F-complete.

(ii) There exists $k \in (0, 1)$ which implies that

$$D(F(x), F(y)) \leq \mathrm{kD}(x, y). \quad (2)$$

Then, F has a unique fixed point $x^* \in X$. Moreover, for any $x_0 \in X$, $n \in \mathbb{N}$, the sequence $\{x_n\} \subset X$ defined by

$$x_{n+1} = F(x_n) \quad (3)$$

is F-convergent to x^*.

Likely, the concept of $\alpha - \psi$-contraction, which appeared in F-metric spaces [14], was considered; in addition, an interesting and important fixed point result was obtained as follows.

Theorem 2. *Assume that (X, D) is an F-metric space, and let the symbol F: $X \longrightarrow X$ be a β-admissible mapping and the following conditions are satisfied:*

(i) The pair (X, D) is F-complete.

(ii) There exist two functions β: $X \times X \longrightarrow [0, +\infty)$ and $\psi \in \Psi$ such that

$$\beta(x, y) D(F(x), F(y)) \leq \psi(M(x, y)), \quad (4)$$

where

$$M(x, y) = \max\{D(x, y), D(x, F(x)), D(y, F(y))\}, \quad (5)$$

for $x, y \in X$.

(iii) There exist $x_0 \in X$ such that $\beta(x_0, F(x_0))$.

Then, the mapping F contains a unique fixed point $x^* \in X$.

In 1969, Kannan [15] published an article, and some important results were obtained; the concept of Kannan mapping is introduced in this article. In 2003, Kirk et al. published a paper, and the intention of this paper is to introduce an important notion, that is, cyclic contraction mapping [16]. In 2010, Karapinar and Erhan [17] published an article, and they proposed a new concept and named it as Kannan-type cyclic mapping; then, they established some fixed point theorems. In 2018, Weng et al. published some articles, and the concept of LW-type Lipschitz cyclic mapping was proposed in the setting of B-metric-like space, and they got some fixed point results under certain suitable conditions. In 2019, Weng et al. published a paper, and a new cyclic mapping concept was proposed, that is, the general LW-type cyclic mapping, and some interesting results were obtained in some proper conditions. In 2020, Weng et al. also published an article, and in this article, a new kind of cyclic mapping was given and some meaningful results were obtained under some suitable conditions. In 2021, a new cyclic mapping was studied by Weng et al. and they called it as φ-type cyclic mapping [18]; some fixed point results from 2018 to 2021 can also be found in [18]. On the other hand, Huang et al. [19] in 2021 introduced the concept of F-contraction and established some fixed point theorems for such contractions in B-metric spaces. Thus, it can be seen that the generalization of research work of metric space and the extended research work of many kinds of mappings in different generalized metric spaces are simultaneous. In particular, it is essential to verify the applicability of Banach contraction mapping in every generalized metric spaces.

Through the continuous efforts of mathematical researchers in mathematical research, from the perspective of history, we can see further development of fixed point theory. In fact, it is a fascinating research hot spot to discuss and study the fixed point problem of new mapping in generalized metric space. Many researchers continue to explore the three basic properties of metric space, and they focus on nonnegativity, symmetry, and triangle inequality of metric spaces, take some examples of specific functions in some (two or three) dimensional spaces to show the reasonability of newly defined generalized metric spaces, and extend some classical conclusions to these newly defined generalized metric spaces. It is even very attractive for us to propose more new content according to the structure of some of these generalized metric spaces.

The research work for cyclic mappings is continuous and a series of works. The research content of this paper also belongs to the continuation of this series. In fact, the research work involving cyclic mapping is important, and it is not over yet. There is some research work that can be done. Inspired and encouraged by earlier research work

[4, 8, 11, 20, 21], another new cyclic mapping concept is proposed in this paper, that is, $L\beta$-type cyclic mapping (Definition 7 of Section 2). Moreover, the main results of this paper will be presented.

2. Preliminaries

The concept of B-metric-like space has been mentioned in Section 1; now the details of its definition will be presented as follows.

Definition 2 (see [12]). A B-metric-like space on a nonempty set X is a function r: $X \times X \longrightarrow [0, +\infty)$ such that for all $x, y, z \in X$ and a constant $s \geq 1$, the following three conditions hold true:

($\mathbf{r}1$) If $r(x, y) = 0 \Longrightarrow x = y$.

($\mathbf{r}2$) $r(x, y) = r(y, x)$.

($\mathbf{r}3$) $r(x, y) \leq s(r(x, z) + r(z, y))$.

The pair (X, r) is called a B-metric-like space.

In fact, let $X = [0, +\infty)$, suppose that the function r: $X \times X \longrightarrow [0, \infty)$ is the metric function, it can be defined concretely by $r(x, y) = (x + y)^4$, and it is easy to show the fact that conditions (r1) and (r2) also are satisfied. On the other hand, for arbitrary $x, y, z \in X$,

$$\begin{aligned} r(x, y) &= (x + y)^4 \\ &\leq (x + z + z + y)^4 \\ &= \left[(x + z)^2 + (z + y)^2 + 2(x + z)(z + y)\right]^2 \\ &= (x + z)^4 + (z + y)^4 + 6(x + z)^2 (z + y)^2 \\ &\quad + 4\left((x + z)^2 + (z + y)^2\right)\left)(x + z)(z + y)\right. \\ &\leq (x + z)^4 + (z + y)^4 + 3\left[(x + z)^4 + (z + y)^4\right] \\ &\quad + 2\left[(x + z)^2 + (z + y)^2\right)\Big] \\ &\leq 6(x + z)^4 + 6(z + y)^4 + 4(x + z)^2 (z + y)^2 \\ &\leq 8\left((x + z)^4 + (z + y)^4\right) \\ &= 8(r(x, z) + r(z + y)), \end{aligned} \tag{6}$$

so (r3) also holds. Then, it is obvious that (X, r) is a B-metric-like space with the parameter $s = 8$. Next, the definition of convergence, completeness, Cauchy sequence, and so on is given in a B-metric-like space.

Definition 3 (see [12]). Let (X, r) be a B-metric-like space. Suppose that $\{x_n\}$ is a sequence of points of X. A point $x \in X$ is said to be the limit of the sequence $\{x_n\}$ if $\lim_{n \longrightarrow \infty} r(x, x_n) = r(x, x)$, and we say that the sequence $\{x_n\}$ is convergent to x and denote it by $x_n \longrightarrow x$ as $n \longrightarrow \infty$.

Definition 4 (see [12]). Let (X, r) be a B-metric-like space. A sequence $\{x_n\}$ is called Cauchy if and only if $\lim_{n \longrightarrow \infty} r(x_n, x_m)$ exists and is finite. A B-metric-like space (X, r) is said to be complete if and only if every Cauchy sequence $\{x_n\}$ in X converges to $x \in X$ so that $\lim_{m,n \longrightarrow \infty} r(x_n, x_m) = r(x, x) = \lim_{n \longrightarrow \infty} r(x_n, x)$.

Definition 5 (see [12]). Let G_1, G_2 be nonempty sets of metric space; if $B(G_1) \subset G_2$ and $S(G_2) \subset G_1$, then the mapping (B, S): $G_1 \times G_2 \longrightarrow G_2 \times G_1$ is called as a pair semicyclic mapping, where B is said to be a lower semicyclic mapping and S is said to be a upper semicyclic mapping. If $B = S$, then B is said to be a cyclic mapping.

Definition 6 (see [22]). If $\prec$ is partially ordered in B-metric-like spaces (X, r), then $(X, r, \prec)$ is a partially ordered B-metric-like space.

The definition of $L\beta$-type cyclic mapping is given as follows.

Definition 7. Let (X, r) be a B-metric-like space. Assume that G_1 and G_2 are nonempty closed two subsets of X. If (B, S) is a pair semicyclic mapping in $G_1 \times G_2$, there exists γ, δ, L, β, and γ, δ, L, β are certain real nonnegative numbers, and $L \leq \beta$, $\gamma + \delta = \beta$, $\beta \in (0, 1]$ such that for all $x \in G_1$, $y \in G_2$, the following inequality is satisfied:

$$r(\mathrm{Bx}, \mathrm{Sy}) \leq L[\beta r(x, y) - M(x, y)], \tag{7}$$

where $M(x, y) = \max\{\delta r(y, \mathrm{Sy}), \gamma r(x, \mathrm{Bx})\}$. Thus, (B, S) is called a $L\beta$-type cyclic mapping.

What calls for special attention is that the $L\beta$-type cyclic mapping is a W-type cyclic mapping while $\beta = 1$, and the expression of $L\beta$-type cyclic mapping is different from the W-type cyclic mapping, and the expression of $L\beta$-type cyclic mapping is further promotion of W-type cyclic mapping; then, it is innovative.

3. Fixed Point Results

The main content of this article is to research the fixed point problem of a new mapping. In this section, some main results for $L\beta$-type cyclic mapping are given under some suitable conditions.

Theorem 3. *Assume that (X, r) is a B-metric-like space and (X, r) is complete. Suppose that (B, S) is a $L\beta$-type cyclic mapping, and G_1 and G_2 are nonempty closed two subsets of X, $G_1 \cap G_2 \neq \varnothing$. Let $p = \max\{L\beta/(1 + L\delta), L\delta, L\gamma, L\beta/(1 + L\gamma)\}$ and $p < 1/s$. Then, there exists a unique z and $z \in G_1 \cap G_2$ such that $Bz = z = Sz$. In other words, the mappings B and S have a unique common fixed point.*

Proof. Let the sequence $\{z_n\}$ be defined in the following manner:

$$\begin{aligned} z_0 \in G_1,\ z_1 = \mathrm{Bz}_0,\ z_2 = \mathrm{Sz}_1,\ z_3 = \mathrm{Bz}_2,\ z_4 = \mathrm{Sz}_3, \ldots, z_{2n+1} \\ = \mathrm{Bz}_{2n}, z_{2n+2} = \mathrm{Sz}_{2n+1}, \ldots,\ n \geq 0. \end{aligned} \tag{8}$$

□

Step 1. The first step is to show that $\{z_n\}$ is a Cauchy sequence. Through the Definition 7, this shows that

$$L[\beta r(z_0, z_1) - M(z_0, z_1)] \geq r(\mathrm{Bz}_0, \mathrm{Sz}_1), \tag{9}$$

that is,

$$L[\beta r(z_0, z_1) - M(z_0, z_1)] \geq r(z_1, z_2). \tag{10}$$

Because $M(x, y) = \max\{\delta r(y, \mathrm{Sy}), \gamma r(x, \mathrm{Bx})\}$, this is obvious, and it needs to be discussed case by case.

While $M(z_0, z_1) = \gamma r(z_0, \mathrm{Bz}_0) = \gamma r(z_0, z_1)$, it shows by (10) that

$$L[\beta r(z_0, z_1) - \gamma r(z_0, z_1)] \geq r(z_1, z_2), \tag{11}$$

that is,

$$L(\beta - \gamma) r(z_0, z_1) \geq r(z_1, z_2). \tag{12}$$

In fact, $\gamma + \delta = \beta$, and this implies that

$$r(z_1, z_2) \leq L\delta r(z_0, z_1). \tag{13}$$

While $M(z_0, z_1) = \delta r(z_1, \mathrm{Sz}_1) = \delta r(z_1, z_2)$, this show that

$$L[\beta r(z_0, z_1) - \delta r(z_1, z_2)] \geq r(z_1, z_2), \tag{14}$$

after appropriate transformation can be obtained, that is,

$$r(z_1, z_2) \leq \frac{L\beta}{1 + L\delta} r(z_0, z_1). \tag{15}$$

Because $L[\beta r(z_2, z_1) - M(z_2, z_1)] \geq r(\mathrm{Bz}_2, \mathrm{Sz}_1) = r(z_3, z_2)$, it follows the above discussion.

While $M(z_2, z_1) = \gamma r(z_2, \mathrm{Bz}_2) = \gamma r(z_2, z_3)$, it shows that

$$L[\beta r(z_2, z_1) - \gamma r(z_2, z_3)] \geq r(z_3, z_2), \tag{16}$$

and after appropriate transformation, the following can be obtained:

$$r(z_3, z_2) = r(z_2, z_3) \leq \frac{L\beta}{1 + L\gamma} r(z_2, z_1) = \frac{L\beta}{1 + L\gamma} r(z_1, z_2), \tag{17}$$

that is,

$$r(z_2, z_3) \leq \frac{L\beta}{1 + L\gamma} r(z_1, z_2). \tag{18}$$

While $M(z_2, z_1) = \delta r(z_1, \mathrm{Sz}_1) = \delta r(z_1, z_2)$, this implies that

$$L[\beta r(z_2, z_1) - \delta r(z_1, z_2)] \geq r(z_3, z_2), \tag{19}$$

and after appropriate transformation can be obtained, by $\gamma + \delta = \beta$, we can get

$$r(z_2, z_3) \leq L\gamma r(z_1, z_2). \tag{20}$$

Since $L \leq \beta$, $\gamma + \delta = \beta$, $\beta \in (0, 1]$,

$$0 < L\delta < 1, 0 < L\gamma < 1, 0 < \frac{L\beta}{1 + L\gamma} < 1, 0 < \frac{L\beta}{1 + L\delta} < 1. \tag{21}$$

Because $p = \max\{L\delta, L\gamma, (L\beta/(1 + L\gamma)), (L\beta/(1 + L\delta))\}$, it shows that $p \in (0, 1)$. By virtue of (13)–(20), this implies that

$$r(z_2, z_3) \leq \mathrm{pr}(z_1, z_2) \leq p^2 r(z_0, z_1). \tag{22}$$

Repeat the above work over and over again, so as to get

$$r(z_n, z_{n+1}) \leq p^n r(z_0, z_1), \forall n \in N. \tag{23}$$

Furthermore, for arbitrary $m, n \in N$, and $n \leq m$, we have

$$\begin{aligned} r(z_n, z_m) &\leq \mathrm{sr}(z_n, z_{n+1}) + \mathrm{sr}(z_{n+1}, z_m) \\ &\leq \mathrm{sr}(z_n, z_{n+1}) + s^2 r(z_{n+1}, z_{n+2}) + s^2 r(z_{n+2}, z_m) \\ &\leq s(p)^n r(z_0, z_1) + s^2 (p)^{n+1} r(z_0, z_1) + \cdots + s^{m-n} \\ &\quad \cdot (p)^{m-1} r(z_0, z_1) \\ &\leq \left[s(p)^n + s^2 (p)^{n+1} + \cdots + s^{m-n} (p)^{m-1}\right] r(z_0, z_1) \\ &= \frac{1 - (\mathrm{sp})^{m-n}}{1 - \mathrm{sp}} \mathrm{sp}^n r(z_0, z_1) \\ &\leq \frac{1}{1 - \mathrm{sp}} (p)^{n-1} r(z_0, z_1). \end{aligned} \tag{24}$$

Since $p < 1/s$, from (24) and letting $n \longrightarrow \infty$, we get that

$$\lim_{n \longrightarrow \infty} r(z_n, z_m) = 0. \tag{25}$$

By (25), it shows that $\{z_n\}$ is a Cauchy sequence.

Step 2. Since $G_1 \cap G_2 \neq \emptyset$, (X, r) is complete, so it is easy to know that there exists a point $z \in X$ such that

$$z_n \longrightarrow z \ (n \longrightarrow \infty). \tag{26}$$

Then, it shows that

$$z_{2n} \longrightarrow z;\ z_{2n+1} \longrightarrow z \ (n \longrightarrow \infty). \tag{27}$$

Because $\{z_{2n}\} \subset G_1$, $\{z_{2n+1}\} \subset G_2$, and G_1, G_2 are closed, then

$$z \in G_1 \cap G_2. \tag{28}$$

Step 3. Next, the fact that a point z is a common fixed point of the mapping S and B will be proved in this step. Since (B, S) is a $L\beta$-type cyclic mapping, the following inequality is easy to obtain:

$$r(\mathrm{Bz}, \mathrm{Sz}) \leq L[\beta r(z, z) - M(z, z)]. \tag{29}$$

Now, it needs to be discussed case by case.

While $M(z, z) = \gamma r(z, \mathrm{Bz})$, it shows that

$$r(\mathrm{Bz}, \mathrm{Sz}) \leq L[\beta r(z, z) - \gamma r(z, \mathrm{Bz})], \tag{30}$$

that is,

$$r(\mathrm{Bz}, \mathrm{Sz}) + L\gamma r(z, \mathrm{Bz}) \leq L\beta r(z, z), \tag{31}$$

Based on the equality $\lim_{n,m\longrightarrow\infty} r(z_n, z_m) = r(z, z)$ and making use of (25) and (31), we can get

$$r(Bz, z) = 0. \tag{32}$$

Conditions of Definition 7 are considered, and it shows that

$$z = \mathrm{Bz}. \tag{33}$$

While $M(z, z) = \delta r(z, \mathrm{Sz})$, this shows that

$$r(\mathrm{Bz}, \mathrm{Sz}) \le L[\beta r(z, z) - \delta r(z, \mathrm{Sz})], \tag{34}$$

that is,

$$r(\mathrm{Bz}, \mathrm{Sz}) + L\delta r(z, \mathrm{Sz}) \le L\beta r(z, z). \tag{35}$$

Since $\lim_{n,m\longrightarrow\infty} r(z_n, z_m) = r(z, z)$ and making use of equality (25) and inequality (35), it implies that

$$r(\mathrm{Sz}, z) = 0. \tag{36}$$

By virtue of the condition of Definition 7, it shows that

$$z = \mathrm{Sz}. \tag{37}$$

By (33) and (37), it implies that

$$\mathrm{Bz} = z = \mathrm{Sz}. \tag{38}$$

Step 4. In this step, the fact that B and S have a unique common fixed point will be fully proved. Now, suppose that $z^*, z \in X$ are two different common fixed points for the pair B and S in (X, r). So, making use of Definition 7 and (38), we have

$$r(\mathrm{Bz}, \mathrm{Sz}^*) \le L[\beta r(z, z^*) - M(z, z^*)]. \tag{39}$$

Discussion is presented next.

While $M(z, z^*) = \gamma r(z, \mathrm{Bz}) = \gamma r(z, z)$, it shows that

$$r(\mathrm{Bz}, \mathrm{Sz}^*) \le L[\beta r(z, z^*) - \gamma r(z, \mathrm{Bz})], \tag{40}$$

$$r(z, z^*) \le L[\beta r(z, z^*) - \gamma r(z, z)], \tag{41}$$

$$r(z, z^*) \le L\beta r(z, z^*) - L\gamma r(z, z), \tag{42}$$

$$r(z, z^*) + L\gamma r(z, z) \le L\beta r(z, z^*). \tag{43}$$

While $M(z, z^*) = \delta r(z^*, \mathrm{Bz}^*) = \delta r(z^*, z^*)$, it shows that

$$r(\mathrm{Bz}, \mathrm{Sz}^*) \le L[\beta r(z, z^*) - \delta r(z^*, z^*)], \tag{44}$$

$$r(z, z^*) \le L[\beta r(z, z^*) - \delta r(z^*, z^*)], \tag{45}$$

$$r(z, z^*) \le L\beta r(z, z^*) - L\delta r(z^*, z^*), \tag{46}$$

$$r(z, z^*) + L\gamma r(z^*, z^*) \le L\beta r(z, z^*). \tag{47}$$

From Definition 7 and formulas (25), (43), and (47), these show that

$$r(z, z^*) \le L\beta r(z, z^*). \tag{48}$$

Since $L \le \beta$, $\beta \in (0, 1]$, L, β are nonnegative real constants, it shows that $L\beta \le 1$. In fact, through inequality (48), it is easy to know that $L\beta \ge 1$, and this is a contradiction. Thus,

$$r(z, z^*) = 0. \tag{49}$$

That is,

$$z = z^*. \tag{50}$$

This completes the proof.

In Section 2, the definition of the partially ordered B-metric-like space is given, and then the result of Theorem 3 is extended to this partially ordered B-metric-like space, and thus the following corollary can be obtained.

Corollary 1. *Assume that (X, r) is a partially ordered B-metric-like space and (X, r) is complete. Suppose that (B, S) is a $L\beta$-type cyclic mapping, and G_1 and G_2 are nonempty closed two subsets of X, $G_1 \cap G_2 \neq \varnothing$. Let $p = \max\{L\beta/(1 + L\delta), L\delta, L\gamma, L\beta/(1 + L\gamma)\}$ and $p < 1/s$. Then, there exists a unique $z \in G_1 \cap G_2$ such that $Bz = z = Sz$. In other words, the mappings B and S have a unique common fixed point.*

Remark 1. From Definition 7 and Theorem 3, when $\beta = 1$, a $L\beta$-type cyclic mapping can be seen as a W-type cyclic mapping. In fact, the condition $\gamma + \delta = 1$ is replaced by $\gamma + \delta = \beta$. Thus, a W-type cyclic mapping is a special case of $L\beta$-type cyclic mapping. It is obvious that the $L\beta$-type cyclic mapping is more extensive than a W-type cyclic mapping. The research extension work involving cyclic mappings and B-metric-like space is important. This set of results will become the part of the basis for discovering more properties in B-metric-like spaces. In order to show the effectiveness of the $L\beta$-type cyclic mapping, a concrete example is given as follows.

4. Example

In this section, a concrete example is given to illustrate the effectiveness of the $L\beta$-type cyclic mapping and show the rationality of the obtained theorems.

4.1. Part One. In this part, a concrete B-metric-like space is presented; at the same time, the mappings B and S are also defined in this part. Based on these preparations, the validity of the definition of the $L\beta$-type cyclic mapping and the correctness of Theorem 3 are verified.

Consider the set $X = \{0, 1, 2\}$ and let the metric function r: $X \times X \longrightarrow [0, +\infty)$ be defined by

$$\begin{aligned} &r(0, 0) = 0, r(1, 1) = 0, r(2, 2) \\ &\quad = 1, r(0, 1) = 4, r(1, 0) = 4, \end{aligned} \tag{51}$$

$$r(0, 2) = 2, r(2, 0) = 2, r(1, 2) = 1, r(2, 1) = 1. \tag{52}$$

It is clear that (X, r) is a B-metric-like complete space with the constant $s = 1.4$ (see [23]).

Recall $L\beta$-type cyclic mapping concept. Assume that G_1 and G_2 are nonempty closed two subsets of X. If (B, S) is a pair semicyclic mapping in $G_1 \times G_2$, there exists γ, δ, L, β, γ, δ, L, β are some real nonnegative numbers, and $L \leq \beta$, $\gamma + \delta = \beta$, $\beta \in (0, 1]$ such that for all $x \in G_1$, $y \in G_2$, the following inequality is satisfied:

$$r(\mathrm{Bx}, \mathrm{Sy}) \leq L[\beta r(x, y) - M(x, y)], \tag{53}$$

and we should need to pay attention to the symbol $M(x, y)$, where $M(x, y) = \max\{\gamma r(x, \mathrm{Bx}), \delta r(y, \mathrm{Sy})\}$. Thus, (B, S) is called as the $L\beta$-type cyclic mapping. Then, the mappings B and S are defined by the following way. Let

$$G_1 = \{0, 1\},\ G_2 = \{1, 2\}, G_1, G_2 \in X, \tag{54}$$

$$B(0) = 2, B(1) = 1, S(1) = 1, S(2) = 1. \tag{55}$$

Next, we explore the existence and uniqueness of fixed points for B and S under the corresponding conditions of Definition 7 and Theorem 3 in a concrete, complete B-metric-like space. In fact, when the mappings B and S are concretely defined, it can get four points, that is, $r(x, y)$ should be $r(0, 1), r(0, 2), r(1, 1), r(1, 2)$.

Thus, some concrete calculation will be shown as follows.

Situation 1. If we choose $r(x, y) = r(0, 1) = 4$, then

$$r(B(0), S(1)) \leq L[\beta r(0, 1) - M(0, 1)], \tag{56}$$

that is,

$$r(2, 1) \leq L[\beta r(0, 1) - M(0, 1)]. \tag{57}$$

In fact,

$$1 \leq L[4\beta - M(0, 1)], \tag{58}$$

where

$$\begin{aligned} M(0, 1) &= \max\{\gamma r(0, B(0)), \delta r(1, S(1))\} \\ &= \max\{\gamma r(0, 2), \delta r(1, 1)\} \\ &= \max\{2\gamma, 0\}, \end{aligned} \tag{59}$$

and thus it should be discussed case by case.

While $M(0, 1) = 2\gamma$, it can get the inequality

$$1 \leq L(4\beta - 2\gamma), \tag{60}$$

that is,

$$1 \leq 2L\beta + L\delta. \tag{61}$$

While $M(0, 1) = 0$, it can get the inequality

$$1 \leq L(4\beta - 0), \tag{62}$$

that is,

$$1 \leq 4L\beta. \tag{63}$$

Situation 2. If we choose $r(x, y) = r(0, 2) = 2$, then

$$r(B(0), S(2)) \leq L[\beta r(0, 2) - M(0, 2)], \tag{64}$$

that is,

$$r(2, 1) \leq L[\beta r(0, 2) - M(0, 2)]. \tag{65}$$

In fact,

$$1 \leq L[2\beta - M(0, 2)], \tag{66}$$

where

$$\begin{aligned} M(0, 2) &= \max\{\gamma r(0, B(0)), \delta r(2, S(2))\} \\ &= \max\{\gamma r(0, 2), \delta r(2, 1)\} \\ &= \max\{2\gamma, \delta\}, \end{aligned} \tag{67}$$

and thus it should be discussed case by case.

While $M(0, 2) = 2\gamma$, it can get the inequality

$$1 \leq L(2\beta - 2\gamma), \tag{68}$$

that is,

$$1 \leq 2L\delta. \tag{69}$$

While $M(0, 2) = \delta$, it can get the inequality

$$1 \leq L(2\beta - \delta), \tag{70}$$

that is,

$$1 \leq L\beta + L\gamma. \tag{71}$$

Situation 3. If we choose $r(x, y) = r(1, 1) = 0$, then

$$r(B(1), S(1)) \leq L[\beta r(1, 1) - M(1, 1)], \tag{72}$$

that is,

$$0 \leq L[0 - M(1, 1)], \tag{73}$$

where

$$\begin{aligned} M(1, 1) &= \max\{\gamma r(1, B(1)), \delta r(1, S(1))\} \\ &= \max\{\gamma r(1, 1), \delta r(1, 1)\} \\ &= \max\{0, 0\} \\ &= 0, \end{aligned} \tag{74}$$

and thus it should be discussed case by case. In fact, arbitrary value of parameters γ, δ, β, L is true in this situation.

Situation 4. If we choose $r(x, y) = r(1, 2) = 1$, then

$$r(B(1), S(2)) \leq L[\beta r(1, 2) - M(1, 2)], \tag{75}$$

that is,

$$r(1, 1) \leq L[\beta r(1, 2) - M(1, 2)]. \tag{76}$$

In fact,

$$0 \leq L[\beta - M(1, 2)], \tag{77}$$

where

$$\begin{aligned} M(1,2) &= \max\{\gamma r(1,B(1)), \delta r(2,S(2))\} \\ &= \max\{\gamma r(1,1), \delta r(1,1)\} \\ &= \max\{0,\delta\}, \end{aligned} \tag{78}$$

and thus it should be discussed case by case.

While $M(1,2) = 0$, it can get the inequality

$$0 \le L\beta. \tag{79}$$

While $M(1,2) = \delta$, it can get the inequality

$$0 \le L(\beta - \delta), \tag{80}$$

that is,

$$0 \le L\gamma. \tag{81}$$

4.2. Part Two. In this part, the validity of the conditions of the theorem is verified. Furthermore, three sets of data are used to verify the existence and the rationality of the theorem conditions.

Let $\gamma = \delta = 1/2, \beta = 1, L = 1$, and take this set of data into the above inequalities of situations 1-4 for verification, and all the inequalities will be true. Because $p = \max\{L\beta/(1+L\delta), L\delta, L\gamma, L\beta/(1+L\gamma)\}$, then $p = \max\{1/1.5, 0.5, 0.5, 1/1.5\} = 0.6667$. So, $\text{sp} = 1.4 \times 0.6667 = 0.9334 < 1$. It is easy to know that the condition $s < 1/p$ is satisfied.

Let $\gamma = 0.47, \delta = 0.52, \beta = 0.99, L = 0.98$, and take this set of data into the above inequalities of situations 1-4 for verification, and all the inequalities will be true. Because $p = \max\{L\beta/(1+L\delta), L\delta, L\gamma, L\beta/(1+L\gamma)\}$, then $p = \max\{0.98 \times 0.99/(1+0.98\times 0.52), 0.98\times 0.52, 0.98\times 0.47, 0.98 \times 0.99/(1+0.98\times 0.47)\} = 0.6642$. So, $\text{sp} = 1.4 \times 0.6642 = 0.9299 < 1$. It is obvious that the condition $s < 1/p$ is satisfied.

Let $\gamma = 0.45, \delta = 0.53, \beta = 0.98, L = 0.96$, and take this set of data into the above inequalities of situations I–4 for verification, and all the inequalities will be true. Because $p = \max\{L\beta/(1+L\delta), L\delta, L\gamma, L\beta/(1+L\gamma)\}$, then $p = \max\{0.96 \times 0.98/(1+0.96\times 0.53), 0.96\times 0.53, 0.96\times 0.45, 0.96 \times 0.98/(1+0.96\times 0.45)\} = 0.65698$. So, $\text{sp} = 1.4 \times 0.65698 = 0.9198 < 1$. It is easy to verify the fact that the condition $s < 1/p$ is satisfied.

Overall, after concrete data validation above, because G_1 and G_2 are two nonempty closed subsets in (X, r), B and S have a unique common fixed point. Therefore, it is verified by an example to show that the theorem holds.

5. Conclusions

A special note is needed. In fact, B-metric-like space is one of the generalized metric spaces, and certain related research work in different generalized metric spaces for promotion is very important. Some research work about cyclic mappings under the B-metric-like space frame is presented in this paper. The $L\beta$-type cyclic mapping proposed in this paper is discussed and researched, and the corresponding fixed point theorems are obtained. The correctness of the theoretical results is verified by a concrete real example. The correctness of the main results in this paper has been verified and is correct. Actually, $L\beta$-type cyclic mappings are not difficult to find to become a W-type cyclic mapping while $\beta = 1$. Thus, the $L\beta$-type cyclic mapping is more extensive than a W-type cyclic mapping. Before this, phi-type cyclic mapping, W-type cyclic mapping [21], and some other cyclic mappings have been proposed and studied by us under the frame of B-metric-like spaces; when these cyclic mappings were further studied and extended, some fixed point theorems were obtained. In fact, these research works in B-metric spaces are beneficial to the further enrichment of B-metric-like space theory Of course, the current research work is not deep; in fact, some of the conditions presented in these research papers may be able to be further adjusted and optimized; as for the follow-up research, we will do our best. Also, it is interesting to apply our results to network systems [24, 25] and stochastic differential equation [26].

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

All authors contributed equally to this study.

Acknowledgments

This study was jointly supported by the High-level Talent Sailing Project of Yibin University (2021QH07).

References

[1] S. Banach, "Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales," *Fundamenta Mathematicae*, vol. 3, no. 1, pp. 133–181, 1922.

[2] S. Czerwik, "Contraction mappings in B-metric spaces," *Acta Mathematica et Informatica Universitatis Ostraviensis*, vol. 1, no. 1, pp. 5–11, 1993.

[3] A. Amini Harandi, "Metric-like spaces, partial metric spaces and fixed points," *Journal of Fixed Point Theory and Appllication*, vol. 2012, Article ID 204, 2012.

[4] T. Kamran, M. Samreen, and Q. UlAin, "A generalization of B-metric space and some fixed point theorems," *Mathematics*, vol. 5, Article ID 19, 2017.

[5] N. Mlaiki, H. Aydi, N. Souayah, and T. Abdeljawad, "Controlled metric type spaces and the related contraction principle," *Mathematics*, vol. 6, no. 10, Article ID 194, 2018.

[6] T. Abdeljawad, N. Mlaiki, H. Aydi, and N. Souayah, "Double controlled metric type spaces and some fixed point results," *Mathematics*, vol. 6, Article ID 320, 2018, https://doi.org/10.3390/math6120320.

[7] N. Mlaiki, "Double controlled metric-like spaces," *Journal of Inequalities and Applications*, vol. 2020, Article ID 189, 2020, https://doi.org/10.1186/s13660-020-02456-z.

[8] A. Mbarki and R. Oubrahim, "Probabilistic B-metric spaces and nonlinear contractions," *Journal of Fixed Point Theory and Application*, vol. 29, p. 15, 2017.

[9] S. Sedghi, N. Shobeand, and T. Doenovic, "Fixed point results in S-metric spaces," *Nonlinear Functional Analysis and Applications*, vol. 20, no. 1, pp. 55–67, 2015.

[10] A. Singh and N. Hooda, "Coupled fixed point theorems in S-metric spaces," *Inter. J. Math. Stat. Invent*, vol. 2, pp. 33–39, 2014.

[11] S. Shukla, "Partial B-metric spaces and fixed point theorems," *Mediterranean Journal of Mathematics*, vol. 11, no. 2, pp. 703–711, 2014.

[12] M. A. Alghamdi, N. Hussain, and P. Salimi, "Fixed point and coupled fixed point theorems on B-metric-like spaces," *Journal of Inequalities and Applications*, vol. 2013, no. 1, Article ID 402, 2013.

[13] M. Jleli and B. Samet, "On a new generalization of Metric Spaces," *Journal of fixed point theory and Application*, vol. 20, p. 20, 2018.

[14] A. Hussain and T. kanwal, "Existence and uniqueness for a neutral differential problem with unbounded delay via fixed point results," *Transactions of A. Razmadze Mathematical Institute*, vol. 172, no. 3, pp. 481–490, 2018.

[15] R. Kannan, "Some results on fixed points-II," *The American Mathematical Monthly*, vol. 76, no. 4, pp. 405–408, 1969.

[16] W. A. Kirk, P. S. Srinivasan, and P. Veeramani, "Fixed points for mappings satisfying cyclical contractive conditions," *Fixed Point Theory*, vol. 4, pp. 79–89, 2003.

[17] E. Karapinar and I. M. Erhan, "Best proximity on different type contractions," *Appl.Math.Inf.Sci*, vol. 25, no. 5, pp. 558–569, 2010.

[18] S. Q. Weng and Q. X. Zhu, "Some fixed point theorems on generalized cyclic mappings in B-metric-like spaces," *Complexity*, vol. 2021, Article ID 9042402, 7 pages, 2021.

[19] H. Huang, Z. D. Mitrović, K. Zoto, S. Radenović, and R. Stojan, "On convex F-contraction in B-metric spaces," *Axioms*, vol. 10, no. 2, p. 71, 2021, https://doi.org/10.3390/axioms10020071.

[20] T. Dosenovic, M. De La Sen, L. Paunovic, R. Dusan, and S. Radenovic, "Some new observations on generalized contractive mappings and related results in B-metric-like spaces," *Journal of Mathematics*, vol. 2021, Article ID 6634822, 9 pages, 2021.

[21] S. Q. Weng and L. Liang, "A fxied-point theorem of cyclic mappings in complete B-metric-like spaces," *Journal of Yibin University*, vol. 21, no. 6, pp. 77–80, 2021.

[22] A. HAydi, S. Felhi, and Sahmim, "On common fixed points for (α, ψ)-contractions and generalized cyclic contractions in B-metric-like spaces and consequences," *The Journal of Nonlinear Science and Applications*, vol. 9, no. 5, pp. 2492–2510, 2016.

[23] H. Nashine and Z. Kadelburg, "Existence of solutions of cantilever beam problem via (α-β-FG)-contractions in B-metric-like spaces," *Filomat*, vol. 31, no. 11, pp. 3057–3074, 2017.

[24] H. Su, Y. Sun, and Z. Zeng, "Semiglobal observer-based nonnegative edge consensus of networked systems with actuator saturation," *IEEE Transactions on Cybernetics*, vol. 50, no. 6, pp. 2827–2836, 2020.

[25] F. Kong and Q. Zhu, "New fixed-time synchronization control of discontinuous inertial neural networks via indefinite Lyapunov-Krasovskii functional method," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 2, pp. 471–495, 2021.

[26] X. Zhang, S. P. He, V. Stojanovic, X. Luan, and F. Liu, "Finite-time asynchronous dissipative filtering of conic-type nonlinear Markov jump systems," *Science China Information Sciences*, vol. 64, no. 5, Article ID 152206, 2021.

Hindawi
Complexity
Volume 2021, Article ID 4167340, 17 pages
https://doi.org/10.1155/2021/4167340

WILEY | Hindawi

Research Article

The Influence of Substituting Prices, Product Returns, and Service Quality on Repurchase Intention

Jian Xue,[1] **Zeeshan Rasool**,[1] **Mohsin Ali Khan**,[2] **Ahmad Imran Khan**,[3] **Farooq Khan**,[2] **Anum Afzal Khan**,[3] **Rabail Shoukat**,[4] **and Hina Ali**[5]

[1]*School of Economics and Management, Shaanxi University of Science and Technology, Xi'an 710021, China*
[2]*National University of Modern Languages Multan Campus, Multan 60000, Pakistan*
[3]*Putra Business School, University of Putra, UPM, Serdang, Seri Kembangan 43400, Selangor, Malaysia*
[4]*Management Science Department, Bahria University Karachi Campus, Karachi, Pakistan*
[5]*The Women University Multan, Multan, Pakistan*

Correspondence should be addressed to Zeeshan Rasool; zeeshan_rasool114@hotmail.com

Received 28 August 2021; Revised 18 October 2021; Accepted 25 October 2021; Published 27 November 2021

Academic Editor: Yue Song

The textile industry is production-intensive and incorporates diverse transactions made by multiple suppliers, corporate buyers, and supply chain members. In the business-to-business context, the cost for attracting a new customer is notably much higher than that needed to retain a present one. Customer loyalty in terms of customer repurchase intention has, therefore, been considered as a key determinant for textile companies to improve their efficiency and competitive advantage. This study aims to investigate the business-to-business repurchase intentions of Pakistan textile and clothing industry customers. The study framework specifically consolidates the mutual dynamics of appealing (service quality), facilitating (product returns), and averting (switching costs) factors altogether and the effect of these variables on customer satisfaction and thus on customer retention (repurchase intent) in the textile's transactional scenario. A sample survey method is used for this study. The data collected through self-administered questionnaires ($n = 325$) from All Pakistan Textile Mills Association enlisted the employees of the companies. The structural equational modeling technique was applied to examine the study hypotheses. The findings contended that service quality and switching costs are essential determinants that shape the repurchase intentions. Therefore, product returns do not contribute toward customer satisfaction and also do not shape the intentions of business-to-business customers to repurchase from the same supplier after having even a good product return experience in past.

1. Introduction

Pakistan is currently positioned as the fifth largest cotton producer in the world and has capitalized on its capability for promoting and developing its textile sector. The country's established textile industry is steered by the availability of locally produced cotton [1]. Globally, Pakistan is ranked third in yarn production and it has a domestically well-expanded textile industry from ginning to made-ups (ginning, spinning, weaving, processing, dyeing, printing, and made-ups) [2]. The industry is considered the backbone of Pakistan's economy and it is one of the major contributors, both in terms of exports and employment. The industry eventually contributes a big part in the country's economy [3].

Regardless of the mature and progressive textile value chain, the Pakistan textile industry's performance is disappointing and it is not generating the returns according to its potential [4]. In a broader spectrum, the major issues of the textile industry have been the inadequate implementation of quality practices, inadequate customer services, delay or hesitation in product returns, poor complaint management system, failure in observing the production lead times, and cost of the value chain [5]. Despite Pakistan's government textile industry bailout packages worth Rs. 29 billion, the year 2018 was also a disappointing year for the textile sector as it failed to deliver the required results despite a mass devaluation of the Pak rupee (29%) throughout the calendar year [4]. Frederick and Daly [1] grouped the

Pakistan textile industry, the least coherent with successful quality practices in its supply chain in South Asia. The main hurdles in the expansion of the textile sector are lack of product diversification, imperfect competition, market diversification, financial crises, import duties, and some other factors like poor customer retention policy, whereas marketing literature frequently addressed the positive association of these factors, especially service quality with customer satisfaction and loyalty [6–9]. Kanat and colleagues blamed Pakistan textile and clothing industry that buyers are getting low service quality responses, thus reducing their confidence to do business with their existing suppliers [10]. They further argue that textile customers are conveniently switching from one supplier to another. This frequent customer switching rate signaled an alarming issue that instigated to initiate this study too. The model was developed to explain the relationships between switching costs, product returns, and service quality, arguing that customer satisfaction influences the corporate customers' loyalty with their present suppliers. It hypothesizes that a dynamic returns management system, improved service quality, and overcoming switching costs could advantage the supplier for gaining customer satisfaction that eventually generates the customer repurchase intentions [11]. However, there is little evidence about how foreseen concepts relate to the textile sector. This study framework sought to evaluate and particularly integrate these interactive antecedents of customer satisfaction to assess their loyalty (repurchase intent) with the present supplier within the B2B textile industry context.

In the textile value chain, customers, suppliers, and supply chain members develop relationships that mutually benefit all the stakeholders of the industry. This value chain comprises several activities that are essential to transform a product from conception to consumption through various other intermediary manufacturing stages (including physical transformation and other manufacturer services) [12]. In the textile industry, both buyers and sellers are, therefore, involved in value cocreation practices for economic exchange [13]. The textile firms want to keep this mutually beneficial process continue for the long term through superior customer services because the cost for appealing a new customer is multiple times higher than that required to sustain an existing one in the B2B transactional scenario [14]. So, customer loyalty with the current supplier firm has, therefore, been viewed as a significant factor for the firm's performance and the competitive advantage [15].

From the last few decades, the manufacturing industry has transformed itself by offering complete solutions or adding services to its actual product [16]. The provision of better service quality is one of the most essential factors for the satisfaction of the customers, and eventually, it increases their loyalty and strengthens repurchase intentions [17]. The findings of Chen also concluded that the provision of better services to the customers increased their satisfaction level and then it led toward the repurchase intentions [18]. The literature also recommended that strategically addressing product returns could also result in improved customer loyalty and increased customer satisfaction [19, 20]. A substantial returns management system could also lead to competitive advantages and superior performance [20, 21]. It creates an opportunity for the firms to manage positive long-term relationships with their clients, enhance customer satisfaction, and make a positive impact on their repurchase intention [22].

Theoretically, Pakistan's textile industry is increasing its production and export output because of its conventional goods, due to its intrinsic quality. However, major investment in machines and new technology is necessary for sustaining its situation, growing its portion, and expanding hooked on high-quality goods. Workforce preparation, performance improvement, research and development, product diversification, and branding are the immediate areas of focus.

Along with an effective returns management system and superior service quality, B2B markets use a traditional switching cost or switching barrier strategy to discourage customers to switch to another supplier [23]. The supply chain management literature has consistently shown that switching cost is an influential mechanism that discourages customers to switch and forces them to repurchase from the existing supplier [23, 24]. Most organizations attempt to form switching barriers that eventually increase the switching costs for their customers. The literature has found the positive impact of switching costs on repurchase intentions and that the customers stay connected with the present supplier if the perceived switching cost is higher for them [23].

Pakistan's textile industry's economic contribution is shown in Table 1, which indicates that the textile business funded 54% to exports, 46% to built-up production, 38% to employment, 8.5% to GDP in total GDP, 7.0 billion USD to investment, and 5.0% to market capitalization (listed companies). During the 5th Five-Year Plan, the synthetic filament yarn industry has identified impulses as demand increases, and therefore, imports have been increased as well as investments by the private sector in improving the market conditions (Pakistan Economic Survey, 2018–19). The textile sector is the second largest segment to contribute in GDP (gross domestic product) after agriculture in Pakistan.

The research questions are as follows:

RQ1: Do the product returns influence the repurchase intention of B2B customers in the textile industry?

RQ2: Do the switching costs create an exit barrier for the B2B textile industry customers to switch to the other supplier?

RQ3: Does service quality enhance customer satisfaction and push B2B textile industry customers to repurchase from the present supplier?

RQ4: Does customer satisfaction mediate the relationship between the repurchase intention of B2B textile industry customers and the antecedents?

2. Literature Review

2.1. Switching Costs. In business literature, switching costs are also known as switching barriers. These costs refer to a customer's costs of time, money, and efforts associated with customer switching from one supplier to another, rather than continuing to have business with the current supplier

Table 1: Sociodemographic characteristics of the participants ($n = 325$).

Variable	Categories	Frequency	Percentage
Current position/designation	General manager	27	8.3
	Manager	87	26.8
	Deputy/asst. manager	171	52.6
	Other staff	40	12.3
Department	Management	56	17.2
	Procurement/sourcing	127	39.1
	Commercial	88	27.1
	Others	54	16.6
No. of years on current position	1–5 years	133	40.9
	6–10 years	127	39.1
	11–15 years	46	14.2
	Above 15 years	19	5.8
No. of years in industry	1–5 years	62	19.0
	6–10 years	111	34.0
	11–15 years	107	33.0
	Above 15 years	45	14.0
Education level	Intermediate/diploma	21	6.5
	Graduate	121	37.5
	Masters/post graduate	152	46.8
	Others	31	9.5
Gender	Male	296	91.0
	Female	29	9.0
Marital status	Single	115	35.4
	Married	208	64.0
	Widowed	0	0.0
	Divorced	02	0.6
Age	25–30 years	67	20.5
	31–40 years	148	45.5
	41–50 years	94	29.0
	51 and above	16	5.0
Length of relationship with the main supplier	1–5 years	156	48.0
	6–10 years	133	40.9
	10–15 years	29	8.9
	Above 15 years	07	2.2
Does the main supplier has exclusive status (contract)	Yes	40	12.3
	No	285	87.7

[15]. It has also been shown in the literature that even when the customer is not satisfied with the present supplier, the customer retains and maintains a long-term relationship with the same company due to perceived switching costs, which is associated with breaching of the relationship with the present supplier [25]. High switching costs force customers to pursue even if the relationship with the current supplier is not satisfactory; it exists due to the customer assessment of additional costs to cease the current relationship [26, 27]. Bozzo revealed that corporate customers often repurchase from the present supplier despite not having any positive association and good links with the supplier due to perceived high switching costs [28]. Switching costs are not only considered as the loss in monetary terms, while these also have an impact on the psychological behaviour of the customer because it involves lots of effort looking for a new supplier for the same kind of products or services [29–31]. In a B2B setting, purchase decisions are not that much simple as these are in a B2C setting because the switching costs factor is highly complex and decisive in the B2B transactional scenario [32]. The switching process is quite painful, and the switching costs factor is substantial, even displeased customers want to keep going business relationships with the current product or service suppliers and retain from suspending the connection [27]. An enormous effect of switching costs on customer retention can be witnessed in commerce literature, where it is seen as useful for the firms for utilizing it as an instrument associated with marketing self-protecting strategy for maintaining and fortifying its connections with the present customer [33].

Generally in the satisfaction-loyalty relationship, switching costs are assumed to be a good moderator [23, 34] but substantial industrial marketing literature has taken the switching costs as a criterion variable that independently influences customer satisfaction and loyalty (e.g., [9, 23]). So, the study took switching costs as an independent variable to define the hypothesized relationship that switching costs affect the B2B customer satisfaction and the repurchase intent.

The study hypothesizes the following:

Product returns: for the corporate sector, product returns management involves substantial operational challenges and

enormous costs for the suppliers. Suppliers usually recall the products for inventory repositioning, or these recalls are initiated due to quality or security concerns that could be voluntary or instructed by government agencies [19]. Product returns management is more than just a cost [35], but generally, it is not so, as suppliers often recall their products to reposition the inventory. This repositioning of inventory and strategically managing returns could result in superior customer services, higher levels of customer satisfaction, and improved customer loyalty [19, 20]. The performance of product returns management associates with how effectively and efficiently an organization manages and handles customer issues and returns. In managing product returns, product vendors are provided with a chance to sought out customer objections and ensure the transition from unsatisfied to satisfied customers who were displeased due to initial supply experiences. Previous literature also emphasized that returns management could also lead to competitive advantages and superior performance [20, 21].

There exists a considerable body of commerce literature that witnessed that an effective product returns management can improve customer satisfaction, reduce costs, enhance profitability [20], shape favourable customer reviews, and make positive referrals [36], and thus, it leads to continuing customer relationship [37]. There are several reasons for the customers that they may return the product to the supplier. The reasons include product discontinuation, product replacement, in-transit damage, product off-season, expired date code, product defects, and high inventory of supplier [38]. Poorly managed product returns result in a significant negative effect on the firm's performance and customers' repurchase intentions. A bad product return experience leads toward switching behaviour and has a negative effect on customers' repurchase intentions [39].

Autry argues that a dynamic product returns management process provides an opportunity to form a positive effect on customers' repurchase intentions who earlier suffered service failures [40]. Although product returns management includes momentous operative and enormous costs, but simultaneously, it is an opening for managing and maintaining long-term relations with the customers and has a significant positive impact on their repurchase intentions [22]. Hence, product returns could also be viewed as customer satisfaction and service recovery opportunity. For a better understanding of the factors affecting repurchase intent, this study examined the role of product returns on B2B textile customer satisfaction and thus on their repurchase intent. Contrarily, Russo's study on the Italian audiology industry found no effect of managing product returns in the B2B context. He revealed that a high or low level of customer satisfaction does not trigger a positive effect on the repurchase intention of B2B customers.

2.2. Service Quality. Service quality is the provision of service that meets or surpasses the expectations of customers [41]. One has to compare the acquired service of the customer with perceived service to examine customer satisfaction with service quality [7]. The service quality-satisfaction-loyalty chain has been widely discussed in past literature [7, 9, 42]. Studies on consumer behaviour have frequently addressed the service quality issue, and the service quality has been proven to positively associate with customer satisfaction and loyalty [43].

Over the past decades, the manufacturing industry has experienced a shift by offering complete solutions or adding services to their actual product that is termed as "servitization" by [16]. The idea of providing products along with necessary services has created great value for the customers as compared to when these are acquired separately, and customer loyalty and satisfaction have also been increased [44]. This relationship between service quality and repurchasing intentions was frequently examined in the past (e.g., [45–47] and thus proven to positively associate with customer satisfaction and loyalty.

Sattari and colleagues state that the firm that needs to maintain a competitive edge over its rivals should come up with superior services to its customers. To survive and compete in the long run, the firm is required to focus on investing a substantial budget and time for the provision of improved quality services [48]. The provision of quality services is one of the most essential factors for customer satisfaction, and thus, it leads to customer loyalty and repurchase intentions of B2B customers [17].

The study hypothesizes the following:

Customer satisfaction: customer satisfaction is "the degree to which a customer perceives that an individual, firm or organization has effectively provided a product or service that meets the customer needs in the context in which the customer is aware of and/or using the product or service" [49], p. 79). Satisfaction is a socially fabricated response to the customer-product-supplier relationship as it is not coherent to an individual or the product. Customer satisfaction is a growing managerial interest of business managers as a means of evaluating service quality. A higher customer satisfaction score is extensively believed to be a fair indicator of a firm's future profit. Satisfaction can be widely characterized as a postpurchase assessment of product quality compared to the prepurchase expectation [50]. It is a cumulative assessment of all kinds of interactions of costumer while interacting with the firm. Therefore, customers' experiences of all kinds of supplier offerings have a significant effect on customer satisfaction. Researchers call customer satisfaction an emotional response that a customer shows right for the first time when the transaction is done. Once a customer gets satisfied with the service of the supplier, he/she is more likely to be returned for another transaction for the same or different product [51]. In B2B markets, the major differences among end-consumers arise when the company policymakers evaluate the features of a product/service of their industrial client. They consider and assess the satisfaction of the various stakeholders of the buying-house who are currently interacting with the industrial supplier [52]. Although these stakeholders are extensively guided by the firm's objectives, still they assess the product or service performance as per their self-reference standards due to their intrinsic motivations and objectives.

Several past studies that discussed a behavioural course of action found B2B customer satisfaction to be significantly and positively associated with repurchase intentions [45, 46, 51, 53, 54]. These studies illustrated that customer satisfaction plays a mediating role between switching costs, service quality, and customer loyalty [39, 55–61] and also have a consensus that customer satisfaction can explain the relationship between repurchase intentions (criterion variable) and their antecedents (predictor variables).

The study hypothesizes the following:

Repurchase intention: the repurchase intention is defined as customers' behavioural actions to purchase a similar product and available services more than one time from the same supplier [62]. In marketing, customer retention is the most important variable that commonly terms "repurchase" [43, 63–66]. In customer retention management, retaining an existing customer is quite cheaper than attracting a new one for the same product as it costs five times higher [14]. When a customer anticipates repurchasing from the present supplier and the intention to keep these relationships continues for the foreseeable future, it shows the loyalty of the customer to the supplier [67]. The loyalty with the present supplier or retailer could also be considered when a customer recommends the same product or service to the other corporate counterparts.

2.3. Theoretical Framework. For understanding the research hypotheses, Figure 1 presents the proposed theoretical model and comprehended the idea of the study that elaborated in diagrammatical form.

3. Research Methodology

3.1. Research Design. The study has chosen the raw material procurement employees of textile firms as the study population that is located in the main textile manufacturing regions, i.e., Karachi, Lahore, Faisalabad, and Multan. The survey questionnaire draft comprises 33 questions and was used to collect data from the target population. A data screening procedure was applied to the collected data. Missing data, outliers, normality, and inter-item correlations were checked for every single item. Adequacy of covariances and Cronbach's alpha was measured for all five constructs. Moreover, confirmatory factor analysis was conducted for each study construct.

After the data screening process, the conceptual model was partially interpreted through the AMOS model, containing a measurement part (confirmatory factor analysis, CFA) and a structural equation part (structural equational modeling, SEM). SEM technique was applied to simultaneously observe a chain of unified dependent association among a set of constructs and for accounting the measurement error at the same time (refer to [68]). The technique suggested how the observed variables relate to the latent constructs and the unified dependent association among these sets of variables.

3.2. Measures. The literature drawn measuring scales were adopted from the past literature that has already been tested for measuring a similar concept. These applied multi-item instruments were found to be valid and reliable for employing and measuring the behavioural intention of industrial customers. Table 2 illustrates the number of items adopted from existing scales in the five-point Likert scale.

3.3. Procedure/Data Collection. This study was designed to populate statistics from a considerable sample size ($n = 325$) to gain statistically valid results. The target population was comprised of the procurement-related employees or intermediaries of the APTMA (All Pakistan Textile Mills Association) member firms to ascertain the valuable inputs of B2B respondents. The respondents (unit of analysis) were designated to purchase raw materials for their companies, which are listed in APTMA (i.e., spinning and weaving mills, dyeing and finishing mills, garment stitching, apparel factories, home textile units, trading companies, and buying houses). The specific population characteristics led the study to follow the specific criterion for sample selection. So, the purposive sampling technique was applied for extracting the sample since the target population has particular traits in terms of their industrial affiliation, responsibilities, and nature of the job.

The questionnaire draft, consisting of closed-ended questions, was designed with the measures of relevant constructs. This questionnaire draft had two main sections. One of the sections comprised respondents' demographics, duration of association with the supplier firms, and a screening question (i.e., Did you ever purchase or involve in the purchase process of textile raw material from your main supplier?), whereas the other section consists of construct measurements where agreement or disagreement was judged on a five-point Likert scales (1 = *strongly disagree* and 5 = *strongly agree*). The questionnaire draft, consisting of closed-ended questions, was made very neutral by keeping the researcher's personal opinions aside to avoid response bias. Questions were easy to understand and clear as there was no item ambiguity in question statements to ensure nonresponse bias. There was no item similarity, or proximity vis-à-vis length of the questionnaire was also moderate. The respondents were conveyed to fill up questionnaires and rate their raw material suppliers in terms of their service quality, returns management system, and switching barriers. A sum of 450 questionnaires was couriered to the respondents of various textile firms, and 375 out of them were received back, producing 83.3% response rate. After data screening, the rest of (n) 325 valid responses were analyzed via statistical package for the social sciences (SPSS) (version 21). A total of 50 questionnaires were excluded from the data due to having some issues. These issues were related to either nonresponse bias or the presence of acquiescence bias where the respondents tended to agree with all questions or they concurred with a particular position. The rest questionnaires were adequate for further analysis.

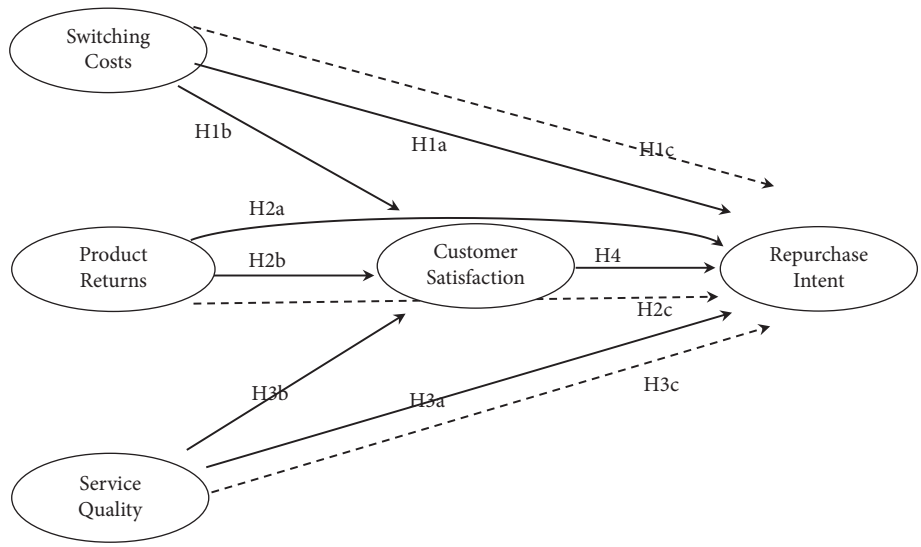


Figure 1: Diagrammatic representation of the relationships among study variables (theoretical model).

Table 2: Measurement of study variables.

Variable	No. of items	Measurement
Switching costs	5	[23]
Product returns	3	[22]
Service quality	15	[69]
Customer satisfaction	6	[70]
Repurchase intent	4	[71]

3.4. Participants. The survey carries the responses of (n) 325 B2B textile materials procurement officials. The total number of respondents included 296 men and 29 women. All respondents were asked to choose from a list of answer choices whether they agree or disagree with a question statement.

4. Results

4.1. Descriptive Statistics. In Table 3, mean values (3.24–3.83) are clearly showing that the values are evidently spread around the central value and ideally describe the central tendency of data. Similarly, low standard deviation values (0.41–0.68) are also showing that the values are not much deviating from the central point and the data points tend to be close to the mean of the data set. It shows fairly close agreement among respondents and a little variance about the answers to the questions. The mean and the standard deviation values confirmed the significance of the data that there were very few chances of error.

4.2. Multivariate Normality. Skewness and Kurtosis values were observed within the acceptable range of normality, and the test established the multivariate normality of the data. The values (see Table 4) showed that the distribution is negatively skewed and the extreme values or outliers are not present in the distribution. Simultaneously, Kurtosis values describe that the sample is slightly platykurtic. Its peak is just a bit shallower than the peak of a perfect distribution. So, it is confirmed that the existing data were normally distributed.

Kaiser–Meyer–Olkin (KMO) and Bartlett's test confirmed the adequacy and suitability of the data. Taken together, these tests satisfied the minimum standards that are required to accomplish before conducting CFA on data. The KMO values of all five study constructs are greater than the recommended range (>0.6) (refer to [73]) and show the adequacy of the percentage of variance in data. Thus confirmed, sampling was adequate and the data were suitable for conducting confirmatory factor analysis (CFA) on it.

4.3. Validation of the Measurement Model. Convergent and discriminant validities of study instruments were assessed before undertaking the confirmatory factor analysis that how thoroughly these constructs measure the intended concepts. As per the criterion set by Fornell and Larcker [74], the items with poor factor loadings were removed from the measurement model to achieve convergent validity. The removal of the seven items (i.e., SC-money, SC-effort, SQ-accurate, SQ-damage, SQ-response, SQ-safe, and SQ-specific) had lower regression weights than the criterion (<.4) set by Avkiran and Ringle [75]. The remaining factor loadings significantly exceeded the threshold value (≥0.70). The removal of these items strengthened the values of different model fit indices, for instance, GFI, AGFI, CFI, TLI, and RMSEA for the achievement of their respective threshold values. These removals also enhanced the AVE up to the acceptable level (i.e., ≥0.50) with the values ranging between 0.50

Table 3: Descriptive statistics of the study.

	N Statistics	Minimum statistics	Maximum statistics	Mean statistics	Std. deviation statistics	Variance statistics
Service quality	325	2.40	4.93	3.816	0.411	0.169
Switching costs	325	1.00	4.80	3.244	0.659	0.435
Product returns	325	1.67	5.00	3.805	0.678	0.460
Customer satisfaction	325	1.67	5.00	3.829	0.503	0.253
Repurchase intention	325	2.00	5.00	3.706	0.627	0.393
Valid N (listwise)	325					

Table 4: Frequencies of the study constructs.

		Service quality	Switching costs	Product returns	Customer satisfaction	Repurchase intent
N	Valid	325	325	325	325	325
	Missing	0	0	0	0	0
Skewness		−0.662	−0.272	−0.958	−1.011	−0.277
Std. Error of skewness		0.135	0.135	0.135	0.135	0.135
Kurtosis		1.393	−0.192	0.568	2.307	0.008
Std. Error of kurtosis		0.233	0.270	0.270	0.270	0.270
P-value		0.001	0.002	0.319	0.001	

Note. Skewness value >+1 or <-1 (balance distribution) – Kurtosis value <1 (flat distribution) [72].

and 0.52, as shown in Table 5. So, all the factors satisfied the discriminant validity, and these were precise in nature (truly measure the characteristics being represented by the variables).

To ensure the uniformity and stability of the measures, the internal and the composite reliabilities were measured. Cronbach's alpha test confirmed the internal consistency and reliability of the concepts with the values (ranging 0.72–0.85) and satisfied the criterion set by Wollack et al. [76]. The composite reliability values (ranging 0.75–0.90) were also above the proposed level ≥0.70 (refer to [77]). Thus, convergent validity was also confirmed by the empirical data.

Multicollinearity was assessed by measuring tolerance and variance inflation factor (VIF), where the VIF values exist around 1. It postulated that predictor variables are either not correlated or are moderately correlated with other variables. The VIF values of all study constructs are less than the recommended threshold value (<3) (refer to [78, 79]) and confirmed that there exists no multicollinearity issue. Also, the model is free of common method bias (CMB) and it is adequate for further statistical analysis. Common method bias (Table 6) is also checked through Harman's single factor method as the total variance extracted by the first common factor that is an account for the 27.6% covariance that is far below the criterion percentage (50%) (criterion set by the developer), whereas CMB statistics (Table 7) also confirmed the absence of bias in data using the common latent factor method.

The correlation matrix (Table 8) was drawn to observe the interconstruct correlation. The matrix indicated that these variables are not mutually correlating with each other. Bivariate correlational values are below the recommended threshold value (<0.7) (refer to [80]). Service quality tends to have a strong positive correlation with customer satisfaction and thus moderate positive correlation with the repurchase intention ($r = 0.66$ and $r = 0.52$, respectively). Similarly, customer satisfaction also has a moderate correlation with repurchase intention ($r = 0.54$). Contrarily, product return has a weak correlation with customer satisfaction and repurchase intention ($r = 0.22$ and $r = 0.25$, respectively) as mentioned in Table 9. Similarly, service quality is also demonstrating a week correlation with switching costs ($r = 0.27$).

4.4. Model's Measurement. The study models and their indicators were specified and labeled as the error terms. The model was also identified as it had enough pieces of information in an equation to produce unique estimates of unknown parameters. In the third stage, different model fit indices, for instance, GFI, AGFI, CFI, TLI, and RMSEA, were estimated. The degree of exactitude of the model was measured through chi-square (x^2) value, as the value is sensitive to a big sample size, i.e., ($n > 200$) (refer to [81]). Finally, the model was respecified till acquiring the value of chi-square/df is < 3 by covarying error terms and by enforcing some path coefficient constraints. The respecified model achieved the acceptable values of fit indices that were observing poor fit otherwise. Therefore, some assumptions, for instance, no equality constraints, were enforced on the factor loadings and were also considered for these indicators.

The distribution of latent constructs is significantly different from each other such as CMIN/df < 3. Likewise, the values of GFI, AGFI, CFI, and TLI are above 0.9, which shows the overall good fit of the model. For population, root mean square error of approximation (RMSEA) value (≤0.07) (refer to [82]) indicates a good fit of the model in relation to the degrees of freedom (df).

The re-specified model fit indices confirmed that the hypothesized model explained a good fit to the data where ($n = 325$), $p < 0.001$, GFI = 0.918, AGFI = 0.905, CFI = 0.917, TLI = 0.976, and RMSEA = 0.054, available in Table 10. With overall fit indices, the study model is confidently supposed to be a fairly good fit to the present data due to significant and meaningful path indications in the model. The items of the

Table 5: KMO and Bartlett's test.

KMO and Bartlett's test	Switching costs	Product returns	Service quality	Customer satisfaction	Repurchase intent	Overall
Kaiser–Meyer–Olkin measure of sampling adequacy	0.710	0.678	0.831	0.853	0.736	0.877
Bartlett's test of sphericity approx. chi-square	346.0	235.0	1533.4	744.9	364.2	4616.5
Df	10	3	105	15	6	528
Sig.	0.001	0.001	0.001	0.001	0.001	0.001

Note. Value (KMO >0.5 or ideally >0.7) for adequacy of percentage of variance [73].

Table 6: Common method bias (CMB) test through common latent factor method.

Standardized regression weights (with common latent factor)				Standardized regression weights (without common latent factor)				Difference
Items			Estimate	Items			Estimate	
SC_money	<---	SC	0.456	SC_money	<---	SC	0.465	**0.009**
SC_effort	<---	SC	0.585	SC_effort	<---	SC	0.555	**−0.03**
SC_time	<---	SC	0.649	SC_time	<---	SC	0.691	**0.042**
SC_technical	<---	SC	0.531	SC_technical	<---	SC	0.632	**0.101**
SC_uncertain	<---	SC	0.516	SC_uncertain	<---	SC	0.597	**0.081**
PR_guarantee	<---	PR	0.702	PR_guarantee	<---	PR	0.83	**0.128**
PR_care	<---	PR	0.618	PR_care	<---	PR	0.679	**0.061**
PR_convenient	<---	PR	0.593	PR_convenient	<---	PR	0.623	**0.03**
CS_services	<---	CS	0.445	CS_services	<---	CS	0.596	**0.151**
CS_pleased	<---	CS	0.535	CS_pleased	<---	CS	0.712	**0.177**
CS_experience	<---	CS	0.528	CS_experience	<---	CS	0.705	**0.177**
CS_overall	<---	CS	0.442	CS_overall	<---	CS	0.628	**0.186**
CS_choose	<---	CS	0.224	CS_choose	<---	CS	0.425	**0.201**
CS_easier	<---	CS	0.466	CS_easier	<---	CS	0.649	**0.183**
RI_next	<---	RI	0.634	RI_next	<---	RI	0.793	**0.159**
RI_relation	<---	RI	0.391	RI_relation	<---	RI	0.567	**0.176**
RI_percent	<---	RI	0.454	RI_percent	<---	RI	0.621	**0.167**
RI_pincrease	<---	RI	0.583	RI_pincrease	<---	RI	0.628	**0.045**
SQ_accurate	<---	SQ	0.148	SQ_accurate	<---	SQ	0.344	**0.196**
SQ_comply	<---	SQ	0.106	SQ_comply	<---	SQ	0.293	**0.187**
SQ_ontime	<---	SQ	0.24	SQ_ontime	<---	SQ	0.422	**0.182**
SQ_damage	<---	SQ	0.002	SQ_damage	<---	SQ	0.201	**0.199**
SQ_prodstatus	<---	SQ	−0.092	SQ_prodstatus	<---	SQ	0.41	**0.502**
SQ_response	<---	SQ	−0.144	SQ_response	<---	SQ	0.268	**0.412**
SQ_busy	<---	SQ	−0.085	SQ_busy	<---	SQ	0.264	**0.349**
SQ_enthusi	<---	SQ	0.055	SQ_enthusi	<---	SQ	0.196	**0.141**
SQ_expertise	<---	SQ	0.046	SQ_expertise	<---	SQ	0.211	**0.165**
SQ_safe	<---	SQ	0.27	SQ_safe	<---	SQ	0.483	**0.213**
SQ_person	<---	SQ	0.546	SQ_person	<---	SQ	0.669	**0.123**
SQ_imp	<---	SQ	0.844	SQ_imp	<---	SQ	0.63	**−0.214**
SQ_specific	<---	SQ	0.257	SQ_specific	<---	SQ	0.366	**0.109**
SQ_needs	<---	SQ	0.217	SQ_needs	<---	SQ	0.435	**0.218**
SQ_status	<---	SQ	0.185	SQ_status	<---	SQ	0.359	**0.174**

model are significantly related to their constructs, so all hypothesized relationships are supported between constructs and their indicators.

4.5. Structural Equation Modeling (SEM)

4.5.1. Structural model assessment. Standardized path coefficients with a bootstrap value of 2,000 resamples were measured to test the study hypotheses. A large number of replicates were required to achieve the same level of the estimate due to longer alignment. In this study, the mean statistics of switching costs were very close to the central (mean) value, so test statistics can fall into one of the critical regions either. So, two-tailed significance with a 95% confidence interval was determined due to both types of perceived positive or negative relationships. The data results confirmed the projected path structure (direct and indirect) authenticity and adequacy. Table 11 summarizes the test results of study hypotheses and explains the direct, indirect, and total effects along with the path coefficients and p-values of the variables.

Table 7: Common method bias (CMB) through Harman's single factor method.

Component	Total variance explained					
	Initial eigenvalues			Extraction sums of squared loadings		
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	9.112	27.613	27.613	9.112	27.613	27.613
2	2.569	7.786	35.399			
3	2.417	7.326	42.725			
4	1.442	4.370	47.095			
5	1.381	4.184	51.279			
6	1.352	4.096	55.376			
7	1.280	3.879	59.254			
8	1.082	3.279	62.533			
9	0.955	2.894	65.427			
10	0.876	2.656	68.083			
11	0.849	2.574	70.657			
12	0.825	2.501	73.158			
13	0.729	2.209	75.367			
14	0.697	2.113	77.480			
15	0.647	1.961	79.442			
16	0.586	1.776	81.218			
17	0.570	1.727	82.945			
18	0.538	1.631	84.576			
19	0.519	1.573	86.149			
20	0.480	1.454	87.603			
21	0.455	1.380	88.982			
22	0.420	1.272	90.255			
23	0.387	1.172	91.427			
24	0.370	1.120	92.547			
25	0.344	1.043	93.590			
26	0.312	0.946	94.535			
27	0.306	0.927	95.462			
28	0.292	0.884	96.345			
29	0.278	0.843	97.188			
30	0.254	0.771	97.959			
31	0.244	0.738	98.697			
32	0.232	0.702	99.399			
33	0.198	0.601	100.000			

Extraction method: principal component analysis.

Table 8: Measurement model result summary.

	Factor loadings		Reliability and validity			Common method bias (collinearity statistics)	
Constructs	Item	Main loading	AVE	Composite reliability (CR)	Cronbach's alpha	Tolerance	VIF
Switching costs	SC-time	0.72	0.50	0.75	0.72	0.74	1.19
	SC-technical	0.70					
	SC-uncertain	0.71					
Product returns	PR-guarantee	0.69	0.51	0.76	0.75	0.70	1.25
	PR-care	0.72					
	PR-convenient	0.74					
Service quality	SQ-needs	0.71	0.51	0.90	0.85	0.51	1.96
	SQ-status	0.70					
	SQ-comply	0.75					
	SQ-ontime	0.69					
	SQ-prodstatus	0.77					
	SQ-enthusi	0.69					
	SQ-expertise	0.72					
	SQ-person	0.72					

Table 8: Continued.

	Factor loadings		Reliability and validity			Common method bias (collinearity statistics)	
Constructs	Item	Main loading	AVE	Composite reliability (CR)	Cronbach's alpha	Tolerance	VIF
Customer satisfaction	SQ-busy	0.71	0.52	0.87	0.85	0.55	1.83
	SQ-imp	0.74					
	CS-services	0.76					
	CS-pleased	0.69					
	CS-experience	0.69					
	CS-overall	0.71					
	CS-choose	0.76					
	CS-easier	0.74					
Repurchase intent	RI-next	0.75	0.51	0.81	0.76		
	RI-relation	0.72					
	RI-percent	0.79					
	RI-pincrease	0.69					

Note. Dependent variable: repurchase intention; AVE: average variance extracted; VIF: variance inflation factor.

Table 9: Pearson's correlation matrix.

Model	Service quality	Switching costs	Product returns	Customer satisfaction	Repurchase intent
Service quality	1				
Switching costs	0.273**	1			
Product returns	0.382**	0.311**	1		
Customer satisfaction	0.655**	0.313**	0.217**	1	
Repurchase intent	0.515**	0.444**	0.254**	0.542**	1

Note. **Correlation is significant at the 0.01 level (2-tailed).

Table 10: Model fit summary (original and revised model indices).

Model	Items	CMIN/*df*	*df*	GFI	AGFI	CFI	TLI	RMSEA
Original model	33	2.324	461	0.836	0.800	0.857	0.836	0.064
Revised model	26	2.190	255	0.918	0.905	0.917	0.976	0.054

Note. GFI: goodness of fit index; AGFI: adjusted goodness of fit index; CFI: comparative fit index; TLI: Tucker–Lewis index; RMSEA: root mean square error of approximation.

The structural model assessment revealed that except for product returns, the other two constructs, switching costs and service quality, have significant direct positive effects on customer satisfaction and thus direct and indirect positive effects on repurchase intentions of the B2B textile industry customers. The results showed that customer satisfaction is a significant determinant of repurchase intention and it effectively explains the relationship between the predictors and the outcome variable of the study.

Hypothesis H1b postulated that switching costs are directly positively associated with customer satisfaction and thus directly and indirectly positively associated (H1a and H1c) with repurchase intent mentioned in Table 12. SEM results exhibited proper support for the hypotheses H1a ($\beta = 0.224$, $p < 0.05$), H1b ($\beta = 0.167$, $p < 0.05$), and H1c ($\beta = 0.111$, $p < 0.05$) and proved that switching costs are the salient determinant of customer satisfaction and repurchase intention.

Hypotheses H2a, H2b, and H2c were assumed that product return has direct positive associations with customer satisfaction and repurchase intention and also has indirect positive relationships with repurchase intention. The analysis results found no support for these three hypotheses and exhibited no relationships of product returns with customer satisfaction and repurchase intention of B2B customers. SEM results for these hypotheses are H2a (β = -0.038, $p < 0.05$), H2b ($\beta = -0.100$, $p < 0.05$), and H2c ($\beta = -0.066$, $p < 0.05$).

Hypotheses H3a and H3c postulated that direct and indirect relationships exist between service quality and repurchase intent. Results endorsed these significant positive relationships with the values ($\beta = 0.618$, $p < 0.05$, and $\beta = 0.426$, $p < 0.05$), respectively, and thus established that service quality is an effective indicator of generating repurchase intention of B2B textile industry customers. Similarly, hypothesis H3b proposed a positive relationship between service quality and customer satisfaction. Results also endorsed this significantly positive relationship with the value ($\beta = 0.646$, $p < 0.05$).

Finally, the study hypothesized a relationship (H4) between the mediator (customer satisfaction) and the DV (repurchase intention) of the study. Results also

Table 11: Hypothesis testing result summary.

Hypothesis	Relationships	Path coefficients	p-value	Confidence interval	Results
H1a	SWTCST⟶REPINT	0.224**	0.002	0.152–0.301	Supported
H1b	SWTCST⟶CUSTSATF	0.167**	0.003	0.075–0.259	Supported
H1c	SWTCST⟶CUSTSATF⟶REPINT	0.111**	0.002	0.051–0.176	Supported
H2a	PRODRTN⟶REPINT	−0.038	0.319	−0.031–0.127	Not supported
H2b	PRODRTN⟶CUSTSATF	−0.100	0.059	−0.218–0.004	Not supported
H2c	PRODRTN⟶CUSTSATF⟶REPINT	−0.066	0.055	−0.150–0.001	Not supported
H3a	SERVQLTY⟶REPINT	0.618**	0.001	0.551–0.711	Supported
H3b	SERVQLTY⟶CUSTSATF	0.646**	0.002	0.543–0.724	Supported
H3c	SERVQLTY⟶CUSTSATF⟶REPINT	0.426**	0.002	0.337–0.521	Supported
H4	CUSTSATF⟶REPINT	0.660**	0.001	0.570–0.752	Supported

Note. * mean significant at level $p < 0.05$; ** mean significant at level $p < 0.01$.SWTCST: switching costs; PRODRTN: product returns; SERVQLTY: service quality; CUSTSATF: customer satisfaction; REPINT: repurchase intention.

Table 12: Standardized direct/indirect effects with two-tailed significance.

	Standardized direct effects (group number 1: default model)			
	PRODRTN	SERQLTY	SWTCST	CUSTSATF
CUSTSATF	−0.100	0.646	0.167	0.000
REPINT	−0.038	0.618	0.224	0.660
	Standardized direct effects: lower bounds (BC) (group number 1: default model)			
	PRODRTN	SERQLTY	SWTCST	CUSTSATF
CUSTSATF	−0.218	0.543	0.075	0.000
REPINT	−0.031	−0.103	0.152	0.570
	Standardized direct effects: upper bounds (BC) (group number 1: default model)			
	PRODRTN	SERQLTY	SWTCST	CUSTSATF
CUSTSATF	0.004	0.724	0.259	0.000
REPINT	0.127	0.111	0.301	0.752
	Standardized direct effects: two-tailed significance (BC) (group number 1: default model)			
	PRODRTN	SERQLTY	SWTCST	CUSTSATF
CUSTSATF	0.059	0.002	0.003	0.0.0.
REPINT	0.319	0.001	0.002	0.001
	Standardized indirect effects (group number 1: default model)			
	PRODRTN	SERQLTY	SWTCST	CUSTSATF
CUSTSATF	0.000	0.000	0.000	0.000
REPINT	−0.066	0.426	0.111	0.000
	Standardized indirect effects: lower bounds (BC) (group number 1: default model)			
	PRODRTN	SERQLTY	SWTCST	CUSTSATF
CUSTSATF	0.000	0.000	0.000	0.000
REPINT	−0.150	0.337	0.051	0.000
	Standardized indirect effects: upper bounds (BC) (group number 1: default model)			
	PRODRTN	SERQLTY	SWTCST	CUSTSATF
CUSTSATF	0.000	0.000	0.000	0.000
REPINT	0.001	0.521	0.176	0.000
	Standardized indirect effects: two-tailed significance (BC) (group number 1: default model)			
	PRODRTN	SERQLTY	SWTCST	CUSTSATF
CUSTSATF	...	...	...	...
REPINT	0.055	0.002	0.002	...

endorsed this significant positive relationship as ($\beta = 0.660$, $p < 0.05$). It indicated that customer satisfaction is a strong indicator of forming repurchase intention of the B2B customers and mediates between predictors (i.e., switching costs and service quality) and the outcome variable (repurchase intent). Therefore, it does not mediate between product returns and repurchase intention mentioned in Table 1.

5. Discussion

The statistical results verified the strong influence of customer satisfaction on the repurchase intention of industrial customers. Customer satisfaction is a strong apparatus for shaping the repurchase intention of the B2B textile industry customers. These findings are consistent as these are in tandem with the various past studies with similar context

(e.g., [9, 23, 34, 83, 84]). Results confirmed that customer satisfaction effectively explains the relationship and function as a mediator between B2B repurchase intention and its antecedents (i.e., switching costs and service quality). The results confidently confirmed that satisfied customers are more likely to repurchase. Therefore, customer satisfaction did not explain any relationship between product returns and repurchase intent, since no significant direct or indirect relationship was found between these two constructs, specifically in the case of B2B behavioural course of action. So, there exists no partial mediation or indirect relationship between product returns and repurchase intention as it was initially assumed in hypothesis H2c. The results somehow supported all perceived study hypotheses except H2a, H2b, and H2c. The product returns did not establish any significant positive or negative relationship with customer satisfaction and thus with the DV (repurchase intent) through the mediator. It pays neither a positive nor a negative effect on B2B customer satisfaction and thus on their repurchase intentions.

The data analysis verified that switching costs and service quality are significantly important determinants that positively influence customer satisfaction and repurchase intention that eventually force the B2B textile industry customer to repurchase from the present supplier. Hypotheses H1a, H1b, and H1c were effectively supported by the study results, since switching cost has a significant direct and indirect positive association with DV (repurchase intent) and direct association with the mediator (customer satisfaction). Results endorsed that switching cost is an extended and ideally important construct to be considered in the B2B transactional scenario of the textile industry. It is an essential determinant that builds positive repurchase intentions and influences B2B customers to be with their present suppliers. Customers become more comfortable with the present supplier as their switching costs to replace the present supplier increase.

The switching cost has a pivotal role in describing B2B customer repurchase intent, as it resulted in enhancing customer satisfaction and transforming positive repurchase intentions for having a long-term relationship with the present supplier. Mostly, past studies affirmed switching costs, a basic requirement, as a barrier to stop customers from switching (e.g., [85–89]). Simultaneously, the past literature also describes that the absence of switching barriers is the main reason for the convenient switching of the customer from one supplier to another, as the state of vulnerability of the supplier is quite high in the dynamic disposition of market space [90]. In the B2B transactional scenario, empirical evidence revealed that investing to raise switching costs resists customers to switch and these are beneficial for the firms in retaining their current clientele. This could be achieved through making changes in procedures, investing in new technology, revising payment terms, and training personnel to utilize advanced technology techniques to retain the customers if other suppliers try to step in.

Therefore, in the case of product returns, the findings are quite different from the other two antecedents of customer satisfaction and loyalty (i.e., switching costs and service quality). Product return does not increase customer satisfaction and also does not shape the intentions of B2B customers to repurchase from the same supplier even after having a good product return experience in past. More importantly, the most critical and commonly discussed weak point that is associated with the product return policy of the firm is specifically the weak management of customer complaints and product returns. This weak point is also noticed by Lambert and Enz. Their study on supply chain management revealed that organizations should utilize customer knowledge to manage returns [91]. Organizations must handle customer information or knowledge strategically to develop value for the return processes.

The customer repurchase intentions are extensively shaped by their satisfaction with the current service supplier, but a good product return experience does not form any added effect to enhance B2B customer satisfaction. A field experiment study by Petersen and Kumar also revealed that product return has a significant effect as it determines the future attitudes and behavior of customers—their willingness to refer the company to friends [92]. Therefore, other markets, for instance, B2C, might be effectively encouraged or influenced by the firm's effective product returns management system. It ascertains that once the B2B customers have experienced good or bad product returns from the supplier side, it does not have any added influence on their satisfaction and loyalty. Although this may be surprising since there is a mass discussion about the important role of product returns in supply chain literature, the past literature witnessed that effective product returns management can improve customer satisfaction, reduce costs, enhance profitability [20], shape favourable customer reviews, and make positive referrals [36], and thus, it leads to continuing customer relationship [37]. Askariazad and Babakhani [39] found that the bad product return experience of customers contributes toward dissatisfaction and has a significant negative impact on the repurchase intention of B2B customers, but this study found that neither good nor bad product return experience contributes toward customer satisfaction. There is inconsistency with the findings, given by the past literature on product returns, but the findings are consistent with Russo's study where he conducted research on the Italian audiology industry [84]. So, the hypotheses H2a, H2b, and H2c are not supported by the study results.

The delivery of superior quality services and high switching costs are both equally effective ways to enhance industrial customer loyalty [15, 32]. The relationship of service quality with customer satisfaction (mediator) and then with repurchase intention (DV) was found significant. Service quality has a significant positive direct effect (H3b) on the mediator and also significant positive direct and indirect effects (H3a and H3c) on DV. The link between service quality and customer satisfaction has shown the strong relationship that eventually forces B2B customers to be with and repurchase from the present supplier. The results showed that service quality is an indicator of building or transforming B2B customer loyalty for considering the current supplier for repurchasing the textile raw materials.

Service quality is a salient determinant of customer satisfaction and repurchase intention, and the respondents had overall a favourable evaluation of supplier quality services and they generally acknowledge it. The results showed that service quality is quite useful in upholding customer satisfaction that stimulates the repurchase intentions in the B2B transactional context. Thus, the textile firms should provide superior quality services to their customers to increase customer satisfaction and to stop other suppliers to step in. The organization, which needs to form and sustain a competitive edge over competitors, should deliver greater services to their organizational customers. These firms are required to focus on investing substantial efforts and time to provide superior quality services to their customers to be in the market and to survive in the long run. Otaigbe Otaigbe [93] also revealed in his study that business managers who consistently review their supply chain management system and processes to ensure service quality experience more satisfied and loyal customers and increased profitability [94]. Hence, the study results provided empirical evidence that service quality tends to be a fair indicator of customer satisfaction and the repurchase intention of B2B textile industry customers. A better service quality increases customer satisfaction and thus assures the retention of the current organizational clientele. The finding of this study is also consistent with that of previous studies that showed similar relationships between these variables [42, 95–97].

Finally, marketing literature widely focuses on customer satisfaction and verifies that firms extensively focus on their customers and their success entirely depends upon customer satisfaction. Baidya and Ghosh [98] assessed customer repeat purchase as the output of a firm's successful customer satisfaction strategy. The study results also showed a significant relationship of customer satisfaction with the B2B repurchase intent, as satisfied industrial customers showed higher behavioural loyalty toward supplier firms. These findings are consistent with past research studies on customer satisfaction [99–104].

Firms are required to be more focused on customer satisfaction than sales since the customer is the key entity for the organizational business, and customer satisfaction is a substantial measuring tool for corporate sustainability. The logic behind the relationship between customer satisfaction and its positive effect on customer retention is that an increase in customer satisfaction will boost customer retention and generate future revenues for the business [99, 105]. Customer satisfaction programs, greater customer switching barriers, and service quality practices can upgrade business processes.

So, a firm needs to develop firm and long-term relationships with its customers by constantly targeting and trying to retain them with distinctive attention. A satisfied customer becomes a loyal customer, provides positive word-of-mouth, and repeats future purchases that lead to increased revenue and that improve the stock prices of the business. For a firm, there is a need to pay more focus on customer satisfaction by all means, [98] as customer satisfaction boosts customer retention [99, 105]. The study results, therefore, confirmed, if an organization elevates customer satisfaction by taking necessary measures, for instance, it increases the service quality and enhances the switching costs to create an exit barrier, it succeeds to generate repurchase intentions of their existing customers.

5.1. Study Contributions. From a theoretical perspective, the study contributes to the literature by combining switching costs, product returns, and service quality on the satisfaction-retention framework. The study depicts a reviewed framework of B2B-specific variable combination in the context of the Pakistan textile industry. The study framework is unique, and it does add up a distinct concept by exploring the behavioural intention of a textile industry-specific population, which is comparatively a new population so far. The outcome of the study regarding product returns contradicts the broader consensus in supply chain literature. It proved insignificant as product return has not any added influence on B2B customer satisfaction and repurchase intent, no matter the customers experienced a vigorous product return even in the past. However, the results of this study are consistent with Russo's findings on the Italian audiology industry [84].

The study particularly contributes to commerce literature and work practices differently. Contrary to the past studies, which have taken either service quality or switching costs to predict customer satisfaction and loyalty, this study explicitly and simultaneously integrated both switching costs (repulsive force) and service quality (attractive force) in the study framework. It examines the outcome of these repulsive and attractive forces in the context of the Pakistan textile industry. The study framework consolidates the mutual dynamics of appealing (service quality), facilitating (product returns), and averting (switching costs) factors altogether and their effect on customer satisfaction and thus on customer retention (repurchase intent) in B2B transactional scenario.

5.2. Managerial Implications. In a practical sense, these study inferences will be valuable for the planning managers to assess the behavioural intention of B2B customers in developing or reviewing their marketing strategies for the Pakistan textile markets. The study recommended that the textile seller firms should design such a viable system that supports and appeals customers toward obtaining their satisfaction by forming exit barriers and improving the service quality. It is essential because mostly customers question the integrity and after-sale services of the supplier firms while performing a transaction. So, the firms must, therefore, convince their industrial customers to have trust in their service quality since it is found that service quality is a significant determinant of B2B customer satisfaction and their repurchase intent. The findings suggested that firms need to strengthen the exit barriers for their clientele to attract and motivate them more to repurchase form the present supplier.

The firms that are still targeting to project their businesses outside of their current portfolio may utilize this study as a guide for the transition from conventional to

progressive firms. These firms can focus even better on upholding the current clientele by exerting a stronger influence on customer satisfaction and their repurchase intention, which they could have missed otherwise.

6. Limitations and Future Research Directions

There are some limitations such as the study sample was limited to an explicit segment of industrial society (B2B textile material procurement officials) and of a specific geographical area (firms located in the main textile manufacturing regions, i.e., Karachi, Lahore, Faisalabad, and Multan) So, extending the findings to some other geographical area or service sectors may lead some different results.

Second, the study only addresses the behavioural intention of corporate buyers who are technically sound and well familiar with the quality and features of the product. So, the statistics may not appeal to B2C or virtual markets having different populations. Even, within the textile industry, a longitudinal study may demonstrate some different results due to changing market dynamics at two different points of time.

A future research pattern could be carried out by testing the study framework within the B2B service industry or virtual markets to enlarge the research agenda. Culture exerts an interdependent and complementary influence on the behaviours that drive demographic change [106]. So, cross-cultural impact or the specific demographical pattern that might base on a specific gender, designation, or job role can be taken to design the future pattern of the industrial population.

Data Availability

The data are not available in the main manuscript, and result analysis is provided through tables and figures. All data are associated with their findings.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was funded by the "Shaanxi Provincial Social Science Fund Project" (grant number: 2021R005) and The New Style Think Tank of Shaanxi Universities: Research Center for Auxiliary Chemistry and New Material Development, Shaanxi University of Science and Technology (grant number: ACNM-202202).

Supplementary Materials

The graphical representation of the abstract is provided. The graph visually represents the primary findings of the article. (*Supplementary Materials*)

References

[1] Frederick, & Daly. (2019). Pakistan in the Apparel Global Value Chain. Retrieved from Duke University Global Value Chains Center: https://gvcc.duke.edu/wp-content/uploads/Pakistan_Apparel_GVC_Report_2019.pdf.

[2] PBIT, *Textile Sector of Pakistan*, PBIT, Lahore, Pakistan, 2019.

[3] NNI, *Textile Industry in Pakistan an Open Example of Resistance Economy*, The Nation, New York, NY, USA, 2018, https://nation.com.pk/03-Jun-2018/textile-industry-in-pakistan-an-open-example-of-resistance-economy.

[4] A. Mansoor, *Textile Industry Continues to Disappoint with its Inefficiency in 2018*, The News, London, UK, 2018.

[5] M. Tanveer and S. Zafar, "The stagnant performance of textile industry in Pakistan," *European Journal of Scientific Research*, vol. 77, pp. 362–372, 2012.

[6] P.-L. Huang, B. Lee, and C.-C. Chen, "The influence of service quality on customer satisfaction and loyalty in B2b technology service industry," *Total Quality Management and Business Excellence*, vol. 30, no. 13-14, pp. 1449–1465, 2019.

[7] R. Hussain, A. Al Nasser, and Y. K. Hussain, "Service quality and customer satisfaction of a uae-based airline: an empirical investigation," *Journal of Air Transport Management*, vol. 42, pp. 167–175, 2015.

[8] I.-M. Wang, C.-J. Shieh, J. Hsiao, and O. Sciences, "The relationship between service quality, customer satisfaction and customer loyalty: a study on the management consulting industry," *Journal of Information and Optimization Sciences*, vol. 26, no. 2, pp. 371–384, 2005.

[9] C.-C. Yang and C.-C. Chao, "How relationship marketing, switching costs, and service quality impact customer satisfaction and loyalty in Taiwan's airfreight forwarding industry?" *Transportmetrica: Transportation Science*, vol. 13, no. 8, pp. 679–707, 2017.

[10] S. Kanat, S. A. Abbasi, M. H. Peerzada, and T. Atilgan, "Swot analysis of Pakistan's textile and clothing industry," *Industria Textila*, vol. 69, no. 6, p. 502, 2018.

[11] W. H. Kunz and J. Hogreve, "Toward a deeper understanding of service marketing: the past, the present, and the future," *International Journal of Research in Marketing*, vol. 28, no. 3, pp. 231–247, 2011.

[12] R. Kaplinsky and M. Morris, *A Handbook for Value Chain Research*, University of Sussex, Brighton, UK, 2001.

[13] A. L. Dixon and J. F. Tanner, "Transforming selling: why it is time to think differently about sales research," *Journal of Personal Selling and Sales Management*, vol. 32, no. 1, pp. 9–13, 2012.

[14] J. L. Heskett, T. O. Jones, G. W. Loveman, W. E. Sasser, and L. A. Schlesinger, "Putting the service-profit chain to work," *Journal Harvard Business Review*, vol. 72, no. 2, pp. 164–174, 1994.

[15] I. Russo, I. Confente, D. M. Gligor, and C. W. Autry, "To Be or not to Be (loyal): is there a recipe for customer loyalty in the B2b context?" *Journal of Business Research*, vol. 69, no. 2, pp. 888–896, 2016.

[16] S. Vandermerwe and J. Rada, "Servitization of business: adding value by adding services," *European Management Journal*, vol. 6, no. 4, pp. 314–324, 1988.

[17] X. N. Deng, T. Wang, and R. D. Galliers, "More than providing 'solutions': towards an understanding of customer-oriented citizenship behaviours of IS professionals,"

Information Systems Journal, vol. 25, no. 5, pp. 489–530, 2015.

[18] C.-F. Chen, "Investigating structural relationships between service quality, perceived value, satisfaction, and behavioral intentions for air passengers: evidence from taiwan," *Transportation Research Part A: Policy and Practice*, vol. 42, no. 4, pp. 709–717, 2008.

[19] D. S. Rogers, D. M. Lambert, K. L. Croxton, and S. J. García-Dastugue, "The returns management process," *International Journal of Logistics Management*, vol. 13, no. 2, pp. 1–18, 2002.

[20] J. Stock, T. Speh, and H. Shear, "Managing product returns for competitive advantage," *MIT Sloan Management Review*, vol. 48, no. 1, p. 57, 2006.

[21] D. A. Mollenkopf, R. Frankel, and I. Russo, "Creating value through returns management: exploring the marketing-operations interface," *Journal of Operations Management*, vol. 29, no. 5, pp. 391–403, 2011.

[22] D. Mollenkopf, I. Russo, and R. Frankel, "The returns management process in supply chain strategy," *International Journal of Physical Distribution & Logistics Management*, vol. 37, 2007.

[23] S. Y. Lam, V. Shankar, M. K. Erramilli, and B. Murthy, "Customer value, satisfaction, loyalty, and switching costs: an illustration from a business-to-business service context," *Journal of the Academy of Marketing Science*, vol. 32, no. 3, pp. 293–311, 2004.

[24] K. Matzler, A. Strobl, N. Thurner, and J. Füller, "Switching experience, customer satisfaction, and switching costs in the ict industry," *Journal of Service Management*, vol. 26, no. 1, pp. 117–136, 2015.

[25] T. B. White, K. N. Lemon, and J. E. Hogan, "Customer retention when the customer's future usage is uncertain," *Psychology and Marketing*, vol. 24, no. 10, pp. 849–870, 2007.

[26] R. A. Ping, "The effects of satisfaction and structural constraints on retailer exiting, voice, loyalty, opportunism, and neglect," *Journal of Retailing*, vol. 69, no. 3, pp. 320–352, 1993.

[27] M. E. Porter, "Industry structure and competitive strategy: keys to profitability," *Financial Analysts Journal*, vol. 36, no. 4, pp. 30–41, 1980.

[28] C. Bozzo, "Understanding inertia in an industrial context," *Journal of Customer Behaviour*, vol. 1, no. 3, pp. 335–355, 2002.

[29] A. S. Dick and K. Basu, "Customer loyalty: toward an integrated conceptual framework," *Journal of the Academy of Marketing Science*, vol. 22, no. 2, pp. 99–113, 1994.

[30] M. Kim, D. Kliger, and B. Vale, "Estimating switching costs: the case of banking," *Journal of Financial Intermediation*, vol. 12, no. 1, pp. 25–56, 2003.

[31] P. Klemperer, "Markets with consumer switching costs," *Quarterly Journal of Economics*, vol. 102, no. 2, pp. 375–394, 1987.

[32] M. Blut, C. M. Frennea, V. Mittal, and D. L. Mothersbaugh, "How procedural, financial and relational switching costs affect customer satisfaction, repurchase intentions, and repurchase behavior: a meta-analysis," *International Journal of Research in Marketing*, vol. 32, no. 2, pp. 226–229, 2015.

[33] A. Caruana and M. T. Ewing, "How corporate reputation, quality, and value influence online loyalty," *Journal of Business Research*, vol. 63, no. 9-10, pp. 1103–1110, 2010.

[34] Z. Yang and R. T. Peterson, "Customer perceived value, satisfaction, and loyalty: the role of switching costs," *Psychology and Marketing*, vol. 21, no. 10, pp. 799–822, 2004.

[35] K. Hjort and B. Lantz, "The impact of returns policies on profitability: a fashion E-commerce case," *Journal of Business Research*, vol. 69, no. 11, pp. 4980–4985, 2016.

[36] A. Minnema, T. H. A. Bijmolt, J. A. Petersen, and J. D. Shulman, *Managing Product Returns within the Customer Value Framework*, Springer, Berlin, Germany, 2018.

[37] J. A. Petersen and V. Kumar, "Can product returns make you money?" *MIT Sloan Management Review*, vol. 51, no. 3, p. 85, 2010.

[38] R. Tibben-Lembke, "Life after death: reverse logistics and the product life cycle," *International Journal of Physical Distribution & Logistics Management*, vol. 32, 2002.

[39] M. H. Askariazad and N. Babakhani, "An application of European customer satisfaction Index (ecsi) in business to business (B2b) context," *Journal of Business Industrial Marketing*, vol. 30, 2015.

[40] C. W. Autry, "Formalization of reverse logistics programs: a strategy for managing liberalized returns," *Industrial Marketing Management*, vol. 34, no. 7, pp. 749–757, 2005.

[41] C. A. Reeves and D. A. Bednar, "Defining quality: alternatives and implications," *Academy of Management Review*, vol. 19, no. 3, pp. 419–445, 1994.

[42] E. Arı and V. Yılmaz, "Consumer attitudes on the use of plastic and cloth bags," *Environment, Development and Sustainability*, vol. 19, no. 4, pp. 1219–1234, 2017.

[43] G. Fullerton, "How commitment both enables and undermines marketing relationships," *European Journal of Marketing*, vol. 39, 2005.

[44] M. Paiola, N. Saccani, M. Perona, and H. Gebauer, "Moving from products to solutions: strategic approaches for developing capabilities," *European Management Journal*, vol. 31, no. 4, pp. 390–409, 2013.

[45] J. J. Cronin and S. A. Taylor, "Measuring service quality: a reexamination and extension," *Journal of Marketing*, vol. 56, no. 3, pp. 55–68, 1992.

[46] A. G. Woodside, L. L. Frey, and R. T. Daly, "Linking sort/ice anlity, customer satisfaction, and behavioral intention," *Journal of Health Care Marketing*, vol. 9, no. 4, pp. 5–17, 1989.

[47] V. A. Zeithaml, L. L. Berry, and A. Parasuraman, "The behavioral consequences of service quality," *Journal of Marketing*, vol. 60, no. 2, pp. 31–46, 1996.

[48] S. Sattari, E. S. Sangari, and K. Peighambari, "Service quality in the Iranian cellular telecommunications market," in *Proceedings of the 2009 Academy of Marketing Science (AMS) Annual Conference*, Portland, OR, USA, May 2015.

[49] E. Cengiz, "Measuring customer satisfaction: must or not," *Journal of naval science engineering*, vol. 6, no. 2, pp. 76–88, 2010.

[50] P. Kotler, *Marketing Management: Analysis, Planning, Implementation and Control*, Prentice-Hall, Englewood, CO, USA, 7th edition, 1997.

[51] E. W. Anderson and M. W. Sullivan, "The antecedents and consequences of customer satisfaction for firms," *Marketing Science*, vol. 12, no. 2, pp. 125–143, 1993.

[52] A. Parasuraman, "Customer service in business-to-business markets: an agenda for research," *Journal of Business & Industrial Marketing*, vol. 13, no. 4/5, pp. 309–321, 1998.

[53] J. B. Gotlieb, D. Grewal, and S. W. Brown, "Consumer satisfaction and perceived quality: complementary or divergent constructs?" *Journal of Applied Psychology*, vol. 79, no. 6, pp. 875–885, 1994.

[54] P. G. Patterson, L. W. Johnson, and R. A. Spreng, "Modeling the determinants of customer satisfaction for business-to-

business professional services," *Journal of the Academy of Marketing Science*, vol. 25, no. 1, pp. 4–17, 1996.
[55] A. Caruana, A. H. Money, and P. R. Berthon, "Service quality and satisfaction-the moderating role of value," *European Journal of Marketing*, vol. 34, 2000.
[56] B. Osayawe Ehigie, "Correlates of customer loyalty to their bank: a case study in Nigeria," *International Journal of Bank Marketing*, vol. 24, no. 7, pp. 494–508, 2006.
[57] C. Guo and Y. Wang, "How manufacturer market orientation influences B2b customer satisfaction and retention: empirical investigation of the three market orientation components," *Journal of Business Industrial Marketing*, vol. 30, 2015.
[58] E. E. Izogo and I.-E. Ogba, "Service quality, customer satisfaction and loyalty in automobile repair services sector," *International Journal of Quality & Reliability Management*, vol. 32, 2015.
[59] F. Olorunniwo and M. K. Hsu, "A typology analysis of service quality, customer satisfaction and behavioral intentions in mass services," *Managing Service Quality: International Journal*, vol. 16, 2006.
[60] I. Santouridis and P. Trivellas, "Investigating the impact of service quality and customer satisfaction on customer loyalty in mobile telephony in Greece," *The TQM Journal*, vol. 22, 2010.
[61] N. Sofnia and R. B. Paramarta, "Perceived quality, customer satisfaction, switching barrier and customer loyalty in business to business context," *International Journal of Scientific Engineering and Research*, vol. 10, no. 2, 2019.
[62] M. Peyrot and D. Doren, "Effect of a class Action suit on consumer repurchase intentions," *Journal of Consumer Affairs*, vol. 28, no. 2, pp. 361–379, 1994.
[63] T. Hennig-Thurau, K. P. Gwinner, G. Walsh, and D. D. Gremler, "Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet?" *Journal of Interactive Marketing*, vol. 18, no. 1, pp. 38–52, 2004.
[64] R. M. Morgan and S. D. Hunt, "The commitment-trust theory of relationship marketing," *Journal of Marketing*, vol. 58, no. 3, pp. 20–38, 1994.
[65] S. A. Narayana, A. A. Elias, and R. K. Pati, "Reverse logistics in the pharmaceuticals industry: a systemic analysis," *International Journal of Logistics Management*, vol. 25, 2014.
[66] M. Zineldin, "The quality of health care and patient satisfaction," *International Journal of Health Care Quality Assurance*, vol. 19, 2006.
[67] G. F. Watson, J. T. Beck, C. M. Henderson, and R. W. Palmatier, "Building, measuring, and profiting from customer loyalty," *Journal of the Academy of Marketing Science*, vol. 43, no. 6, pp. 790–825, 2015.
[68] B. M. Byrne, "Structural equation modeling with amos, eqs, and lisrel: comparative approaches to testing for the factorial validity of a measuring instrument," *International Journal of Testing*, vol. 1, no. 1, pp. 55–86, 2001.
[69] A. Parasuraman, V. Zeithaml, and L. Berry, "Servqual: a multi item scale for measuring consumer perception of service quality," *Journal of Retailing*, vol. 64, pp. 12–40, 1988.
[70] S. A. Taylor and T. L. Baker, "An assessment of the relationship between service quality and customer satisfaction in the formation of consumers' purchase intentions," *Journal of Retailing*, vol. 70, no. 2, pp. 163–178, 1994.
[71] A. H. Liu, *Examining the Role of Customer Value, Customer Satisfaction, and Perceived Switching Costs: A Model of Repurchase Intention for Business-To-Business Services*, Georgia State University, Ann Arbor, MI, USA, 1998.
[72] J. Hair, W. Black, B. Babin, and R. Anderson, *Multivariate Data Analysis*, Prentice Hall, Upper Saddle River, NY, USA, 1998.
[73] H. F. Kaiser, "A second generation little jiffy," *Psychometrika*, vol. 35, no. 4, pp. 401–415, 1970.
[74] C. Fornell and D. F. Larcker, "Evaluating structural equation models with unobservable variables and measurement error," *Journal of Marketing Research*, vol. 18, no. 1, pp. 39–50, 1981.
[75] N. K. Avkiran and C. M. Ringle, *Partial Least Squares Structural Equation Modeling: Recent Advances in Banking and Finance*, Springer, Basingstoke, UK, 2018.
[76] J. A. Wollack, A. S. Cohen, and C. S. Wells, "A method for maintaining scale stability in the presence of test speededness," *Journal of Educational Measurement*, vol. 40, no. 4, pp. 307–330, 2003.
[77] J. C. Nunnally, *Psychometric Theory*, McGraw-Hill, New York, NY, USA, 2nd edition, 1978.
[78] J. Hair, W. Black, B. Babin, and R. Anderson, *Multivariate Data Analysis: A Global Perspective*, Pearson, London, UK, Seventh edition, 2010.
[79] N. Kock, "Common method bias in PLS-SEM," *International Journal of E-Collaboration*, vol. 11, no. 4, pp. 1–10, 2015.
[80] B. Craig, D. Moore, and G. McCabe, *Introduction to the Practice of Statistics*, WH Freeman and Company, New York, NY, USA, 6th edition, 2009.
[81] R. P. Bagozzi and Y. Yi, "On the evaluation of structural equation models," *Journal of the Academy of Marketing Science*, vol. 16, no. 1, pp. 74–94, 1988.
[82] Arbuckle, *Amos 6.0 User's Guide*, Amos Development Corporation, Crawfordville, GA, USA, 2005.
[83] R. S. Padilla, S. K. Milton, L. W. Johnson, and M. W. Nyadzayo, "Impact of service value on satisfaction and repurchase intentions in business-to-business cloud computing," *Service Science*, vol. 9, no. 1, pp. 5–13, 2017.
[84] I. Russo, I. Confente, D. M. Gligor, and N. Cobelli, "The combined effect of product returns experience and switching costs on B2b customer Re-purchase intent," *Journal of Business & Industrial Marketing*, vol. 32, no. 5, pp. 664–676, 2017.
[85] S. Hsu, Hung, and B. Bayarsaikhan, "Factors influencing on online shopping attitude and intention of Mongolian consumers," *The Journal of International Management Studies*, vol. 7, no. 2, 2012.
[86] J. Jiang, C. Chen, and C. Wang, "Knowledge and trust in E-consumers' online shopping behavior," in *Proceedings of the 2008 International Symposium on Electronic Commerce and Security*, Guangzhou, China, August 2008.
[87] Y. B. Limbu, M. Wolf, and D. Lunsford, "Perceived ethics of online retailers and consumer behavioral intentions," *The Journal of Research in Indian Medicine*, vol. 6, no. 2, pp. 133–154, 2012.
[88] A. Mukherjee and P. Nath, "Role of electronic trust in online retailing," *European Journal of Marketing*, vol. 41, no. 9/10, pp. 1173–1202, 2007.
[89] J. Sutanonpaiboon and M. N. Hamimah, "Malay, Chinese, and internet banking," *Chinese Management Studies*, vol. 4, no. 2, pp. 141–153, 2010.
[90] C. Whyte, "Developed states' vulnerability to economic disruption online," *Orbis*, vol. 60, no. 3, pp. 417–432, 2016.

[91] D. M. Lambert and M. G. Enz, "Issues in supply chain management: progress and potential," *Industrial Marketing Management*, vol. 62, pp. 1–16, 2017.
[92] J. A. Petersen and V. Kumar, "Perceived risk, product returns, and optimal resource allocation: evidence from a field experiment," *Journal of Marketing Research*, vol. 52, no. 2, pp. 268–285, 2015.
[93] E. E. Otaigbe: Exploring Satisfaction and Retention Strategies for Generation Y Customers.
[94] J. Álvarez-García, M. D. L. C. del Río-Rama, and B. Simonetti, "Quality management and customer results: the tourist accommodation sector in Spain," *International Journal of Quality & Reliability Management*, vol. 34, no. 5, pp. 701–719, 2017.
[95] Y.-C. Huang, M.-L. Yang, and Y.-J. Wong, "Institutional pressures, resources commitment, and returns management," *Supply Chain Management: International Journal*, vol. 32, 2016.
[96] S.-M. Meng, G.-S. Liang, K. Lin, and S.-Y. Chen, "Criteria for services of air cargo logistics providers: how do they relate to client satisfaction?" *Journal of Air Transport Management*, vol. 16, no. 5, pp. 284–286, 2010.
[97] W. Shen, W. Xiao, and X. Wang, "Passenger satisfaction evaluation model for urban rail transit: a structural equation modeling based on partial least squares," *Transport Policy*, vol. 46, pp. 20–31, 2016.
[98] M. Baidya and G. Ghosh, "An empirical investigation of repeat buying behavior of customers of two brands in India," *Journal of Indian Business Research*, vol. 6, 2014.
[99] S. Basuroy, K. C. Gleason, and Y. H. Kannan, "Ceo compensation, customer satisfaction, and firm value," *Review of Accounting and Finance*, vol. 13, 2014.
[100] K. L. Bernhardt, N. Donthu, and P. A. Kennett, "A longitudinal analysis of satisfaction and profitability," *Journal of Business Research*, vol. 47, no. 2, pp. 161–171, 2000.
[101] J. Eskildsen, A. H. Westlund, and K. Kristensen, "The predictive power of intangibles," *Measuring Business Excellence*, vol. 7, 2003.
[102] K. Kristensen, L. S. Mørch, and H. Sørensen, "Relationship between performance measures and financial results in a large nordic bank," *Journal of Performance Measurement Management*, vol. 2006, pp. 1011–1017, 2006.
[103] A. Rucci, S. P. Kirn, and R. T. Quinn, "The employee-customer-profit chain at sears," vol. 76, pp. 82–98, 1998.
[104] W. Shi, L. Tang, X. Zhang, Y. Gao, and Y. Zhu, "How does word of mouth affect customer satisfaction?" *Journal of Business Industrial Marketing*, vol. 31, 2016.
[105] A. M. Dean, M. Griffin, and A. Kulczynski, "Applying service logic to education: the Co-creation experience and value outcomes," *Procedia-Social and Behavioral Sciences*, vol. 224, pp. 325–331, 2016.
[106] R. A. Pollak and S. C. Watkins, "Cultural and economic approaches to fertility: proper marriage or mesalliance?" *Population and Development Review*, vol. 19, no. 3, pp. 467–496, 1993.

Hindawi
Complexity
Volume 2021, Article ID 6911443, 11 pages
https://doi.org/10.1155/2021/6911443

Research Article

Some Further Results on the Reduction of Two-Dimensional Systems

Dongmei Li, Yingying Gui, Jinwang Liu, and Man Wu

School of Mathematics and Computing Sciences, Hunan University of Science and Technology, Xiangtan, Hunan 411201, China

Correspondence should be addressed to Yingying Gui; 19010701003@mail.hnust.edu.cn

Received 5 August 2021; Accepted 29 October 2021; Published 17 November 2021

Academic Editor: Xiao Ling Wang

The reduction of two-dimensional systems plays an important role in the theory of systems, which is closely associated with the equivalence of the bivariate polynomial matrices. In this paper, the equivalence problems on several classes of bivariate polynomial matrices are investigated. Some new results on the equivalence of these matrices are obtained. These results are useful for reducing two-dimensional systems.

1. Introduction

Multidimensional (nD) systems, especially two-dimensional (2D) systems, are widely used in the field of circuits, image, signal processing, control systems, etc. [1–5, 12]. And the theory of 2D systems has received increasing attention, among which the reduction of 2D systems is an important research content. Usually, a given system is desirable to reduce to an equivalent system with fewer equations or unknowns (named the equivalence of 2D systems). In this way, the characteristics of the original system can be studied in a better and simpler way. A 2D system can be represented with two types of dynamical elements, so 2D systems are often described by bivariate polynomial matrices. Hence, the equivalence problems of 2D systems are usually transformed into the equivalence problems of polynomial matrices [6–11].

The equivalence problem of univariate polynomial matrices was solved by Rosenbrock in 1970 [6]. Using the Euclidean division property of the univariate polynomial ring, he proved that every univariate polynomial matrix in this ring is equivalent to the Smith form. The equivalence problem of the bivariate polynomial matrix is more complex. Lee and Zak gave an example of bivariate polynomial matrix $\begin{pmatrix} z_2 & -z_1-1 \\ -z_2^2 & z_2 \end{pmatrix}$, which is not equivalent to the Smith form [8]. Note that none of the bivariate polynomial rings is a Euclidean ring, and the Euclidean division property does not hold. So, many researchers study the equivalence of some special types of bivariate polynomial matrices. In 1986, Frost and Boudellioua proved that a full row rank bivariate polynomial matrix $T(z,w)$ is equivalent to the Smith form if and only if there exists a unimodular column vector U such that $(T\ U)$ has a right inverse [9]. In 2012, Boudellioua wrote an algorithm to enable the decomposition and equivalence of some bivariate polynomial matrices to be realized in Maple [10]. In 2019, Li et al. presented some criteria on an $l\times l$ bivariate polynomial matrix $F(z,w)$ which is equivalent to the Smith form diag$(I_{l-1}, \det F(z,w))$ with $\det F(z,w)$ being an irreducible polynomial [11]. There are also some results on the equivalence of multivariate polynomial matrices in special cases [13, 14, 16, 18, 19], among which when an $l\times l$ polynomial matrix $F(z)$ which is equivalent to the Smith form diag $(I_{l-1}, \det F(z))$ is investigated, such as $\det F(z) = x_1 - f(x_2,\ldots,x_n)$ [13], $\det F(z) = (x_1 - f(x_2,\ldots,x_n))^q$ [14], and $\det F(z) = (x_1 - f_1(x_2,\ldots,x_n))^{q_1}(x_2 - f_2(x_2,\ldots,x_n))^{q_2}$ [19], diag $(I_{l-1}, \det F(z))$ is a very important kind of matrices in the equivalence of multidimensional systems [9, 15, 17].

In this paper, $K[z,w]$ denotes a bivariate polynomial ring with K being a field, and we consider arbitrary polynomial $f(z,w)$ in $K[z,w]$ as a polynomial in w with

coefficient in $K[z]$, written as $f(z,w)=\sum_{i=0}^{n}\alpha_i w^i$, where $\alpha_i=\alpha_i(z)\in K[z]$. Note that the coefficient ring $K[z]$ is a Euclidean ring, and combined with the Euclidean division property of $K[z]$, we will investigate some classes of bivariate polynomial matrices with their entries in $K[z,w]$. The following three problems are also considered.

Problem 1. Let $F(z,w)\in K^{l\times l}[z,w]$, and $\det F(z,w)=h^r(z)$, where $h(z)\in K[z]$ is irreducible and r is a positive integer. When is $F(z,w)$ equivalent to

$$\begin{pmatrix} I_{l-r} & & & \\ & h(z) & & \\ & & \ddots & \\ & & & h(z) \end{pmatrix}? \tag{1}$$

Problem 2. Let $F(z,w)\in K^{l\times l}[z]$ with $\det F(z,w)=h^{q\cdot r}(z)$, $h(z)\in K[z]$ be irreducible, and q,r be positive integers. When is $F(z,w)$ equivalent to

$$\begin{pmatrix} I_{l-r} & & & \\ & h^q(z) & & \\ & & \ddots & \\ & & & h^q(z) \end{pmatrix}? \tag{2}$$

Problem 3. Let $F(z,w)\in K^{l\times m}[z]$ $(l\le m)$ with $d_l(F)=h^{q\cdot r}(z)$, where $d_l(F)$ denotes the greatest common divisor of the $l\times l$ minor of $F(z,w)$, $h(z)\in K[z]$ is irreducible, and q,r are positive integers. When is $F(z,w)$ equivalent to its Smith form?

2. Preliminaries

In the following, K is an arbitrary field, $K[z]$ is the univariate polynomial ring in variable z with coefficients in K, $K[z,w]$ is the bivariate polynomial ring in variables z,w whose coefficients are in K, $\overline{K}$ is the algebraic closed field of K, $0_{m,n}$ is the $m\times n$ zero matrix, and I_m is the $m\times m$ identity matrix. For $F(z,w)\in K^{l\times m}[z,w]$, $d_i(F)$ will be the greatest common divisor (g.c.d) of the $i\times i$ minors of $F(z,w)$, $i=1,\ldots,l$. Set $F(z,w)=(f_{ij})\in K^{l\times m}[z,w]$, where $f_{ij}\in K[z,w]$, $h(z)\in K[z]$ is irreducible, and $\overline{F(z,w)}=(\overline{f_{ij}})\in K^{l\times m}[z,w]$, where $\overline{f_{ij}}$ denotes $f_{ij}\bmod h(z)$. For convenience, the argument (z,w) is omitted whenever its omission does not cause confusion throughout this paper.

Definition 1 (see [16]). Let $F(z,w)\in K^{l\times m}[z,w]$ $(l\le m)$ with rank r and Φ_i be a polynomial defined as follows:

$$\Phi_i=\begin{cases}\dfrac{d_i(F)}{d_{i-1}(F)}, & 1\le i\le r,\\[2ex] 0, & r<i\le l,\end{cases} \tag{3}$$

where $d_0(F)\equiv 1$, $d_i(F)$ is the g.c.d of the $i\times i$ minors of $F(z,w)$, and Φ_i satisfies

$$\Phi_1|\Phi_2|\cdots|\Phi_r. \tag{4}$$

Then, the Smith form of $F(z,w)$ is given by

$$S=\begin{pmatrix}\operatorname{diag}\{\Phi_i\} & 0_{r,m-r}\\ 0_{l-r,r} & 0_{l-r,m-r}\end{pmatrix}. \tag{5}$$

Definition 2 (see [16]). Let $F(z,w)\in K^{l\times m}[z,w]$ with full row rank; $F(z,w)$ is said to be zero left prime if the $l\times l$ minors of $F(z,w)$ have no common zero in $\overline{K}^2$.

Definition 3. Let $U(z,w)\in K^{l\times l}[z,w]$; then, we say $U(z,w)$ to be a unimodular matrix if the determinant of $U(z,w)$ is a unit of K.

Definition 4 Let $F_1(z,w),F_2(z,w)\in K^{l\times m}[z,w]$; $F_1(z,w)$ is said to be equivalent to $F_2(z,w)$ if there are unimodular matrices $U(z,w)\in K^{l\times l}[z,w]$ and $V(z,w)\in K^{m\times m}[z,w]$ such that

$$F_1(z,w)=U(z,w)F_2(z,w)V(z,w). \tag{6}$$

Lemma 1 (see [16]). *Suppose $F(z,w),Q(z,w)\in K^{l\times l}[z,w]$; if $F(z,w)$ is equivalent to $Q(z,w)$, then $d_i(F(z,w))=d_i(Q(z,w))$, for $i=1,\ldots,l$.*

Lemma 2 (see [16]). *Let $F(z,w)$, $F_1(z,w)$, $F_2(z,w)\in K^{l\times l}[z,w]$, and $F(z,w)=F_1(z,w)F_2(z,w)$. If the $(l-r)\times(l-r)$ $(r\le l)$ minors of $F(z,w)$ have no common zero in $\overline{K}^2$, then the $(l-r)\times(l-r)$ minors of F_i have no common zero in $\overline{K}^2$ for $i=1,2$.*

3. Main Results

In this section, we investigate the three problems presented in Section 1 and give the main results of this paper.

In the following, for $f(z,w)\in K[z,w]$, we consider it as an element in $K[z][w]$, written as $f(z,w)=\sum_{i=0}^{n}\alpha_i w^i$, where $\alpha_i=\alpha_i(z)\in K[z]$. If $h(z)\in K[z]$ is irreducible, then $\overline{f(z,w)}$ denotes $f(z,w)\bmod h(z)$ and $\overline{f(z,w)}\equiv\sum_{i=0}^{n_1}\beta_i w^i\pmod{h(z)}$ or 0, where $\beta_i=\beta_i(z)\in K[z]$, $\deg\beta_i(z)<\deg h(z)$. Hence, $\beta_{n_1}(z)$ and $h(z)$ are relatively prime in $K[z]$; by Euclidean algorithm, there are $x(z),y(z)\in K[z]$ such that $x(z)\cdot\beta_{n_1}(z)=1-y(z)\cdot h(z)$, i.e., $x(z)\cdot\beta_{n_1}(z)\equiv 1\pmod{h(z)}$. Then, $x(z)\cdot\overline{f(z,w)}\equiv f^*(z,w)\pmod{h(z)}$, where $f^*(z,w)$ is monic. In other words, $x(z)\overline{f(z,w)}$ can be reduced to a monic polynomial $f^*(z,w)$ by $h(z)$.

Denote

$$P(z)=\begin{pmatrix} I_{l-r} & & & \\ & h(z) & & \\ & & \ddots & \\ & & & h(z)\end{pmatrix}. \tag{7}$$

First, we investigate Problem 1.

Theorem 1. *Let $F(z,w) \in K^{l\times l}[z,w]$ with $\det F(z,w) = h^r(z)$, where $h(z) \in K[z]$ is irreducible. If $h(z) | d_{l-r+1}(F(z,w))$, then $F(z,w)$ is equivalent to the Smith form $P(z)$.*

Proof. If r rows of $F(z,w)$ are zero vectors mod $h(z)$, that is, r rows of $\overline{F(z,w)}$ are zero vectors, we obtain that

$$F(z,w) = U(z,w)\begin{pmatrix} I_{l-r} & & & \\ & h(z) & & \\ & & \ddots & \\ & & & h(z) \end{pmatrix} Q(z,w), \quad (8)$$

and then $F(z,w) = U(z,w)P(z)Q(z,w)$, where $U(z,w)$ is unimodular; by computing, $\det Q(z,w) = 1$, so $Q(z,w)$ is unimodular. Hence, $F(z,w)$ is equivalent to the Smith form $P(z)$.

If $F(z,w)$ has no r rows which are the zero vectors mod $h(z)$, then $F(z,w)$ has r_1 rows of zero vectors mod $h(z)$; in other words, $\overline{F(z,w)}$ has r_1 rows of zero vectors, $0 \le r_1 \le r-1$. We premultiply and postmultiply $\overline{F(z,w)}$ by unimodular matrices U_{11} and V_{11} such that

$$U_{11}\overline{F(z,w)}V_{11} = \begin{pmatrix} \overline{f(z,w)} & \\ \overline{f_2(z,w)} & \\ \vdots & X \\ \overline{f_{l-r_1}(z,w)} & \\ 0_{r_1,1} & 0_{r_1,l-1} \end{pmatrix}, \quad (9)$$

where $\overline{f(z,w)} \neq 0$, $\deg_w \overline{f(z,w)} \le \deg_w \overline{f_j(z,w)}$, or $\overline{f_j(z,w)} = 0$, $j = 2,\ldots,l-r_1$. Let $\overline{f(z,w)} = \alpha_0 + \alpha_1 w + \cdots + \alpha_n w^n$, where $\deg_z \alpha_i < \deg_z h(z)$, $i = 0,\ldots,n$, and $\alpha_n \neq 0$. Note that α_n and $h(z)$ are relatively prime, so we can find $x(z), y(z) \in K[z]$ such that $x(z)\cdot\alpha_n + y(z)\cdot h(z) = 1$; then, $x(z)\cdot\alpha_n = 1 - y(z)\cdot h(z)$. We have $x(z)\cdot\overline{f(z,w)} \equiv f'(z,w) (\mathrm{mod}\, h(z))$, where $f'(z,w)$ is monic. There are $q_j(z,w), r_j(z,w) \in K[z,w]$ such that $\overline{f_j(z,w)} = q_j(z,w)\cdot f'(z,w) + r_j(z,w)$, where $\deg_w r_j(z,w) < \deg_w f'(z,w)$ or $r_j(z,w) = 0$, $j = 2,\ldots,l-r_1$. Therefore,

$$\overline{f_j(z,w)} \equiv q_j(z,w)\cdot x(z)\cdot\overline{f(z,w)} + \overline{r_j(z,w)} (\mathrm{mod}\, h(z)), \quad (10)$$

where $\deg_w \overline{r_j(z,w)} < \deg_w \overline{f(z,w)}$ or $\overline{r_j(z,w)} = 0$, $j = 2,\ldots,l-r_1$.

Let

$$U_{12} = \begin{pmatrix} 1 & 0_{1,l-1} \\ -x(z)q_2(z,w) & \\ \vdots & I_{l-1} \\ -x(z)q_{l-r_1}(z,w) & \\ 0_{r_1,l-r_1} & \end{pmatrix}. \quad (11)$$

Then,

$$U_{12}U_{11}\overline{F(z,w)}V_{11} = \begin{pmatrix} \overline{f(z,w)} & \\ \overline{r_2(z,w)} & \\ \vdots & X \\ \overline{r_{l-r_1}(z,w)} & \\ 0_{r_1,1} & 0_{r_1,l-1} \end{pmatrix} \mathrm{mod} h(z), \quad (12)$$

where $\deg_w \overline{r_j(z,w)} < \deg_w \overline{f(z,w)}$ or $\overline{r_j(z,w)} = 0$, $j = 2,\ldots,l-r_1$. If some of $\overline{r_j(z,w)}$ are not 0, $j = 2,\ldots,l-r_1$, do some row transformations to the matrix $U_{12}U_{11}\overline{F(z,w)}V_{11}$ such that the nonzero polynomial of the least degree in w among its first column is at position (1, 1). Repeating the previous steps, we obtain that

$$U_1\overline{F(z,w)}V_{11} = \begin{pmatrix} \overline{d}_{11} & X \\ 0_{l-1,1} & \overline{F_1(z,w)} \end{pmatrix} \mathrm{mod}\, h(z), \quad (13)$$

where $U_1 = U_{1t}\cdots U_{12}U_{11}$ is a unimodular matrix, the last r_1 rows of $\overline{F_1(z,w)}$ are zero vectors, and $\overline{d}_{11} \neq 0$.

If r rows of $\overline{F_1(z,w)}$ are zero vectors mod $h(z)$, we can find two $(l-1)\times(l-1)$ unimodular matrices U'_{20}, V'_{20} such that

$$U'_{20}\overline{F_1(z,w)}V'_{20} = \begin{pmatrix} B(z,w) \\ 0_{r,l-1} \end{pmatrix} \mathrm{mod}\, h(z), \quad (14)$$

where $B(z,w) \in K^{(l-r-1)\times(l-1)}[z,w]$.

Let $U_{20} = \begin{pmatrix} 1 & \\ & U'_{20} \end{pmatrix}$ and $V_{20} = \begin{pmatrix} 1 & \\ & V'_{20} \end{pmatrix}$; we have that

$$U_{20}U_1F(z,w)V_{11}V_{20} = \begin{pmatrix} I_{l-r} & & & \\ & h(z) & & \\ & & \ddots & \\ & & & h(z) \end{pmatrix} Q_1, \quad (15)$$

where $Q_1 \in K^{l\times l}[z,w]$ and $\det Q_1 = 1$. Then, $F(z,w) = UP(z)V$, where $U = U_1^{-1}U_{20}^{-1}$ and $V = Q_1V_{20}^{-1}V_{11}^{-1}$ are unimodular matrices, so $F(z,w)$ is equivalent to the Smith form $P(z)$.

If $\overline{F_1(z,w)}$ has no r rows of zero vectors mod $h(z)$, then $\overline{F_1(z,w)}$ has r_2 rows of zero vectors, $r_1 \le r_2 \le r-2$. Imitating the previous procedure to $\overline{F(z,w)}$, there are two $(l-2)\times(l-2)$ unimodular matrices U'_{21}, V'_{21} such that

$$U'_{21}\overline{F_1(z,w)}V'_{21} = \begin{pmatrix} \overline{d}_{22} & X \\ 0_{l-2,1} & \overline{F_2(z,w)} \end{pmatrix} \mathrm{mod}\, h(z), \quad (16)$$

where $\overline{d}_{22} \neq 0$ and the last r_2 rows of $\overline{F_1(z,w)}$ are zero vectors.

Let $U_{21} = \begin{pmatrix} 1 & \\ & U'_{21} \end{pmatrix}$ and $V_{21} = \begin{pmatrix} 1 & \\ & V'_{21} \end{pmatrix}$; then,

$$U_{21}U_1\overline{F(z,w)}V_{11}V_{21} = \begin{pmatrix} \overline{d}_{11} & * & X \\ & \overline{d}_{22} & \\ 0_{l-2,2} & & \overline{F_2(z,w)} \end{pmatrix} \mathrm{mod}\, h(z). \quad (17)$$

Repeating the procedure above successively, we obtain a series of $\overline{F_i(z,w)}$, $i = 3, \ldots, l-r-1$. If there is some $\overline{F_i(z,w)}$ which contains r rows of zero vectors mod $h(z)$, then the conclusion is straightforward. Otherwise, $\overline{F_i(z,w)}$ has no r rows of zero vectors mod $h(z)$.

Then, we consider the case that $\overline{F_i(z,w)}$ has no r rows of zero vectors mod $h(z)$, $i = 3, \ldots, l-r-1$. In this case, $\overline{F_{l-r-1}(z,w)} \in K^{(r+1)\times(r+1)}[z,w]$, and it has no r rows of zero vectors mod $h(z)$; there are $(r+1)\times(r+1)$ unimodular matrices $U'_{l-r,1}, V'_{l-r,1}$ such that

$$U'_{l-r,1}\overline{F_{l-r-1}(z,w)}V'_{l-r,1} = \begin{pmatrix} \overline{d}_{l-r,l-r} & X \\ 0_{r,l-r} & \overline{F_{l-r}(z,w)} \end{pmatrix} \bmod h(z), \tag{18}$$

where $\overline{d}_{l-r,l-r} \neq 0$.

Let $U_{l-r,1} = \begin{pmatrix} I_{l-r-1} & \\ & U'_{l-r,1} \end{pmatrix}$ and $V_{l-r,1} = \begin{pmatrix} I_{l-r-1} & \\ & V'_{l-r,1} \end{pmatrix}$; then,

$$U_{l-r,1}U_{l-r-1,1}\cdots U_{21}U_1\overline{F(z,w)}V_{11}V_{21}\cdots V_{l-r-1,1}V_{l-r,1} = \begin{pmatrix} \overline{d}_{11} & * & \cdots & & \\ & \overline{d}_{22} & * & \cdots & X \\ & & \ddots & & \\ & & & \overline{d}_{l-r,l-r} & \\ & 0_{r,l-r} & & & \overline{F_{l-r}(z,w)} \end{pmatrix} \bmod h(z). \tag{19}$$

Let

$$A(z,w) = \begin{pmatrix} \overline{d}_{11} & * & \cdots & & \\ & \overline{d}_{22} & * & \cdots & X \\ & & \ddots & & \\ & & & \overline{d}_{l-r,l-r} & \\ & 0_{r,l-r} & & & \overline{F_{l-r}(z,w)} \end{pmatrix}, \tag{20}$$

$$\overline{F_{l-r}(z,w)} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rr} \end{pmatrix},$$

combined with $h(z) | d_{l-r+1}(F(z,w))$; we have that $d_{l-r+1}(\overline{F(z,w)}) \equiv 0 \bmod h(z)$ and $d_{l-r+1}(A(z,w)) \equiv 0 \bmod h(z)$. Considering the $(l-r+1)\times(l-r+1)$ minors of $A(z,w)$, we see that $\overline{d}_{11}\overline{d}_{22}\cdots\overline{d}_{l-r,l-r}\cdot a_{k,j} \equiv 0 \bmod h(z)$ for all $k, j = 1, \ldots, r$. Since $\overline{d}_{i,i} \neq 0$, $i = 1, \ldots, l-r$, $a_{k,j} \equiv 0 \bmod h(z)$, $k, j = 1, \ldots, r$. Hence, $\overline{F_{l-r}(z,w)} \equiv 0_{r,r} \bmod h(z)$. Hence,

$$U_{l-r,1}U_{l-r-1,1}\cdots U_{21}U_1F(z,w)V_{11}V_{21}\cdots V_{l-r-1,1}V_{l-r,1} = A(z,w) = \begin{pmatrix} I_{l-r} & & & \\ & h(z) & & \\ & & \ddots & \\ & & & h(z) \end{pmatrix} Q_2, \tag{21}$$

where $Q_2 \in K^{l\times l}[z,w]$. Note that $\det F(z,w) = h^r(z)$ and U_1, U_{i1}, V_{i1} are unimodular matrices, $i = 1, 2, \ldots, l-r$; then, we obtain that $\det Q_2 = 1$ and Q_2 is a unimodular matrix. Let $U = U_{l-r,1}U_{l-r-1,1}\cdots U_1$ and $V = V_{11}V_{21}\cdots V_{l-r,1}Q_2^{-1}$; then, $F(z,w) = U^{-1}P(z)V^{-1}$ with U^{-1}, V^{-1} being unimodular matrices. Therefore, $F(z,w)$ is equivalent to the Smith form $P(z)$. □

Remark 1. Theorem 1 provides a positive answer to Problem 1.

Corollary 1. *Let $F(z,w) \in K^{l\times l}[z,w]$ and $h(z) \in K[z]$ be an irreducible polynomial. If $h(z) | d_{l-r+1}(F(z,w))$ and $h^r(z) | \det F(z,w)$, then $F(z,w)$ can be factorized as $T(z,w)P(z)Q(z,w)$, where $T(z,w), Q(z,w) \in K^{l\times l}[z,w]$ and $T(z,w)$ is unimodular.*

Theorem 2. *Let $F(z,w) \in K^{l\times l}[z,w]$ and $h(z) \in K[z]$ be an irreducible polynomial. Suppose $\det F(z,w) = h^{q\cdot r}(z)$, r and q are positive integers. If $h^q(z) | d_{l-r+1}(F(z,w))$, then $F(z,w)$ can be factorized as*

$$T_1(z,w)P(z)T_2(z,w)P(z)\cdots T_q(z,w)P(z)T(z,w), \tag{22}$$

where $P(z)$ is defined as above and $T(z,w), T_i(z,w)$ are unimodular matrices, $i = 1, 2, \ldots, q$.

Proof. When $q = 1$, by Theorem 1, $F(z,w) = T_1(z,w)P(z)T(z,w)$, where $T(z,w), T_1(z,w)$ are unimodular matrices, so the conclusion is true.

When $q \geq 2$, since $h^r(z)$ is a factor of $\det F(z,w)$, according to Corollary 1, we see that $F(z,w)$ can be factorized as $T_1(z,w)P(z)F_1(z,w)$, where $T_1(z,w)$ is a unimodular matrix.

Then, we consider $F_1(z,w)$. For $\det(F_1(z,w)) = h^{(q-1)\cdot r}(z)$, according to Lemma 1, we obtain

$$d_i(F(z,w)) = d_i\big(P(z)F_1(z,w)\big), \quad i = 1, \ldots, l. \tag{23}$$

So,

$$d_{l-r+1}(F(z,w)) = d_{l-r+1}\big(P(z)F_1(z,w)\big) = d_{l-r+1}(P(z))d_{l-r+1}\big(F_1(z,w)\big) = h(z)d_{l-r+1}\big(F_1(z,w)\big). \tag{24}$$

Note that $h^q(z)|d_{l-r+1}(F(z,w))$; then, $h^{q-1}(z)|d_{l-r+1}(F_1(z,w))$. According to Corollary 1, $F_1(z,w)$ can be factorized as $T_2(z,w)P(z)F_2(z,w)$. Then,

$$F(z,w) = T_1(z,w)P(z)T_2(z,w)P(z)F_2(z,w). \tag{25}$$

Imitating the procedure above successively, we can obtain that

$$F(z,w) = T_1(z,w)P(z)T_2(z,w)P(z)\cdots T_q(z,w)P(z)T(z,w), \tag{26}$$

where $T(z,w), T_i(z,w)$ are unimodular matrices, $i = 1, 2, \ldots, q$. □

Lemma 3. *Let $P(z)$ be defined as above and $T(z,w) \in K^{l\times l}[z,w]$ be unimodular. Suppose $F(z,w) = P^{s_1}(z)T(z,w)P^{s_2}(z)$. If the $(l-r)\times(l-r)$ minors of $F(z,w)$ have no common zero in $\overline{K}^2$, then $F(z,w)$ is equivalent to $P^{s_1+s_2}(z)$.*

Proof. Let

$$T(z,w) = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}, \tag{27}$$

where $T_{11} \in K^{(l-r)\times(l-r)}[z,w]$, $T_{12} \in K^{(l-r)\times r}[z,w]$, $T_{21} \in K^{r\times(l-r)}[z,w]$, and $T_{22} \in K^{r\times r}[z,w]$.

Note that

$$P(z) = \begin{pmatrix} I_{l-r} & \\ & h(z)I_r \end{pmatrix}. \tag{28}$$

Then,

$$F(z,w) = P^{s_1}(z)T(z,w)P^{s_2}(z) = \begin{pmatrix} T_{11} & T_{12}h^{s_2}(z) \\ T_{21}h^{s_1}(z) & T_{22}h^{s_1+s_2}(z) \end{pmatrix}. \tag{29}$$

By computing, $d_{l-r+1}(F(z,w)) = h^{s_1+s_2}(z)$ $d_{l-r+1}(T(z,w))$. Because $T(z,w)$ is unimodular and the $(l-r)\times(l-r)$ minors of $F(z,w)$ have no common zero,

$$\begin{aligned} &d_1(F(z,w)) = d_2(F(z,w)) = \cdots = d_{l-r}(F(z,w)) = 1, \\ &d_{l-r+k}(F(z,w)) = h^{k(s_1+s_2)}(z), \quad 1 \le k \le r. \end{aligned} \tag{30}$$

Since (T_{11}, T_{12}) is ZLP, the $(l-r)\times(l-r)$ minors of (T_{11}, T_{12}) have no common zero. Set the $(l-r)\times(l-r)$ minors of (T_{11}, T_{12}) to be $e_1, e_2, \ldots, e_t$, where $e_1 = \det(T_{11})$. Then, the $(l-r)\times(l-r)$ minors of $(T_{11}, T_{12}h^{s_2}(z))$ are $e_1, e_2h^{m_2}(z), \ldots, e_th^{m_t}(z)$, where $s_2 \le m_i \le r\cdot s_2$, $i = 2, \ldots, t$.

We prove that $(T_{11}, T_{12}h^{s_2}(z))$ is ZLP.

Suppose the $(l-r)\times(l-r)$ minors of $(T_{11}, T_{12}h^{s_2}(z))$ have a common zero a_0; then, a_0 is a common zero of e_1 and $h(z)$. Note that

$$F(z,w) = \begin{pmatrix} T_{11} & T_{12}h^{s_2}(z) \\ T_{21}h^{s_1}(z) & T_{22}h^{s_1+s_2}(z) \end{pmatrix}. \tag{31}$$

Then, the $(l-r)\times(l-r)$ minors of $F(z,w)$ have common zero a_0. This is a contradiction. So, the $(l-r)\times(l-r)$ minors of $(T_{11}, T_{12}h^{s_2}(z))$ have no common zero. Hence, $(T_{11}, T_{12}h^{s_2}(z))$ is ZLP. By the Quillen–Suslin theorem, we can find a unimodular matrix $M_1 \in K^{(l-r)\times(l-r)}[z,w]$ that satisfies

$$(T_{11}, T_{12}h(z))M_1 = \begin{pmatrix} I_{l-r} & 0_{l-r,r} \end{pmatrix}. \tag{32}$$

Then,

$$F(z,w)M_1 = \begin{pmatrix} I_{l-r} & 0_{l-r,r} \\ N_1 & N_2 \end{pmatrix}, \tag{33}$$

where $N_1 \in K^{r\times(l-r)}[z,w]$ and $N_2 \in K^{r\times r}[z,w]$. Let

$$M_2 = \begin{pmatrix} I_{l-r} & 0_{l-r,r} \\ -N_1 & I_r \end{pmatrix} \tag{34}$$

such that

$$M_2F(z,w)M_1 = \begin{pmatrix} I_{l-r} & 0_{l-r,r} \\ 0_{r,l-r} & N_2 \end{pmatrix}. \tag{35}$$

Let

$$M_3 = \begin{pmatrix} I_{l-r} & 0_{l-r,r} \\ 0_{r,l-r} & N_2 \end{pmatrix}. \tag{36}$$

We know M_3 is equivalent to $F(z,w)$.

According to Lemma 1, $d_i(M_3) = d_i(F(z,w))$, $i = 1, \ldots, l$. Hence,

$$\begin{aligned} &d_1(M_3) = d_2(M_3) = \cdots = d_{l-r}(M_3) = 1, \\ &d_{l-r+k}(M_3) = h^{k(s_1+s_2)}(z), \quad 1 \le k \le r. \end{aligned} \tag{37}$$

Notice that

$$\begin{aligned} M_3 &= \begin{pmatrix} I_{l-r} & 0_{l-r,r} \\ 0_{r,l-r} & N_2 \end{pmatrix} \\ &= \begin{pmatrix} I_{l-r} & & & & \\ & b_{l-r+1,l-r+1} & b_{l-r+1,l-r+2} & \cdots & b_{l-r+1,l} \\ & b_{l-r+2,l-r+1} & b_{l-r+2,l-r+2} & \cdots & b_{l-r+2,l} \\ & \vdots & \vdots & \ddots & \vdots \\ & b_{l,l-r+1} & b_{l,l-r+1} & \cdots & b_{l,l} \end{pmatrix}. \end{aligned} \tag{38}$$

We have that $h(z)$ divides every element of N_2, so

$$M_3 = \begin{pmatrix} I_{l-r} & & & \\ & h^{s_1+s_2}(z) & & \\ & & \ddots & \\ & & & h^{s_1+s_2}(z) \end{pmatrix} M_4. \tag{39}$$

Hence, $M_3 = P^{s_1+s_2}(z) \cdot M_4$, and $\det M_4 = 1$. Note that $M_2 \cdot F(z,w) \cdot M_1 = M_3$, so $M_2 \cdot F(z,w) \cdot M_1 = P^{s_1+s_2}(z) \cdot M_4$; then, $F(z,w) = M_2^{-1} P^{s_1+s_2}(z) \cdot M_4 M_1^{-1}$, where M_2^{-1} and $M_4 M_1^{-1}$ are unimodular matrices. Thus, $F(z,w)$ is equivalent to $P^{s_1+s_2}(z)$. □

Theorem 3. *Let $F(z,w) \in K^{l\times l}[z,w]$ and $h(z) \in K[z]$ be an irreducible polynomial. If $\det F(z,w) = h^{q\cdot r}(z)$ and $h^q(z) | d_{l-r+1}(F(z,w))$, where q and r are positive integers, then $F(z,w)$ is equivalent to the Smith form*

$$P^q(z) = \begin{pmatrix} I_{l-r} & & & \\ & h^q(z) & & \\ & & \ddots & \\ & & & h^q(z) \end{pmatrix}, \tag{40}$$

if and only if the $(l-r)\times(l-r)$ minors of $F(z,w)$ have no common zero in $\overline{K}^2$.

Proof. Sufficiency: by Theorem 2, we have that

$$F(z,w) = T_1(z,w)P(z)T_2(z,w)P(z)\cdots T_q(z,w)P(z)T(z,w), \tag{41}$$

where

$$P(z) = \begin{pmatrix} I_{l-r} & & & \\ & h(z) & & \\ & & \ddots & \\ & & & h(z) \end{pmatrix}. \tag{42}$$

$T(z,w), T_i(z,w)$ are unimodular matrices, $i = 1, 2, \ldots, q$.

We first prove that $P(z)T_2(z,w)P(z)$ is equivalent to $P^2(z)$. From the $(l-r)\times(l-r)$ minors of $F(z,w)$ which have no common zero in $\overline{K}^2$, we know that the $(l-r)\times(l-r)$ minors of $P(z)T_2(z,w)P(z)$ have no common zero in $\overline{K}^2$ by combining with Lemma 2. Then, according to Lemma 3, $P(z)T_2(z,w)P(z)$ is equivalent to $P^2(z)$.

Repeating the procedure above, we obtain that $P(z)T_2(z,w)P(z)\cdots T_q(z,w)P(z)$ is equivalent to the matrix $P^q(z)$, so there exist $U_1(z,w), V_1(z,w)$ such that

$$P(z)T_2(z,w)P(z)\cdots T_q(z,w)P(z) = U_1(z,w)P^q(z)V_1(z,w), \tag{43}$$

where $U_1(z,w), V_1(z,w)$ are unimodular matrices. Hence, we have that

$$F(z,w) = U(z,w)P^q(z)V(z,w), \tag{44}$$

where $U(z,w) = T_1(z,w)U_1(z,w)$ and $V(z,w) = V_1(z,w)T(z,w)$ are unimodular matrices. Therefore, $F(z,w)$ is equivalent to the Smith form $P^q(z)$.

Necessity: since $F(z,w)$ is equivalent to $P^q(z)$, there exist unimodular matrices $U(z,w)$ and $V(z,w)$ such that $F(z,w) = U(z,w)P^q(z)V(z,w)$. For the $(l-r)\times(l-r)$ minors of $P^q(z)$ which have no common zero in $\overline{K}^2$, the $(l-r)\times(l-r)$ minors of $F(z,w)$ have no common zero in $\overline{K}^2$ by Lemma 2. □

Remark 2. According to Theorem 3, Problem 2 is solved, and a criterion for discriminating this class of bivariate polynomial matrices to be equivalent to the Smith form is also presented.

Theorem 4. *Let $F(z,w) \in K^{l\times m}[z,w]$ $(l \le m)$ be of full row rank and $h(z) \in K[z]$ be irreducible. Suppose $d_l(F) = h^{q\cdot r}(z)$ and $h^q(z) | d_{l-r+1}(F)$, where q and r are positive integers. Then, $F(z,w)$ is equivalent to the Smith form*

$$Q(z) = \begin{pmatrix} P^q(z) & 0_{l\times(m-l)} \end{pmatrix}, \tag{45}$$

if and only if the $(l-r)\times(l-r)$ minors of $F(z,w)$ have no common zero in $\overline{K}^2$.

Proof. Sufficiency: by Theorem 3.3 in [17], there are $H(z,w) \in K^{l\times l}[z,w]$ and $F_1(z,w) \in K^{l\times m}[z,w]$ satisfying $F(z,w) = H(z,w)F_1(z,w)$, where $\det H(z,w) = d(F)$ and $F_1(z,w)$ is ZLP. Then, the $(l-r)\times(l-r)$ minors of $H(z,w)$ have no common zero by using Lemma 2. Combined with $h^q(z) | d_{l-r+1}(H)$ and Theorem 2, there are unimodular matrices $U_{11}(z,w), U_{12}(z,w) \in K^{l\times l}[z,w]$ such that $H(z,w) = U_{11}(z,w)P^q(z)U_{12}(z,w)$. Then,

$$F(z,w) = U_{11}(z,w)P^q(z)U_{12}(z,w)F_1(z,w). \tag{46}$$

We know that $U_{12}(z,w)F_1(z,w)$ is also ZLP ($U_{12}(z,w)$ is unimodular). According to the Quillen-Suslin theorem, there exists an $m\times m$ unimodular matrix $U(z,w)$ which satisfies that $U_{12}(z,w)F_1(z,w)U(z,w) = \begin{pmatrix} I_l & 0_{l,m-l} \end{pmatrix}$. Then,

$$F(z,w) = U_{11}(z,w)\begin{pmatrix} P^q(z) & 0_{l,m-l} \end{pmatrix} U(z,w) = U_1(z,w)Q(z)U(z,w). \tag{47}$$

According to the $(l-r)\times(l-r)$ minors of $F(z,w)$ which have no common zero and $h^q(z)|d_{l-r+1}(F)$, we see that $Q(z)$ is the Smith form of $F(z,w)$. Note that $U_1(z,w)$ and $U(z,w)$ are invertible matrices, so $F(z,w)$ is equivalent to the Smith form $Q(z)$.

Necessity: since $F(z,w)$ is equivalent to the Smith form $Q(z)$, it is easy to see that the $(l-r)\times(l-r)$ minors of $Q(z)$ have no common zero in $\overline{K}^2$. By Lemma 2, we have that the $(l-r)\times(l-r)$ minors of $F(z)$ have no common zero in $\overline{K}^2$. □

Remark 3. A positive answer to Problem 3 is presented in Theorem 4. And the equivalence of a rectangle matrix $F(z,w)$ and its Smith form is considered, which makes the result more general.

4. An Example

In this section, we use an example to illustrate our results and methods.

Example 1. Consider a 3×3 2D polynomial matrix of $\mathbb{R}^{3\times3}[z,w]$:

$$F(z,w) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \tag{48}$$

where

$$\begin{aligned}
a_{11} &= \left(z+1-w^2\right)(1-2w) - \left(2z^2+2zw+2\right)\left(z^2+z+1\right)^2, \\
a_{12} &= \left(z+1-w^2\right)(z-w) + \left[(zw+1-z)(w+z)+w-1\right]\left(z^2+z+1\right)^2, \\
a_{13} &= w\left(z+1-w^2\right) + \left(zw+z^2+1\right)\left(z^2+z+1\right)^2, \\
a_{21} &= -w(1-2w) - 2z\left(z^2+z+1\right)^2, \\
a_{22} &= -w(z-w) + (zw+1-z)\left(z^2+z+1\right)^2, \\
a_{23} &= -w^2 + z\left(z^2+z+1\right)^2, \\
a_{31} &= z(1-2w) - \left(2z^2+2\right)\left(z^2+z+1\right)^2, \\
a_{32} &= z(z-w) + \left[z(zw+1-z)+w-1\right]\left(z^2+z+1\right)^2, \\
a_{33} &= zw + \left(z^2+1\right)\left(z^2+z+1\right)^2.
\end{aligned} \tag{49}$$

By computing, $d_1(F(z,w)) = 1$, $d_2(F(z,w)) = (z^2+z+1)^2$, $d_3(F(z,w)) = (z^2+z+1)^4$, and $\det F(z,w) = (z^2+z+1)^4$. We know that the 1×1 minors of $F(z,w)$ have no common zero in $\mathbb{C}^2$. Combining $z^2+z+1 \in K[z]$ is irreducible and $(z^2+z+1)^2|d_2(F(z,w))$; by Theorem 3, we have that $F(z,w)$ is equivalent to the Smith form

$$P^2(z) = \begin{pmatrix} 1 & & \\ & \left(z^2+z+1\right)^2 & \\ & & \left(z^2+z+1\right)^2 \end{pmatrix}. \tag{50}$$

Let $\overline{F(z,w)} = (\overline{f_{ij}})$ denote $F(z,w) = (f_{ij})$ mod (z^2+z+1); then,

$$\overline{F(z,w)} = \begin{pmatrix} \left(z+1-w^2\right)(1-2w) & \left(z+1-w^2\right)(z-w) & w\left(z+1-w^2\right) \\ 2w^2-w & w^2-zw & -w^2 \\ -2zw+z & -zw-z-1 & zw \end{pmatrix}, \tag{51}$$

and no row of $\overline{F(z,w)}$ is zero vector mod (z^2+z+1). Note that the nonzero polynomial of the least degree in w among the first column is $-2zw+z$, so we postmultiply $\overline{F(z,w)}$ by a unimodular matrix $U_{11}=\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}$ such that $-2zw+z$ can be changed to the position (1, 1). Then,

$$U_{11}\overline{F(z,w)}=\begin{pmatrix}-2zw+z & -zw-z-1 & zw\\ 2w^2-w & w^2-zw & -w^2\\ \left(z+1-w^2\right)(1-2w) & \left(z+1-w^2\right)(z-w) & w\left(z+1-w^2\right)\end{pmatrix}. \tag{52}$$

We consider the element $-2zw+z$; $-2z$ is the leading coefficient of $-2zw+z$. We know that $-2z$ and z^2+z+1 are relatively prime in $\mathbb{R}[z]$, so we can find $x_1(z)=(1/2)(z+1)$, $y_1(z)=1$ such that $x_1(z)\cdot(-2z)+y_1(z)\cdot(z^2+z+1)=1$ $((1/2)(z+1)\cdot(-2z)+(z^2+z+1)=1)$; then, $(1/2)(z+1)\cdot(-2z)=1-(z^2+z+1)\equiv 1 \bmod (z^2+z+1)$. So, $x_1(z)\cdot(-2zw+z)=(1/2)(z+1)\cdot(-2zw+z)\equiv(w-(1/2)) \bmod (z^2+z+1)$; that is, $-2zw+z$ can be reduced to a monic polynomial $(w-(1/2))$ by mod (z^2+z+1).

Then, we reduce other elements in the first column.

For the element $2w^2-w$, we can find $q_1(z,w)=2w$ and $r_1(z,w)=0$ such that $2w^2-w=(w-(1/2))\cdot q_1(z,w)+r_1(z,w)\equiv(1/2)(z+1)\cdot(-2zw+z)\cdot 2w+0 \bmod (z^2+z+1)$. In reality, $2w^2-w=(1/2)(z+1)\cdot(-2zw+z)\cdot 2w-(z^2+z+1)(w-2w^2)$.

And for the element $(z+1-w^2)(1-2w)$, we can find $q_2(z,w)=2w^2-2z-2$ and $r_2(z,w)=0$ such that

$$\begin{aligned}\left(z+1-w^2\right)(1-2w)&=\left(w-\frac{1}{2}\right)\cdot q_2(z,w)+r_2(z,w)=\left(w-\frac{1}{2}\right)\cdot\left(2w^2-2z-2\right)+\\ 0&\equiv\frac{1}{2}(z+1)\cdot(-2zw+z)\cdot\left(2w^2-2z-2\right)+0 \bmod\left(z^2+z+1\right).\end{aligned} \tag{53}$$

Actually, $(z+1-w^2)(1-2w)=(1/2)(z+1)\cdot(-2zw+z)\cdot(2w^2-2z-2)-(z^2+z+1)(w^2-z-1)(1-2w)$.

Let

$$U_{12}=\begin{pmatrix}1 & 0 & 0\\ -\frac{1}{2}(z+1)2w & 1 & 0\\ -\frac{1}{2}(z+1)\left(2w^2-2z-2\right) & 0 & 1\end{pmatrix}. \tag{54}$$

We have

$$U_{12}U_{11}\cdot\overline{F(z,w)}=\begin{pmatrix}-2zw+z & -zw-z-1 & zw\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\bmod\left(z^2+z+1\right), \tag{55}$$

so

$$U_{12}U_{11}\cdot F(z,w)=\begin{pmatrix}1 & & \\ & z^2+z+1 & \\ & & z^2+z+1\end{pmatrix}F_1(z,w), \tag{56}$$

where

$$\begin{aligned}
F_1(z,w)&=\begin{pmatrix}b_{11} & b_{12} & b_{13}\\ b_{21} & b_{22} & b_{23}\\ b_{31} & b_{32} & b_{33}\end{pmatrix},\\
b_{11}&=z(1-2w)-\left(2z^2+2\right)\left(z^2+z+1\right)^2,\\
b_{12}&=z(z-w)+[z(zw+1-z)+w-1]\left(z^2+z+1\right)^2,\\
b_{13}&=zw+\left(z^2+1\right)\left(z^2+z+1\right)^2,\\
b_{21}&=2w^2-w+\left[-2z+(zw+w)\left(2z^2+2\right)\right]\left(z^2+z+1\right),\\
b_{22}&=\left(w^2-zw\right)-(zw+w)[z(zw+1-z)+w-1]\left(z^2+z+1\right)+(zw+1-z)\left(z^2+z+1\right),\\
b_{23}&=-w^2+\left[z-w(z+1)\left(z^2+1\right)\right]\left(z^2+z+1\right),\\
b_{31}&=(1-2w)\left(-w^2+1+z\right)+\left[-2z^2-2zw-2+(z+1)\left(-w^2+1+z\right)\left(-2z^2-2\right)\right]\left(z^2+z+1\right),\\
b_{32}&=\left(-w^2+1+z\right)(z-w)+[(zw+1-z)(w+z)+w-1]\left(z^2+z+1\right)+(z+1)\left(-w^2+1+z\right)\\
&\quad [z(zw+1-z)+w-1]\left(z^2+z+1\right),\\
b_{33}&=\left(-w^2+1+z\right)w+\left(z^2+zw+1\right)\left(z^2+z+1\right)+(z+1)\left(-w^2+z+1\right)\left(z^2+1\right)\left(z^2+z+1\right).
\end{aligned} \tag{57}$$

Let $U_1=U_{12}U_{11}$; then,

$$\begin{aligned}
F(z,w)&=U_1^{-1}\begin{pmatrix}1 & & \\ & z^2+z+1 & \\ & & z^2+z+1\end{pmatrix}F_1(z,w),\\
\overline{F_1(z,w)}&=\begin{pmatrix}-2zw+z & -zw-z-1 & zw\\ 2w^2-w & w^2-zw & -w^2\\ (1-2w)\left(-w^2+1+z\right) & \left(-w^2+z+1\right)(z-w) & w\left(-w^2+z+1\right)\end{pmatrix}.
\end{aligned} \tag{58}$$

Now, consider $\overline{F_1(z,w)}$, for none of its rows are zero vectors mod (z^2+z+1); repeating the steps above, we can obtain that

$$U_{13}\cdot\overline{F_1(z,w)}=\begin{pmatrix}-2zw+z & -zw-z-1 & zw\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\mathrm{mod}\left(z^2+z+1\right), \tag{59}$$

where

$$U_{13} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2}(z+1)2w & 1 & 0 \\ -\frac{1}{2}(z+1)\left(2w^2 - 2z - 2\right) & 0 & 1 \end{pmatrix}. \tag{60}$$

Hence,

$$U_{13}F_1(z,w) = \begin{pmatrix} 1 & & \\ & z^2+z+1 & \\ & & z^2+z+1 \end{pmatrix} F_2(z,w), \tag{61}$$

where

$$\begin{aligned} F_2(z,w) &= \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}, \\ c_{11} &= z(1-2w) - \left(2z^2+2\right)\left(z^2+z+1\right)^2, \\ c_{12} &= z(z-w) + [z(zw+1-z)+w-1]\left(z^2+z+1\right)^2, \\ c_{13} &= zw + \left(z^2+1\right)\left(z^2+z+1\right)^2, \\ c_{21} &= 2w^2 - w - 2z + (zw+w)\left(2z^2+2\right)\left(z^2+z+2\right), \\ c_{22} &= \left(w^2+1-z\right) - (zw+w)\left(z^2w+z-z^2\right)\left(z^2+z+2\right) - \left(w^2-w\right)(z+1)\left(z^2+z+2\right), \\ c_{23} &= z - w^2 - w(z+1)\left(z^2+1\right)\left(z^2+z+2\right), \\ c_{31} &= (1-2w)\left(-w^2+1+z\right) - \left(2z^2+2zw+2\right) + (z+1)\left(-w^2+1+z\right)\left(-2z^2-2\right)\left(z^2+z+2\right), \\ c_{32} &= \left(-w^2+1+z\right)(z-w) + (zw+1-z)(w+z) + w - 1 + \left(z^2+z+2\right)(z+1)\left(-w^2+1+z\right) \\ &\quad (w-1) + \left(z^2+z\right)\left(-w^2+1+z\right)(zw+1-z)\left(z^2+z+2\right), \\ c_{33} &= \left(-w^2+1+z\right)w + \left(z^2+zw+1\right) + \left(z^2+z+2\right)(z+1)\left(-w^2+z+1\right)\left(z^2+1\right). \end{aligned} \tag{62}$$

So,

$$F_1(z,w) = U_{13}^{-1}\begin{pmatrix} 1 & & \\ & z^2+z+1 & \\ & & z^2+z+1 \end{pmatrix} F_2(z,w). \tag{63}$$

Combining

$$F(z,w) = U_1^{-1}P(z)F_1(z,w), \tag{64}$$

we have that

$$F(z,w) = U_1^{-1}P(z)U_{13}^{-1}P(z)F_2(z,w). \tag{65}$$

Consider $P(z)U_{13}^{-1}P(z)$; combined with Lemma 3, we obtain

$$P(z)U_{13}^{-1}P(z) = \begin{pmatrix} 1 & 0 & 0 \\ (wz+w)\left(z^2+z+1\right) & \left(z^2+z+1\right)^2 & 0 \\ -(z+1)\left(-w^2+1+z\right)\left(z^2+z+1\right) & 0 & \left(z^2+z+1\right)^2 \end{pmatrix}. \tag{66}$$

Let

$$U_{14} = \begin{pmatrix} 1 & 0 & 0 \\ (wz+w)\left(z^2+z+1\right) & 1 & 0 \\ -(z+1)\left(-w^2+1+z\right)\left(z^2+z+1\right) & 0 & 1 \end{pmatrix}. \quad (67)$$

Then,

$$P(z)U_{13}^{-1}P(z) = U_{14}P^2(z). \quad (68)$$

So,

$$F(z,w) = U_1^{-1}U_{14}\begin{pmatrix} 1 & & \\ & \left(z^2+z+1\right)^2 & \\ & & \left(z^2+z+1\right)^2 \end{pmatrix}F_2(z,w), \quad (69)$$

and by computing, $U_1^{-1}U_{14}$ and $F_2(z,w)$ are unimodular matrices. Therefore, $F(z,w)$ is equivalent to $P^2(z)$.

5. Conclusions

In this paper, we have investigated the reduction of several kinds of bivariate polynomial matrices in $K[z,w]$, where K is an arbitrary field. Some new results on these matrices to be reduced to their Smith forms are presented. Furthermore, the conditions of these results are easily verified. An example is given to illustrate our method in the end of the article. All of these are useful for reducing 2D systems.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (11871207 and 11971161).

References

[1] N. Bose, *Applied Multidimensional Systems Theory*, Van Nostrand Reinhold, Newyork, NY, USA, 1982.

[2] N. Bose, B. Buchberger, and J. Guiver, *Multidimensional Systems Theory and Applications*, Kluwer, Dordrecht, Netherlands, 2003.

[3] Y. Liu and H. Su, "Some necessary and sufficient conditions for containment of second-order multi-agent systems with sampled position data," *Neurocomputing*, vol. 378, pp. 228–237, 2020.

[4] Y. F. Liu and H. S. Su, "Containment control of second-order multi-agent systems via intermittent sampled position data communication," *Applied Mathematics and Computation*, vol. 362, Article ID 124522, 2019.

[5] C. Xu, H. Xu, H. Su, and C. Liu, "Disturbance-observer based consensus of linear multi-agent systems with exogenous disturbance under intermittent communication," *Neurocomputing*, vol. 404, pp. 26–33, 2020.

[6] H. Rosenbrock, *State Space and Multivariate Theory*, Nelson-Wiley, New York, NY, USA, 1970.

[7] M. Frost and C. Storey, "Euivalence of a matrix over R[s,z] with its Smith form," *International Journal of Control*, vol. 28, no. 5, pp. 665–671, 1979.

[8] E. Lee and S. Zak, "Smith forms over R[z_1, z_2]," *IEEE Transactions on Automatic Control*, vol. 28, no. 1, pp. 115–118, 1983.

[9] M. G. Frost and M. S. Boudellioua, "Some further results concerning matrices with elements in a polynomial ring," *International Journal of Control*, vol. 43, no. 5, pp. 1543–1555, 1986.

[10] M. S. Boudellioua, "Computation of the Smith form for multivariate polynomial matrices using maple," *American Journal of Computational Mathematics*, vol. 02, no. 01, pp. 21–26, 2012.

[11] D. Li, R. Liang, and J. Liu, "Some further results on the Smith form of bivariate polynomial matrices," *Journal of Systems Science and Mathematical Sciences*, vol. 39, no. 12, pp. 1983–1997, 2019.

[12] D. Li, J. Liu, and L. Zheng, "A zero-dimensional valuation ring is 1-Gröbner," *Journal of Algebra*, vol. 484, pp. 334–343, 2017.

[13] Z. Lin, M. Boudellioua, and L. Xu, "On the equivalence and factorization of multivariate polynomial matrices," in *Proceedings of the IEEE*, pp. 4911–4914, Tampa FL, USA, November 2006.

[14] D. Li, J. Liu, and L. Zheng, "On the equivalence of multivariate polynomial matrices," *Multidimensional Systems and Signal Processing*, vol. 28, no. 1, pp. 225–235, 2017.

[15] D. Quillen, "Projective modules over polynomial rings," *Inventiones Mathematicae*, vol. 36, no. 1, pp. 167–171, 1976.

[16] D. Li, J. Liu, and L. Zheng, "On serre reduction of multidimensional systems," *Mathematical Problems in Engineering*, vol. 2020, Article ID 7435237, 2020.

[17] M. Wang and D. Feng, "On Lin-Bose problem," *Linear Algebra and Its Applications*, vol. 390, pp. 279–285, 2004.

[18] M. Boudellioua, "Further results on the equivalence to Smith form of multivariate polynomial matrices," *Control and Cybernetics*, vol. 42, no. 2, pp. 543–551, 2013.

[19] D. Li and R. Liang, "Serres reduction and the Smith forms of multivariate polynomial matrices," *Mathematical Problems in Engineering*, vol. 2020, Article ID 5430842, 2020.

Hindawi
Complexity
Volume 2021, Article ID 1258730, 9 pages
https://doi.org/10.1155/2021/1258730

Research Article

Cooperation Mechanism in Blockchain by Evolutionary Game Theory

Jinxin Zhang and Meng Wu

School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China

Correspondence should be addressed to Meng Wu; wum@njupt.edu.cn

Received 9 September 2021; Accepted 20 October 2021; Published 8 November 2021

Academic Editor: Hou-Sheng Su

In the blockchain network, to get rewards in the blockchain, blockchain participants pay for various forms of competition such as computing power, stakes, and other resources. Because of the need to pay a certain cost, individual participants cooperate to maintain the long-term stability of the blockchain jointly. In the course of such competition, the game between each other has appeared invisibly. To better understand the blockchain design of cooperation mechanisms, in this paper, we constructed a game framework between participants with different willingness, using evolutionary game theory, and complex network games. We analyzed how the behavior of participants potentially develops with cost and payoff. We consider the expected benefits of participants for the normal growth of the blockchain as the major factor. Considering the behavior of malicious betrayers, the blockchain needs to be maintained in the early stage. Numerical simulation supports our analysis.

1. Introduction

The public blockchain is a peer-to-peer (P2P) network that is used to build a decentralized ledger and can be trusted to achieve autonomy without third-party supervision. The most well known of all blockchains is Bitcoin, in which the Nakamoto consensus protocol [1] creatively integrates a variety of mechanisms and technologies, laying the foundation for the subsequent development of the blockchain. Blockchain technology has the characteristics of decentralization, traceability, non-tampering, programmability, and collective maintenance.

Blockchain is a time-ordered one-in-one-out directed graph, interconnected to form a chain structure that uses cryptography to ensure that the distributed ledger is tamper-evident and unforgeable. In a broad sense, blockchain technology is a new distributed infrastructure and computing paradigm. It uses a blockchain data structure to store data, decentralized node consensus algorithms to generate and verify data, cryptography to secure data transmission and access, and smart contracts composed of automated scripting code to use data.

Besides applying some sophisticated technologies, the blockchain also considers some mechanisms. In the blockchain represented by Bitcoin, the entire nodes (block miners) are economically encouraged to abide by the "longest chain policy" to preserve the stability of the blockchain. Most blockchains follow this policy. The block bookkeeper packs any set of verified transactions into a data structure, announced a candidate "block," and broadcasts it to the entire network. Each blockchain participant consistently uses the longest chain as its local blockchain copy. This design can ensure that we can write the new blocks with the most resources into the chain in time.

Many blockchains also used bitcoin's incentive mechanism. It comprises two parts: effort and reward.

(i) *Effort*. The process of solving intensive problems makes Sybil attacks [2] economically unbearable

(ii) *Reward*. The reward given to the new block producer (miner) will be released after the network admits the block it publishes

Taking Bitcoin as an example, the hash puzzle-solving process is regarded as a computing power competition among miners. After investing in a certain amount of effort, the miners will get two types of rewards during the mining process: new currency rewards for creating new blocks and transaction fees included in the block. To get these rewards, miners are vying to complete a mathematical puzzle based on a cryptographic hash algorithm, which requires powerful computing power. In addition, each transaction may include a transaction fee. Miners who successfully "dig out" a new block during the mining process can receive "tips" for all transactions recorded in the block. The competition of this algorithm and the mechanism by which the winner has the all right to record transactions on the blockchain provide the safety of Bitcoin. Blockchain is a competitive platform for users. The anonymity of its participants and the transparency of network information create innate conditions for pursuing a balanced game. And the periodic expansion of new blocks in the blockchain provides an appropriate arena for repeated games.

Some scholars have applied game models to blockchain issues. How many transactions a new block contains is the miner's choice. Houy [3] defined the Bitcoin mining game between miners, assuming that the number of transactions in a block results from the game, and studied the Nash equilibrium of the game. From the perspective of game theory, Schrijvers et al. [4] defined the mining pool payment function in a single mine pool and analyzed whether several mining pool strategies meet the requirements of incentive compatibility, proportional payment, and budget balance. Eyal discussed the attack between the two mining pools [5], there is a contradiction between individual rationality and collective rationality between the pools, which is like the tragedy game of the commons. He believes that the interception attack reduces the computing power of the entire network.

Security issues in the blockchain have always been the concentrate of attention. 51% attack, selfish mining attack [6], and other threats always exist. The focus of this article is the impact of the strategic choice of each blockchain participant on the cooperation mechanism. We consider individual miners have bounded rationality and will adopt specific participation strategies in the blockchain. We model the strategy choice in the network as an evolutionary game and study the evolutionary stability of the dynamics of strategy selection.

The contributions of this paper are threefold. (1) We formulate the cooperative behavior on the blockchain. (2) We use evolutionary game theory to analyze how the participants of the blockchain change their strategies and their incentives. (3) We also analyzed the influence of network structure on cooperative behavior on the blockchain.

We organized this paper as follows. In Section 2, we review the background on security issues in the blockchain and evolutionary game theory. We formalize the new block generation right game and analyze it in Section 3. In Section 4, we assign values to the parameters in the model and conduct simulations and discuss the results. Finally, we conclude with future directions in Section 5.

2. Background

2.1. The Relationship between Cooperation, Trust, and Security. Cooperation, trust, and security are three concepts that are connected with each other and have their emphasis.

Different security scenarios, relationship types, and internal interactions within the system require different levels of trust. Trust depends on the restraint of the interacting parties, which is usually determined by agreement in the system. Since trust involves both risks and benefits, the need for trust is different for cooperation. Sztompka [7] believes that the simplest definition of trust is the gamble of believing in the future actions of others. Gambetta [8] emphasized that trust required depends on the strength of the mechanism that determines cooperative decision-making, and it also depends on the social context in which we decide and treats trust as "a result of cooperation."

It does not base the basis of cooperation on pure trust, but on the continuity of the relationship. Axelrod and Hamilton [9] believe that, in the long run, whether the conditions for the establishment of a stable cooperation model between the two parties are mature is more important than whether they trust each other. In order for the cooperation to be stable, there must be a large enough incentive in the future. It requires that the participants in the system have a large enough chance to meet again, and the meaning of their encounter again cannot be discounted too much.

Lamsal believes that trust and security are tightly coupled [10]. There are various ways to provide security services. However, encryption is one of the most extensive and effective ways to provide these services. But the management of cryptographic keys becomes a problem. The fundamental problem that any key management system is trying to solve: Is the key trusted?

Trust and security go hand in hand, and trustworthiness is the foundation of security. A trusted mechanism cannot accomplish any security function alone but can be implemented in various systems and work with the security mechanism of the system to play the role of reinforcing and managing security. But traditional security measures cannot really cover the smart environment, so trust is important. "6G White Paper: Research Challenges for Trust, Security and Privacy" [11] pointed out that the 6G network must support embedded trust for an increased level of information security in 6G. Trust modeling, trust policies, and trust mechanisms need to be defined.

In the modern network environment, users do not feel what kind of information has been collected, and they do not know where and who uses this information. With the gradual popularity of such application scenarios, the concept of trust has recently approached the concept of security and is closer to the needs of users. As the cooperative relationship between the participants in the network advances, the trust relationship is established unconsciously.

2.2. Security Issues in Blockchain. Blockchain uses some mature cryptographic principles and consensus mechanisms based on distributed ledger technology to realize trusted interaction without the intervention of a third party in a distributed environment. The improvement of system efficiency and cost reduction brought by blockchain provides new methods for economic and social development and various industries. Blockchain can build trust in a low-cost way. Blockchain has developed rapidly recently, and scholars have tried to use blockchain in fields such as medical care [12], edge computing [13], and the Internet of Things [14, 15].

These application areas involve the following features of blockchain technology.

(i) Decentralization. The blockchain network is based on a peer-to-peer (P2P) network to create information broadcasts, and the failure of a single node has no global impact.

(ii) Untamperable. All nodes maintain the ledger in the blockchain, and the consensus mechanism ensures that most nodes allow the ledger and provides a strong correlation between data blocks with a hash function, and the data on the chain are open to the entire network, and the data are verifiable and traceable.

(iii) Anonymity. Cryptography and other technologies separate the user's real identity from the network identity so that the user's identity can be protected and establish a foundation of trust at the same time.

In a blockchain network without third-party supervision, the essence of security is the confrontation between honest participants and malicious betrayers. Traditional network security threats will also exist in P2P network; using identity authentication, intrusion detection, traffic monitoring, etc., plays a significant role in network security protection. However, defense often lags attacks, and it is intractable to respond to more complex attacks and defenses in the network promptly. With the risk of more complex attacks, it is also a major challenge to coordinate the selection of defense strategies and generate attack responses before the attackers have caused severe losses. Besides the traditional security, the special scenario of the blockchain network takes unique challenges to the network security. This is because traditional data-based network security research methods are problematic to handle with blockchain network attacks that have real monetary benefits.

The blockchain suffers the following security threats:

(i) *51% Attack.* The 51% attack is an attack method that uses computing power as a competitive condition and controls the advantage of computing power to cancel payment transactions that have occurred. Blockchain networks such as Bitcoin have encryption security because of decentralization. After most nodes on the Bitcoin network agree, the public ledger is credible. However, if a single entity or team can control over 50% of the hash rate, then they can change the state of the blockchain. It can get the random number needed to mine the block faster than others and decide the packaged transaction and affect the effectiveness of the block. After launching the attack, the attacker will continuously destroy the value of a cryptocurrency by implementing double-spending or posting various fraudulent transfers on the chain.

(ii) *Selfish Mining.* A selfish mining attack is a method against Bitcoin mining and the incentive mechanism. Its purpose is not to destroy the operating mechanism of Bitcoin, but to pursue additional rewards and make honest miners perform invalid calculations.

In short, the core idea of a selfish mining attack is that malicious mining pools deliberately delay publishing new blocks worked out by them, and create a private branch under their control, making the chain to fork. Honest miners continue to mine based on public branches, while malicious mining pools mine based on private branches controlled by them. If the malicious mining pool calculates more blocks, the length of the private branch they control is naturally ahead of the public branch. The malicious mining pool chooses not to rush to publish these new blocks and strives to further increase the mining revenue. However, because of the limitation of the computing power of the malicious mining pool, the length advantage of the private branch will not be maintained consistently. When the public branch is close to the length of the private branch, the malicious mining pool will announce the branch generated by itself to replace the legal one and take the reward of these new blocks.

(iii) *Eclipse Attacks.* The method of eclipse attacks [16] is different from 51% attack. They attack specific nodes or node clusters so that the network appears "partitioned" to achieve double-spending and other purposes. Due to network spare capacity and computing power distribution limitations, the blockchain usually limits the upper limit of a single node that can receive information and actively link to other nodes. The concentrate of an eclipse attack is to start with the information of the nodes linked to and received. If the node receiving information and the externally linked nodes are all controlled by malicious nodes, it is equivalent that the malicious nodes isolate and control the received information to the target node. Here, it is called the node that suffered an "eclipse attack." If the malicious nodes can control more nodes and launch an eclipse attack on more normal nodes, then the malicious nodes will split the blockchain network into two distinct partitions, just like a fork.

2.3. Application of Evolutionary Game Theory in Blockchain. In game theory, players can adopt several strategies, have a certain influence on other players, and get a certain payoff in the interaction. Game theory is the study of rational choice of strategies between individuals to pursue the maximization of their interests. The core concept is Nash

equilibrium. Daskalakis et al. [17] conclude that it is almost impossible to find Nash equilibrium on a large scale in action-graph games. They proved it belongs to a subset of the NP problem. But when the participants are anonymous, they need only polynomial time to approach the Nash equilibrium.

The evolutionary game theory concentrates on how individuals with bounded rationality can continuously optimize their payoffs by comparing in the repeated game. Similar to the concept of fitness in the evolution of nature, players with a high payoff are more likely to take part in the next round of the game, while a play with a low payoff is more likely to be eliminated. Eventually, a particular strategy will reach equilibrium in the system.

The blockchain is a natural testing ground in this respect because of the following characteristics:

(1) Based on the decentralized P2P communication method, information symmetry becomes workable. In the blockchain network, all individuals share the same database, and all information releases appear as broadcast. Then, not only are the two parties in the transaction holding information, but all communicating parties have corresponding information. This has completely changed the limitations of information asymmetry between the two parties and created conditions for joint decision-making.

(2) The status of participants in the blockchain is equal. Besides the alliance chain, which is between the public chain and the private chain, the rights and status of the nodes in an individual blockchain are equal. It enhances the autonomy of decision-making. In the game, the inequality of social status is prone to cause a party to have the dominant advantage, making the result more predisposed to the dominant party. The blockchain has changed this state and created a nearly free environment, making it easier to analyze the results in a rational decision-making manner. With the blessing of blockchain technology, both parties to the transaction have complete transaction data, and each transaction subject can make independent decisions based on their knowledge, and there is almost no distinction in transaction rights.

(3) Cryptography provides a credible environment. All participants and their information on the blockchain are digital. On the premise that it does not compromise encryption technology, transaction data is irreversible and cannot be denied. Participants in the transaction process are more prone to trust each other.

(4) Generating new blocks with no curb on the rounds creates conditions for repeated games. We can consider the processing of each new block as a repeated game, and participants do not demand to reveal their private keys and remain anonymous. As the chain continues to grow, it reaches the Nash equilibrium.

In a complex network environment, players do not have sufficient ability to get global information, nor have enough ability to choose the best strategy to maximize their profits. This condition of blockchain shows that participants are bounded to rationality in the strategy selection process.

There are two types of behaviors in the system: cooperation and defection. Cooperation refers to the behavior of paying a definite cost to benefit the opponent, but defection is to get benefits from other players. According to the law of survival of the fittest, high payoff players will eliminate low payoff ones.

There will be group selection [18] among players. A group can contain several players, and the players within the group are cooperative, while the groups are competing. The contact relationship of blockchain participants also needs to be considered, which can be represented by the network. In the network, the node means a player, and the edge between nodes serves as the game relationship between players. In each round of the game, players interact according to the game model and update their strategies according to payoffs.

From a security perspective, purely technical cannot completely resolve the security threats of the blockchain system, and the design of the mechanism brings a new perspective to answer security issues.

3. The Evolutionary Game of Blockchain Security

The reason attack in the blockchain is hard to raise is not that these participants have noble morals, but that they have calculated that the benefits of honest participation are higher than malicious attacks. That is why they maintain the blockchain system instead of destroying it. Not only suffering 51% attacks, but the blockchain must also be able to resist a variety of attacks. While using various security technologies, it is necessary to construct an excellent mechanism. For example, in the process of evolution of the blockchain, there are competition mechanisms [19, 20] and voting mechanisms [21], etc. These mechanisms drive the blockchain more secure. Blockchain is a consensus system that pursues Nash equilibrium and the Nash equilibrium point of a blockchain system with an efficient mechanism design can approach the optimal ideal state. This is also a significant reason to treat blockchain as a revolutionary technology.

When the miners successfully dig a block, they become the block producer in this round. They can charge a fee for the included transactions. They can also get additional economic rewards for new block. This is an incentive for miners. With the uninterrupted generation of new blocks, each participant continuously improves its game strategy by learning the experience of the winners and forms a fresh offensive and defensive situation. Motivated by the continuous improvement of the game strategy of the blockchain participants, a developing defense-attack system of the blockchain has finally been established.

3.1. Description and Assumptions of the Model. In the blockchain system, different participants have distinct preferences. To facilitate the model, we arrange all miners

into two groups. Group A presents an inclination for cooperation, while group B is more selfish. In the evolutionary game, both groups A and B can only have access to insufficient information with bounded rationality. And they both imitate the strategies they would consider most helpful and improve the strategies they have worked in previous. Both groups will constantly adjust their strategies to maximize their payoffs.

The strategies of the two sides of the game include the "cooperation" strategy of observing the consensus mechanism and the "defection" strategy of not observing the consensus mechanism. The following are the assumptions and parameter descriptions in the model.

Define the cooperation cost of participants in group A as C_a and the cost of participants in group B as C_b. The distinction between the two groups is that group A tends to abide by the rules and gets payoffs on the blockchain honestly; with more computing power and other resources, participants of group B are more prone to invest in hostile costs to get payoffs on the blockchain. We define the total revenue brought by each round of the game as R, which will be distributed to two groups of participants after each round. When a serious malicious attack occurs in the blockchain, participants will suspect the sustainability of the blockchain system, and the revenue R cannot be maintained. To establish the long-term validity and rationality of the blockchain, the model satisfies the following conditions: $C_a > C_b, R > C_a + C_b$.

Considering that no matter what participation means, the participants chosen, even adopt some attacks, each round of block generation game will gain certain benefits as long as the blockchain can continue to run. However, this type of payoff has distinct expectations for various participants. Participants in group A can get mining rewards and commission income in the forthcoming. This type of expected payoff is defined as K_a, and group B will have hope it can control the blockchain to get illegal revenue, and we define this kind of expected revenue as K_b.

In order to reduce the occurrence of malicious situations, the blockchain system adopts some punitive measures, adding some regulations to the consensus mechanism, such as delaying the issuance of mining rewards and collecting deposits. We use P to represent this kind of punishment.

3.2. Model Analysis. Based on the assumptions and parameters in the previous section, this section establishes a payment matrix between blockchain participants. Players adopt various strategies, get the payoffs, and establish a dynamic system. We show the strategy matrix in Table 1.

We use E_a^h, E_a^m to denote the payoffs of participants in group A using different strategies (honest cooperation, malicious defection), and the payoffs of participants in group B are E_b^h, E_b^m.

$$\begin{aligned} E_a^h &= y\left(K_a + \lambda R - C_a\right) + (1-y)\left(K_a - C_a\right), \\ E_a^m &= y\left(K_a - C_a - P\right) + (1-y)\left(K_a - P\right), \\ E_b^h &= x\left(K_b + (1-\lambda)R - C_b\right) + (1-x)\left(K_b - C_b - P\right), \\ E_b^m &= x\left(K_b - C_b\right) + (1-x)\left(K_b - P\right). \end{aligned} \tag{1}$$

The average expected benefits of the two groups are

$$\begin{aligned} E_a &= xE_a^h + (1-x)E_a^m, \\ E_b &= yE_b^h + (1-y)E_b^m. \end{aligned} \tag{2}$$

Therefore, the dynamic equations for participants to choose cooperation strategies can be expressed as $F(x), F(y)$. Assume that the growth rate of the proportion of participants in various groups is equal to its benefits minus the average expected benefits of this group of participants.

$$\begin{aligned} F(x) &= \frac{\mathrm{d}x}{\mathrm{d}t} = x\left(E_a^h - E_a\right), \\ F(y) &= \frac{\mathrm{d}y}{\mathrm{d}t} = y\left(E_b^m - E_b\right). \end{aligned} \tag{3}$$

Unlike classical one-shot games, in the blockchain system, participants can take part in the game many times, and each block produced is regarded as a fresh round. Participants may shift their strategies. Malicious participants may attack or persuade honest participants to change their strategies in the next game. To investigate the changing state of participants, we employ evolutionary games in our analysis.

Assume that all blockchain participants have designated a specific strategy at the genesis round. All participants are honest or malicious, and they stay in the blockchain environment in a certain proportion. After several rounds of games, we examine the changes in the population.

In the evolutionary game, the evolutionary stable strategy (ESS) is a strategy that other strategies cannot invade. The case with punishment has a strong ESS because, according to the loss aversion theory [22], punishment can influence decision-making more than not getting incentives.

3.3. Blockchain Evolutionary Game in the Complex Network Environment. Game model, network structure, and evolution rules are the three important factors of network evolution. Blockchain is based on a peer-to-peer network, which guarantees the fairness of the status of nodes to the greatest extent. In the classic evolutionary game theory, the contact relationship between players is uniform. However, the relationship between participants in the blockchain is completely random. Therefore, the actual network factor is also an important part to be considered. The network can describe the connection relationship between blockchain participants. Each node

Table 1: Game strategy matrix.

		Group B	
		Honest cooperation	Malicious defection
Group A	Honest cooperation	$K_a + R - C_a, K_b + R - C_b$	$K_a - C_a, K_b - C_b - P$
	Malicious defection	$K_a - C_a - P, K_b - C_b$	$K_a - P, K_b - P$

Table 2: Simulation variables and values.

Variable	Symbol	Value
The expected benefit of group A	K_a	8
The expected benefit of group B	K_b	2
Participation costs of group A	C_a	4
Participation costs of group B	C_b	3
Penalty	P	10
New block benefit	R	10

represents a participant, and the edges between nodes represent the communication relationship between participants. In each round of the blockchain generation process, participants adopt evolution rules to update the game strategy.

In the small-world network, it is easier for collaborators to unite on the random network to form a large cooperative cluster, which promotes the emergence of cooperative behavior on the random network [23]. The emergence of cooperation in BA scale-free networks has emerged in many types of games, and the collaborators occupy the dominant position in the network [24]. In the BA scale-free network, the central nodes exist as stubborn collaborators. These pure collaborators can effectively resist the attacks of malicious betrayers by forming an interconnected subnet.

In the blockchain network, various participants are pursuing payoffs, but the methods used are distinct. From the perspective of blockchain security, besides paying attention to collaborators, the behavior of betrayers needs more attention. Each coin has two sides. The influence of the network structure on cooperative behavior also has a corresponding influence on the betrayers.

Except for some alliance chains and private chains, most of the blockchains have incentive mechanisms to adjust their expectations for future benefits. The blockchain system encourages participants to work actively, so that there are enough nodes in the system to ensure the normal operation of the system. Incentive theory ensures that a rational participant can continue to work honestly. If malicious participants are in a helpful position, blockchain will not reward them. Therefore, the best strategy for a rational blockchain participant is to maintain honest cooperation in the blockchain.

Distributed individuals maintaining this behavior will not pose a serious threat to the system, but malicious players contact each other to form a community structure in the network, which may pose a threat to the system. The mechanism design of the blockchain needs to ensure that rational behavior is honest. Therefore, if participants are dishonest, they will not be rewarded or even punished, so the expected future benefits will decrease. This has strong behavioral constraints on the bounded rationality assumption mentioned above and encourages miners to take part honestly in the blockchain.

In response to the above analysis, we designed corresponding simulation experiments. In the next paragraph, we will examine the development of cooperation in the blockchain from two different perspectives: the proportion of participants and the parameters in the game matrix.

4. Experimental Evaluation and Related Work

4.1. Experimental Evaluation. We evaluate different profit situations through experiments to determine the changes in the proportion of honest collaborators and malicious participants in the population in the blockchain system under distinct game conditions. We ran experiments for the reward matrix of the case (Table 2).

Initially, 67% of group A were honest cooperators, and 20% of group B were malicious betrayers. Blue and red signify the proportions of validators with honest and malicious strategies, respectively. The Y-axis tracks the population proportions, and the X-axis tracks the generations (i.e., block rounds). As shown in Figure 1, we can notice that the honest strategy is an ESS.

We also run simulations by modifying the initial proportions of malicious participants. The initial proportions of malicious participants range from 0 to 33%. As shown in Figure 2, the proportion of malicious participants has shown a downward tendency.

In the simulation experiment, we discover that, although the honest strategy is ESS, it is a weak ESS when group B's expected payoffs become higher, which is shown in Figure 3. The speculation is that when participants notice that there is a high expected payoff with defection, it will motivate them to flourish in a malicious direction.

The above experiment does not consider the influence of the network structure on the blockchain game and the frequency of cooperation. We use the parameters in Table 2 to design the game within the group A population in two typical complex networks in Figure 4. Figure 4(a) runs the WS small-

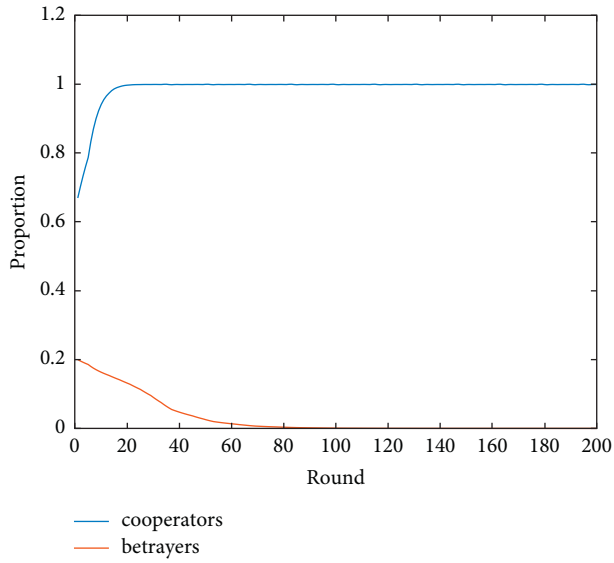


Figure 1: The evolution of population with honest and malicious strategies.

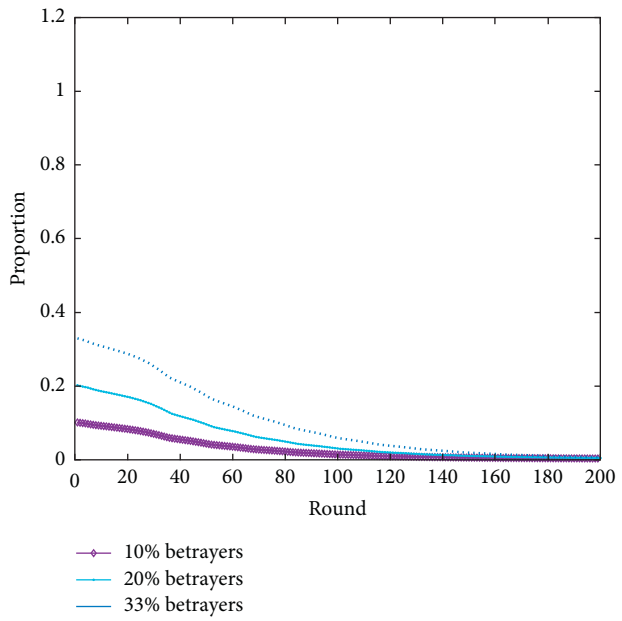


Figure 2: The initial proportions of malicious participants range from 10% to 33%.

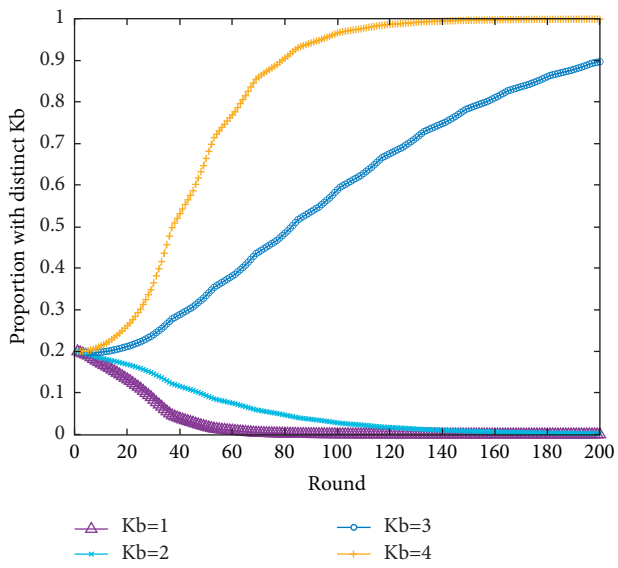


Figure 3: Influence on behavior with distinct K_b.

world model [25], and Figure 4(b) adopts the BA scale-free network model [26]. In the simulation, the NetworkX library of python is used. There are 1000 nodes in the WS network, and each node has 5 neighbors, and the probability of random reconnection is 0.3. There are 1000 nodes in the BA network, and it added 3 edges each time. The player follows the death-birth updating process. The simulation discloses that it takes quite lengthy rounds for honest collaborators to establish a trusting, cooperative relationship. This protracted process provides an opportunity for malicious behavior on the blockchain.

4.2. Related Work. In the research on the combination of blockchain and game theory, it mainly appears on the following topics: the game of mining behavior in the blockchain, security on the blockchain, and cryptocurrency and its applications.

In generating new blocks in the blockchain, there is a competitive mining process. This leads to a game between individuals. Some papers study the distribution of computing power [27] and the choice of forks [28]. With the continuous growth of miners, individual mining competition becomes more and more fierce. Therefore, miners spontaneously form mining alliances (mining pools). Some scholars conduct research on the choice of mining pools and the rewards in the mining pools [4, 29].

There are threats, such as selfish mining, 51% attacks, and denial of service (DoS) on the blockchain. Some papers start with the study of gaming behavior from the aspect of blockchain security. There is a game between the cost of an

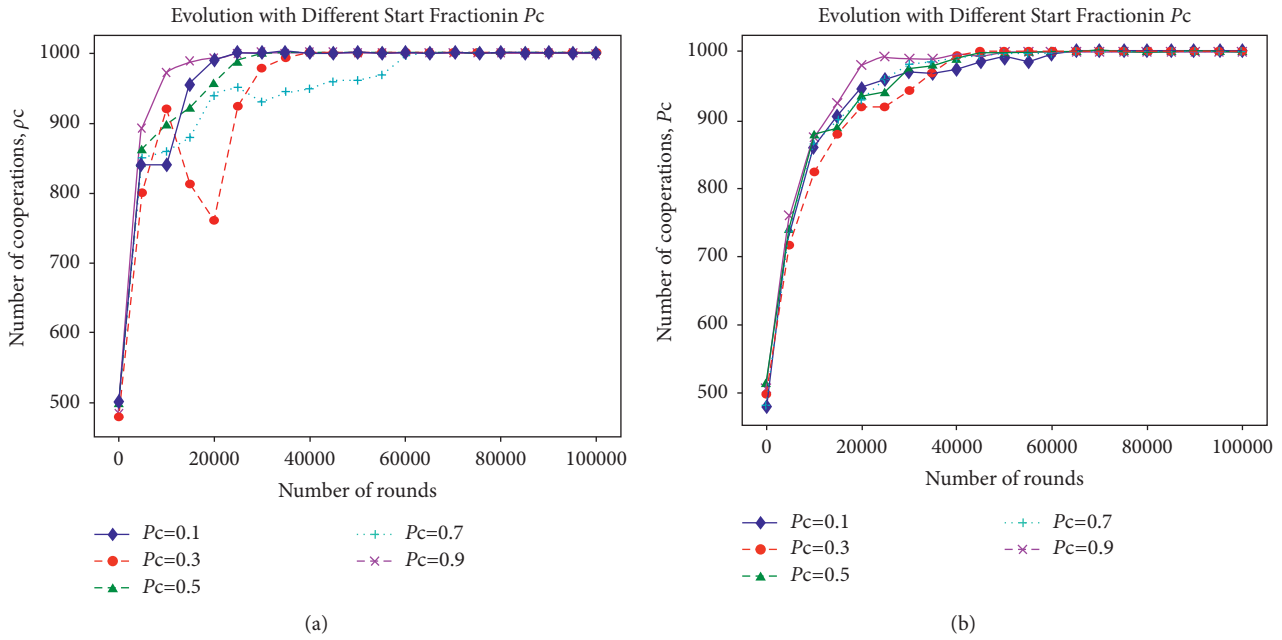


FIGURE 4: Cooperative evolution in two types of complex networks. (a) Cooperative evolution in BA scale-free network. (b) Cooperative evolution in WS small-world network.

attack on the blockchain and whether the system launches an attack and the strength of the attack [5, 30].

An important application of game theory is in the analysis of economic behavior, so a lot of research involves the game of economic behavior in the blockchain, for example, the choice between different cryptocurrencies [31], the choice of transactions when making a new block [32], and some payment behaviors involving actual transactions [33].

The article [34] focuses on the reward mechanism in the blockchain and uses evolutionary games to analyze how the behavior of participants develops with the reward mechanism. The authors also found that punishment plays a central role in maintaining the integrity of the blockchain. Compared with this article, in our blockchain game model, we found that future revenue expectations will affect the behavior of the blockchain. We also consider the influence of the network model on the cooperation mechanism in the blockchain.

5. Conclusion

We formulated the cooperation and defection behavioral game in the blockchain system. Using evolutionary game theory, we found the importance of reducing the profit expectations of malicious participants in maintaining the integrity of the ledger. When considering the influence of network factors, it takes a long time to establish a cooperative process within the blockchain. Thus, more attention needs to be paid to the security protection of the initial stage of the blockchain. These findings provide mechanism guidance for the security of the blockchain. A safe and reliable system should be considered from the most basic physical layer. We will start with the security and credibility requirements of the physical layer and design corresponding technologies and mechanisms, enhance our game model, and construct a stable blockchain mechanism.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this study.

Acknowledgments

This work was funded by the Postgraduate Research and Practice Innovation Program of Jiangsu Province under Grant KYCX18_0903.

References

[1] S. Nakamoto, "Bitcoin: a peer-to-peer electronic cash system," 2008, https://bitcoin.org/bitcoin.pdf.

[2] C. Dwork and M. Naor, "Pricing via processing or combatting junk mail," in *Proceedings of the Annual International Cryptology Conference*, pp. 139–147, Springer, Santa Barbara, CA, USA, 1992.

[3] N. Houy, "The bitcoin mining game," *Ledger*, vol. 1, 2014.

[4] O. Schrijvers, J. Bonneau, D. Boneh, and T. Roughgarden, "Incentive compatibility of bitcoin mining pool reward functions," in *Proceedings of the International Conference on Financial Cryptography and Data Security*, pp. 477–498, Springer, Christ Church, Barbados, 2016.

[5] I. Eyal, "The miner's dilemma," in *Proceedings of the 2015 IEEE Symposium on Security and Privacy*, pp. 89–103, IEEE, San Jose, CA, USA, 2015.
[6] K. Nayak, S. Kumar, A. Miller, and E. Shi, "Stubborn mining: generalizing selfish mining and combining with an eclipse attack," in *Proceedings of the 2016 IEEE European Symposium on Security and Privacy (EuroS&P)*, pp. 305–320, IEEE, Saarbrücken, Germany, 2016.
[7] P. Sztompka, *Trust: A Sociological Theory*, Cambridge University Press, Cambridge, UK, 1999.
[8] D. Gambetta, "Can we trust trust," *Trust: Making and breaking cooperative relations*, vol. 13, pp. 213–237, 2000.
[9] R. Axelrod and W. D. Hamilton, "The evolution of cooperation," *Science*, vol. 211, no. 4489, pp. 1390–1396, 1981.
[10] P. Lamsal, *Understanding Trust and Security*, Department of Computer Science, University of Helsinki, Helsinki, Finland, 2001.
[11] M. Ylianttila, "6g white paper: research challenges for trust, security and privacy," 2020, http://arxiv.org/abs/2004.11665.
[12] J. Zhang and M. Wu, "Blockchain use in IoT for privacy-preserving anti-pandemic home quarantine," *Electronics*, vol. 9, no. 10, p. 1746, 2020.
[13] S. Chen, L. Yang, C. Zhao, V. Varadarajan, and K. Wang, "Double-blockchain assisted secure and anonymous data aggregation for fog-enabled smart grid," *Engineering*, 2020.
[14] M. Wu, K. Wang, X. Cai, S. Guo, M. Guo, and C. Rong, "A comprehensive survey of blockchain: from theory to IoT applications and beyond," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8114–8154, 2019.
[15] Y. Liu, K. Wang, Y. Lin, and W. Xu, "LightChain: a lightweight blockchain system for industrial Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 6, pp. 3571–3581, 2019.
[16] E. Heilman, A. Kendler, A. Zohar, and S. Goldberg, "Eclipse attacks on bitcoin's peer-to-peer network," in *Proceedings of the 24th {USENIX} Security Symposium ({USENIX} Security 15)*, pp. 129–144, Santa Clara, CA, USA, 2015.
[17] C. Daskalakis, G. Schoenebeckt, G. Valiant, and P. Valiant, "On the complexity of Nash equilibria of action-graph games," in *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 710–719, SIAM, New York, NY, USA, 2009.
[18] A. Traulsen and M. A. Nowak, "Evolution of cooperation by multilevel selection," *Proceedings of the National Academy of Sciences*, vol. 103, no. 29, pp. 10952–10955, 2006.
[19] I. Bentov, C. Lee, A. Mizrahi, and M. Rosenfeld, "Proof of activity," *ACM SIGMETRICS-Performance Evaluation Review*, vol. 42, no. 3, pp. 34–37, 2014.
[20] A. Kiayias, A. Russell, B. David, and R. Oliynykov, "Ouroboros: a provably secure proof-of-stake blockchain protocol," in *Proceedings of the Annual International Cryptology Conference*, pp. 357–388, Springer International Publishing, Santa Barbara, CA, USA, 2017.
[21] I. Grigg, "Eos-an introduction," 2017, https://whitepaperdatabase.com/eos-whitepaper.
[22] J. Shalev, "Loss aversion equilibrium," *International Journal of Game Theory*, vol. 29, no. 2, pp. 269–287, 2000.
[23] C. Hauert and G. Szabó, "Game theory and physics," *American Journal of Physics*, vol. 73, no. 5, pp. 405–414, 2005.
[24] F. C. Santos and J. M. Pacheco, "Scale-free networks provide a unifying framework for the emergence of cooperation," *Physical Review Letters*, vol. 95, no. 9, Article ID 098104, 2005.
[25] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
[26] A. L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
[27] N. Dimitri, "Bitcoin mining as a contest," *Ledge*, vol. 2, pp. 31–37, 2017.
[28] B. Biais, C. Bisière, M. Bouvard, and C. Casamatta, "The blockchain folk theorem," *Review of Financial Studies*, vol. 32, no. 5, pp. 1662–1715, 2019.
[29] Y. Lewenberg, Y. Bachrach, Y. Sompolinsky, A. Zohar, and J. S. Rosenschein, "Bitcoin mining pools: a cooperative game theoretic analysis," in *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pp. 919–927, Istanbul, Turkey, 2015.
[30] J. A. Kroll, I. C. Davey, and E. W. Felten, "The economics of bitcoin mining, or bitcoin in the presence of adversaries," *Proceedings of WEIS*, vol. 2013, p. 11, 2013.
[31] L. W. Cong, Y. Li, and N. Wang, "Tokenomics: dynamic adoption and valuation," *Review of Financial Studies*, vol. 34, no. 3, pp. 1105–1155, 2021.
[32] N. Houy, "The economics of Bitcoin transaction fees," *GATE WP*, vol. 1407, 2014.
[33] J. Kang, R. Yu, X. Huang, S. Maharjan, Y. Zhang, and E. Hossain, "Enabling localized peer-to-peer electricity trading among plug-in hybrid electric vehicles using consortium blockchains," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 6, pp. 3154–3164, 2017.
[34] S. Motepalli and H.-A. Jacobsen, "Reward mechanism for blockchains using evolutionary game theory," 2021, http://arxiv.org/abs/2104.05849.

Hindawi
Complexity
Volume 2021, Article ID 3394666, 11 pages
https://doi.org/10.1155/2021/3394666

Research Article

Chaotic Behavior Analysis of a New Incommensurate Fractional-Order Hopfield Neural Network System

Nadjette Debbouche,[1] Adel Ouannas,[2] Iqbal M. Batiha,[3,4] Giuseppe Grassi,[5] Mohammed K. A. Kaabar,[6,7,8] Hadi Jahanshahi,[9] Ayman A. Aly,[10] and Awad M. Aljuaid[11]

[1]*Department of Mathematics and Computer Science, University of Larbi Ben M'hidi, Oum El Bouaghi, Algeria*
[2]*Laboratory of Dynamical Systems and Control, University of Larbi Ben M'hidi, 04000 Oum El Bouaghi, Algeria*
[3]*Department of Mathematics, Faculty of Science and Technology, Irbid National University, 2600 Irbid, Jordan*
[4]*Nonlinear Dynamics Research Center (NDRC), Ajman University, 346 Ajman, UAE*
[5]*Dipartimento Ingegneria Innovazione, Universita del Salento, 73100 Lecce, Italy*
[6]*Gofa Camp, Near Gofa Industrial College and German Adebabay, Nifas Silk-Lafto, 26649 Addis Ababa, Ethiopia*
[7]*Jabalia Camp, United Nations Relief and Works Agency (UNRWA), Palestinian Refugee Camp, Gaza Strip, Jabalya, State of Palestine*
[8]*Institute of Mathematical Sciences, Faculty of Science, University of Malaya, Kuala Lumpur 50603, Malaysia*
[9]*Department of Mechanical Engineering, University of Manitoba, Winnipeg R3T 5V6, Canada*
[10]*Department of Mechanical Engineering, College of Engineering, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia*
[11]*Department of Industrial Engineering, College of Engineering, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia*

Correspondence should be addressed to Mohammed K. A. Kaabar; mohammed.kaabar@wsu.edu

Received 3 May 2021; Revised 6 September 2021; Accepted 7 October 2021; Published 8 November 2021

Academic Editor: Miaomiao Wang

This study intends to examine different dynamics of the chaotic incommensurate fractional-order Hopfield neural network model. The stability of the proposed incommensurate-order model is analyzed numerically by continuously varying the values of the fractional-order derivative and the values of the system parameters. It turned out that the formulated system using the Caputo differential operator exhibits many rich complex dynamics, including symmetry, bistability, and coexisting chaotic attractors. On the other hand, it has been detected that by adapting the corresponding controlled constants, such systems possess the so-called offset boosting of three variables. Besides, the resultant periodic and chaotic attractors can be scattered in several forms, including 1D line, 2D lattice, and 3D grid, and even in an arbitrary location of the phase space. Several numerical simulations are implemented, and the obtained findings are illustrated through constructing bifurcation diagrams, computing Lyapunov exponents, calculating Lyapunov dimensions, and sketching the phase portraits in 2D and 3D projections.

1. Introduction

The artificial neural networks, which are deemed one of the deepest learning technologies that are included under the rubric of the so-called artificial intelligence, have recently received a considerable amount of interest by many researchers whose practical work associates with the human brain [1]. In order to make an effective progress with the development of modern engineering and electronics, it is necessary to continuously attempt to improve this intelligent scheme. In 1943, McCulloch and Pitts are the first scientists who studied the artificial neural networks [2]. In conformity with their investigation, numerous benign engineering and electronic applications have been employed in several applied fields. For instance, switching in electronic circuits, the oscillation of systems in accordance with the impact of an

earthquake, image and signal processing, impacting machines, power circuits, and dry friction are some of such applications [3]. In 1982, with the aim of dealing with some optimization and computational issues and to conquer some specific problems associated with the hardware's execution, a novel memory neural network was established by Hopfield called later on the Hopfield neural network (HNN) [4]. Presently, this type of networks has begun taking its place in different industrial sectors, motivating a lot of investigators to further explore the dynamical properties for its states and moreover deduce other ones [5].

Due to the key role of using the fractional calculus in formulating many phenomena rather than that of using the classical calculus, the HNNs were fractionalized to be later on called the fractional-order Hopfield neural networks (FoHNNs) [6, 7]. The basic idea of the FoHNNs' origin may be owed to Boroomand and Menhaj who carried out a replacement of the fractance instead of the ordinary capacitor within the classical HNN model [8]. The key benefit of such replacement refers to the truth that the fractional-order derivatives can describe the HNN more efficiently due to the infinite memory and some other hereditary properties that can be generated from its various processes [9]. From a different point of view, the inclusion process of a memory term into the HNN model by incorporating the fractional-order derivatives/integrals can provide a super calculation capability, which might be needed in, e.g., the stimulus anticipation and the information processing as well as other calculations associated with the oscillatory neuronal firing [10]. For these reasons and more, the analysis of FoHNNs is recently considered one of the main promising topics that benefits future researchers in different applied science fields.

Several significant numerical findings related to the presence of the chaotic behavior and the limit cycles for the dynamics of the FoNN are discussed in the literature. For instance, the stability of the FoHNN was fully investigated through energy-like function analysis in [8]. However, a theoretical approach, based on the harmonic balance theory, was used to investigate the existence of chaos for a cellular neural network model in [11]. A chaos neuron model was proposed and examined as a novel artificial neuron model in [12]. Moreover, in [13], a fractional-order cellular neural network model was introduced by replacing the traditional first-order cell by the fractional-order one. The chaotic synchronization of such networks was also discussed in several other papers. We find certain chaotic behaviors in the time-delayed FoNN system, studied well in [14]. Also, the chaos feedback control and synchronization systems were constructed for a neuron network system by Zhou et al. in [15]. One year later, Zhou et al. once again illustrated the chaotic synchronization system for a FoNN system in [16]. In [17], the Laplace transform and the generalized Gronwall inequality were employed to examine the FoHNN models in terms of its finite-time stability, whereas various dynamic features like constructing bifurcation diagrams, chaos, stability, and multistability of the FoNNs were studied in [9]. In [18], an α-synchronization and an α-stability were determined and explored in the FoNN models, whereas a uniform stability of such models was described and analyzed with time delay in [19] with the help of using an efficient kind of error norm. Similarly, the complex dynamics of several modern maps have been recently studied such as the initial-dependent extreme multistability and offset-boosted coexisting attractor.

In light of the several applications of the FoHNN models in different applied science fields and in order to move forward in further discovery of more properties of the FoHNNs dynamics, this paper attempts to study different chaotic dynamics of such networks with incommensurate order. Besides, it intends to analyze the stability of the proposed system numerically by continuous varying the fractional-order derivative values as well as the values of system parameters. Such analysis will be carried out by performing several numerical simulations, like constructing the bifurcation diagrams, computing Lyapunov exponents, calculating Lyapunov dimensions, and sketching the phase portraits in 2D and 3D projections. Several rich complex dynamics, including symmetry, bistability, and coexisting chaotic attractors will be investigated and discussed. It will be also shown, through adapting some controlled constants, that the proposed system will possess the offset boosting of three variables. However, the structure of this article is arranged in the following manner: Section 2 will study and discuss different complex dynamics of the incommensurate FoHNN model through illustrating their different corresponding numerical simulations. Dynamics of the incommensurate fractional-order model will be discussed in Section 3. The variable-boostable attractors that are generated from the incommensurate-order model will be discussed and analyzed in Section 4, followed by Section 5 that will summarize the whole work.

2. Preliminaries and Mathematical Model

This part intends to describe some essential fundamentals in regard to the fractional calculus, especially the Caputo differential operator and the Riemann–Liouville integral operator which were reported in [20]. Besides, a new version of the HNN model will be proposed in light of its incommensurate fractional-order derivatives.

Definition 1. The Riemann–Liouville fractional-order integral operator of the function h can be expressed as

$$I^{\delta}h(t) = \frac{1}{\Gamma(\delta)}\int_0^t \frac{h(\tau)}{(t-\tau)^{(1-\delta)}}\mathrm{d}\tau, \tag{1}$$

where $h \in C^m(0,T]$, $\delta > 0$, $T > 0$, and $m \in \mathbb{N}$.

Definition 2. The Caputo fractional-order differential operator of the function h can be expressed as

$$D^{\delta}h(t) = \begin{cases} \dfrac{1}{\Gamma(m-\delta)}\displaystyle\int_0^t (t-\tau)^{m-\delta-1}h^{(m)}(\tau)\mathrm{d}\tau, & \delta \in (m-1,m), \\ h^{(m)}(t), & \delta = m, \end{cases} \tag{2}$$

where $h \in C^m(0,T]$, $T > 0$, $m \in \mathbb{N}$, and $\delta \in [m-1,m]$.

In [9], a mathematical model of FoHNN with a ring structure was established and its stability analysis was discussed in view of some features of its parameters. The model was formulated by the following three-dimensional system:

$$\begin{cases} D^q x(t) = -x(t) + 2\,\sin(x(t)) + a\,\sin(y(t)) + b\,\sin(z(t)), \\ D^q y(t) = -y(t) + b\,\sin(x(t)) + 2\,\sin(y(t)) + a\,\sin(z(t)), \\ D^q z(t) = -z(t) + a\,\sin(x(t)) + b\,\sin(y(t)) + 2\,\sin(z(t)), \end{cases} \tag{3}$$

where a and b are the system parameters, D^q is the Caputo differential operator of order q, and x, y, and z are the states of the system. As a matter of fact, Kaslik and Sivasundaram studied and analyzed the stability of the above system by taking its fractional-order derivatives in commensurate-order case [9]. To illustrate the chaotic motion of this neural network system in its integer-order case (i.e., system (4) when $q = 1$), we plot Figure 1 that shows its phase portraits in different planes generated by its states. The basin of attraction of this integer-order system is also shown in Figure 2, where the initial conditions (ICs) can be shown in yellow and blue regions, leading to the limit cycles and to the chaotic attractors, respectively.

In this work, we assert that if we change these derivatives to be in its incommensurate-order case, then such systems will exhibit more rich complex dynamics and more chaotic patterns than in the previous model reported in [9]. Thus, the new incommensurate fractional-order version of the previous HNN model, which will be considered from now on, can be formulated as follows:

$$\begin{cases} D^{q_1} x(t) = -x(t) + 2\,\sin(x(t)) + a\,\sin(y(t)) + b\,\sin(z(t)), \\ D^{q_2} y(t) = -y(t) + b\,\sin(x(t)) + 2\,\sin(y(t)) + a\,\sin(z(t)), \\ D^{q_3} z(t) = -z(t) + a\,\sin(x(t)) + b\,\sin(y(t)) + 2\,\sin(z(t)), \end{cases} \tag{4}$$

where a, b, x, y, and z are defined above, while D^{q_i} is the Caputo differential operator of order q_i such that $0 < q_i \leq 1$ and $i = 1, 2, 3$. Actually, in order to solve the above system, one can implement the predictor-corrector method which was proposed by Diethelm et al. in [21]. Although this method can provide an accurate solution of a given nonlinear fractional-order system numerically [22], an enhanced approach called Adams–Bashforth–Moulton scheme was established in [23, 24].

3. Dynamics of the Incommensurate Fractional-Order Model

In this section, different complex dynamics of incommensurate FoHNN model (4) will be numerically studied and analyzed, including discussion of the stability analysis versus taking different values of incommensurate fractional-order derivatives and also of system parameters and presenting some special phenomena that could be generated from the proposed model such as presenting the so-called symmetry and coexisting attractors. For this purpose, several numerical tools will be used for performing some required simulations, including constructing bifurcation diagrams, sketching the phase portraits of the system dynamics in 2D and 3D projections, plotting the basin of attractions, and also computing Lyapunov exponents/dimensions.

3.1. Stability vs. Different Incommensurate Fractional-Order Derivatives. Here, we will take the two parameters of the system as $a = 1$ and $b = -9$ as well as we will consider the IC of the states of system (4) as $(x_0, y_0, z_0) = (-1, 1, 1)$. In this part, we aim to study the stability of model (4) according to the following three cases: the first one is carried out by fixing $q_2 = q_3 = 1$ and continuously varying the value of q_1, the second one is performed by fixing $q_1 = q_2 = 1$ and continuously varying the value of q_2, and finally, the third one is implemented by fixing $q_1 = q_2 = 1$ and continuously varying the value of q_3. Accordingly, the bifurcation diagrams and Lyapunov exponents for such three cases are plotted in Figures 3–5, respectively. Based on these figures, one can observe the existence of positive Lyapunov exponents, confirming that FoHNN system (4) shows chaotic behavior. In addition, one might notice, from such figures again, that system (4) is asymptotically stable to a limit cycle when $q_1, q_2, q_3 \in (0.6, 0.73)$, while it begins losing its stability and begins behaving in a chaotic mode when $q_1 \in (0.82, 1)$, $q_2 \in (0.83, 1)$, and $q_3 \in (0.84, 1)$. It might be further noticed that once the values of the incommensurate fractional-order derivatives are increased, different periodic windows will be shown. For instance, to deal with the aforesaid three cases, we may take the values of the fractional-order derivatives as $q_1 = 0.89$ and $q_2 = q_3 = 1$, which, directly, implies the three Lyapunov exponents, $\mathrm{LE}_1 = 0.26$, $\mathrm{LE}_2 = 0$, and $\mathrm{LE}_3 = -4.39$, with $|0.26| < |-4.398|$. On the other hand, if one takes $q_2 = 0.90$ and $q_1 = q_3 = 1$, then the following three Lyapunov exponents will be yielded: $\mathrm{LE}_1 = 0.51$, $\mathrm{LE}_2 = 0$, and $\mathrm{LE}_3 = -2.63$, with $|0.51| < |-2.63|$. Finally, taking $q_3 = 0.87$ and $q_1 = q_2 = 1$ gives the following three Lyapunov exponents: $\mathrm{LE}_1 = 0.31$, $\mathrm{LE}_2 = 0$, and $\mathrm{LE}_3 = -5.06$, with $|0.31| < |-5.06|$. Consequently, the corresponding Lyapunov dimension, which is considered an estimation of the complexity of the generated attractor from the system, can be calculated according to the previous three cases by considering the following formula:

$$D_{KY} = j + \frac{1}{\left|\mathrm{LE}_{j+1}\right|} \sum_{i=1}^{j} \mathrm{LE}_i, \tag{5}$$

where j is the greatest integer number that satisfies $\sum_{i=1}^{j} \mathrm{LE}_i \geq 0$ and $\sum_{i=1}^{j+1} \mathrm{LE}_i < 0$. In other words, the calculated Lyapunov dimensions of system (4) according to the previous values of incommensurate fractional-order derivatives are $D_{KY} = 2.09 > 2$, $D_{KY} = 2.19 > 2$, and $D_{KY} = 2.06 > 2$, respectively. This, consequently, leads us to deduce that there is a chaotic attractor of system (4). To see this, Figures 6–8 present the complex chaotic attractor of system (4) on different 2D projections according to the three considered cases: ($q_1 = 0.89$ and $q_2 = q_3 = 1$), ($q_2 = 0.90$ and $q_1 = q_3 = 1$), and ($q_3 = 0.87$ and $q_1 = q_2 = 1$), respectively. In addition, Figures 9–11 exhibit the corresponding 3D projections of the chaotic attractors that are generated by system (4).

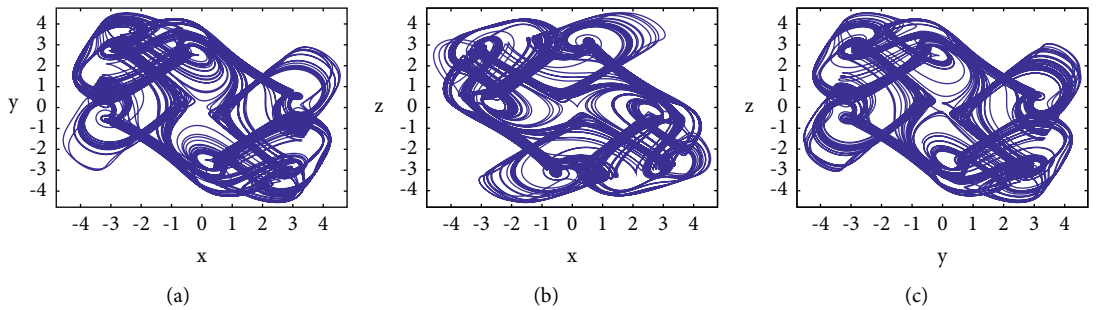


FIGURE 1: Chaotic attractor of system (3) when $q = 1$ according to the system parameters $a = 1$ and $b = -9$ and the IC $(x_0, y_0, z_0) = (-1, 1, 1)$ on (a)xy plane, (b)xz plane, and (c)yz plane.

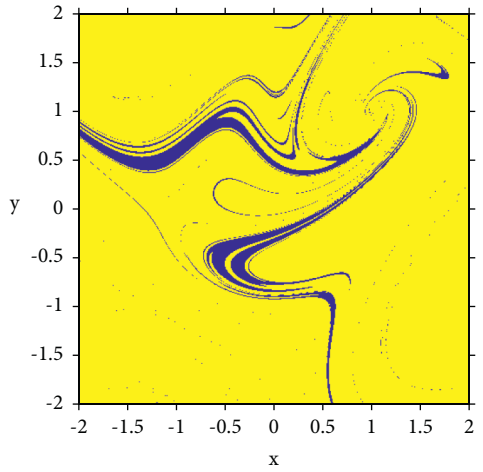


FIGURE 2: Basin of attraction of system (3) when $q = 1$ according to the system parameters $a = 1$ and $b = -9$ and to the initial condition of the third state variable $z = 0$.

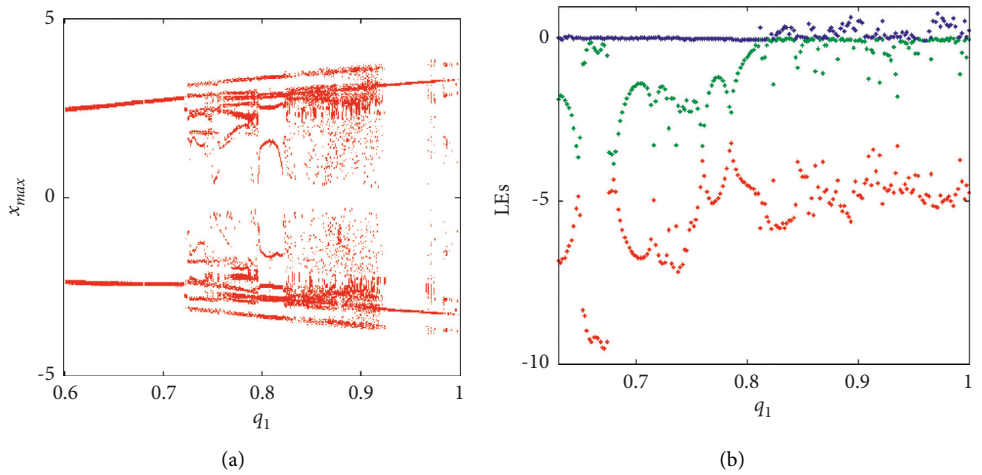


FIGURE 3: (a) The diagram of bifurcation. (b) Lyapunov exponents of system (4) when $q_1 \in (0.6, 1)$ and $q_2 = q_3 = 1$ according to the system parameters $a = 1$ and $b = -9$ and the IC $(x_0, y_0, z_0) = (-1, 1, 1)$.

3.2. Stability vs. Different Values of the System Parameters. This section will explore the stability of system (4) by continuously varying the values of the system parameters a and b and by fixing the incommensurate fractional-order derivative values at $[q_1, q_2, q_3] = [0.8, 0.90, 0.91]$ as well as fixing the IC at $(x_0, y_0, z_0) = (-1, 1, 1)$. Immediately, Figures 12 and 13 represent the bifurcation diagrams together with the Lyapunov exponents of system (4) when $a \in (0, 5)$ and $b \in (-20, 0)$, respectively. One might observe based on such plots that when the values of the system parameters are

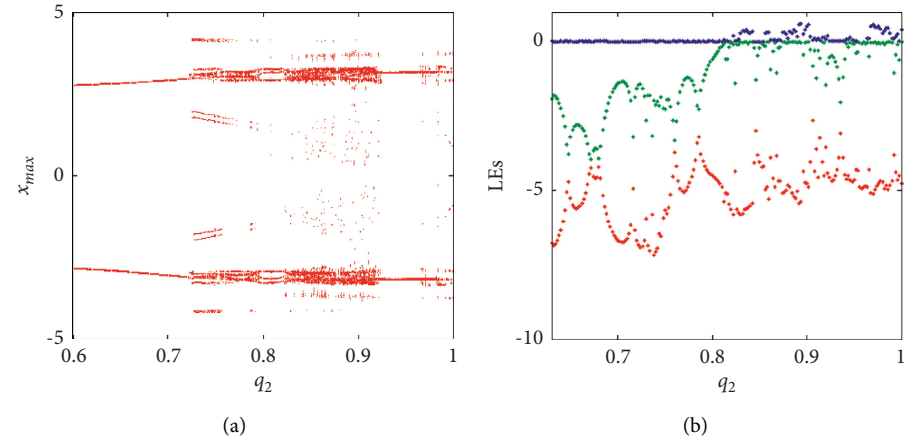


FIGURE 4: (a) The diagram of bifurcation. (b) Lyapunov exponents of system (4) when $q_2 \in (0.6, 1)$ and $q_1 = q_3 = 1$ according to the system parameters $a = 1$ and $b = -9$ and the IC $(x_0, y_0, z_0) = (-1, 1, 1)$.

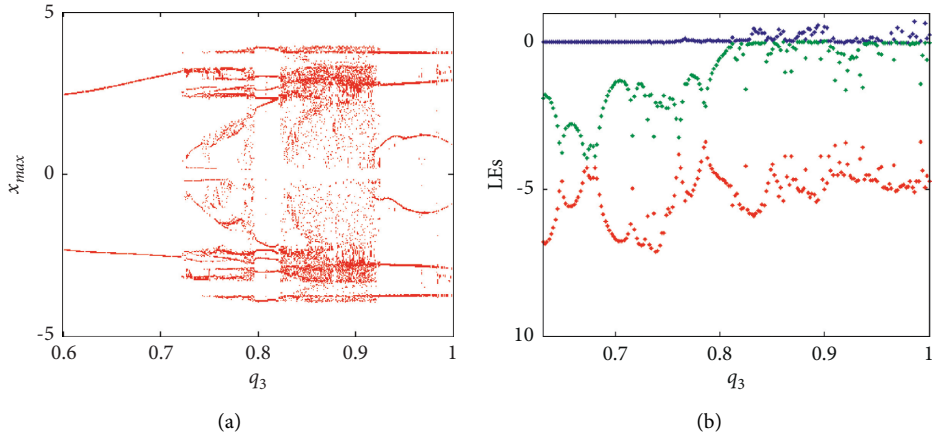


FIGURE 5: (a) The diagram of bifurcation. (b) Lyapunov exponents of system (4) when $q_1 \in (0.6, 1)$ and $q_2 = q_3 = 1$ according to the system parameters $a = 1$ and $b = -9$ and the IC $(x_0, y_0, z_0) = (-1, 1, 1)$.

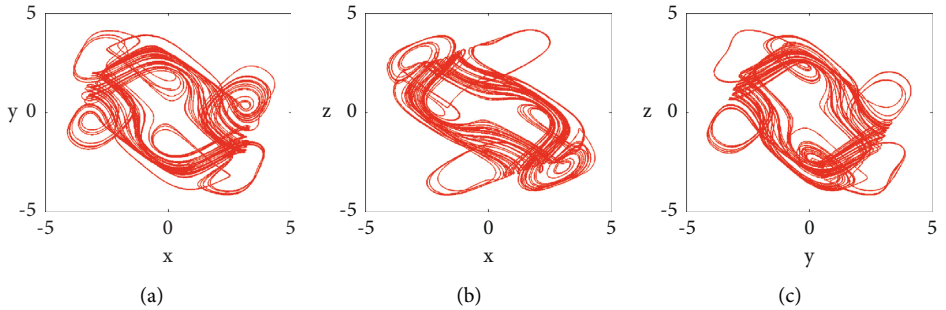


FIGURE 6: Chaotic attractor of system (4) when $q_1 = 0.89$ and $q_2 = q_3 = 1$ according to the system parameters $a = 1$ and $b = -9$ and the IC $(x_0, y_0, z_0) = (-1, 1, 1)$ on (a)xy plane, (b)xz plane, and (c)yz plane.

decreased, system (4) will be asymptotically stable and then it will exhibit periodic oscillations to chaos. Moreover, the chaotic ranges with periodic windows will be of $a \in (0, 2.9)$ and $b \in (-20, -7.5)$. Obviously, there exist certain positive Lyapunov exponents within these ranges, confirming the chaotic behavior of system (4).

3.3. Symmetry, Bistability, and Coexisting Chaotic Attractors. Symmetric dynamical systems are typically obtained when they often exhibit a symmetric pair of coexisting attractors. This property has attracted considerable interest in the field of nonlinear dynamic systems. To obtain a complete overview of this property, the reader may refer to [25]. In regard

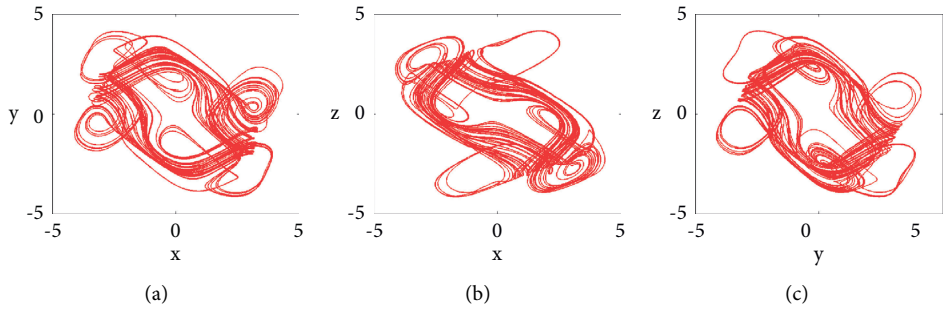


FIGURE 7: Chaotic attractor of system (4) when $q_2 = 0.90$ and $q_1 = q_3 = 1$ according to the system parameters $a = 1$ and $b = -9$ and the IC $(x_0, y_0, z_0) = (-1, 1, 1)$ on (a)xy plane, (b)xz plane, and (c)yz plane.

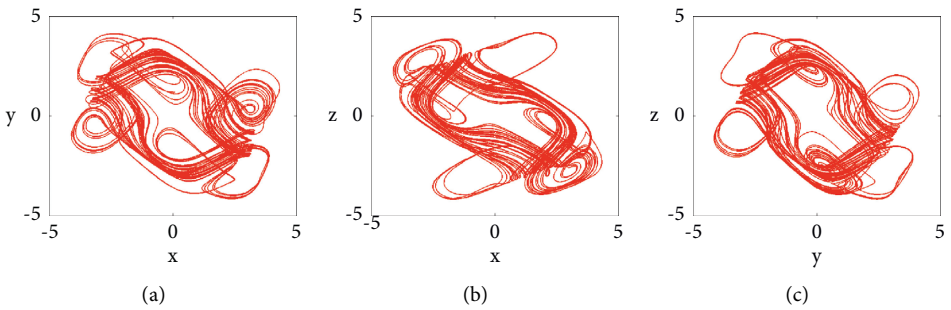


FIGURE 8: Chaotic attractor of system (4) when $q_3 = 0.87$ and $q_1 = q_2 = 1$ according to the system parameters $a = 1$ and $b = -9$ and the IC $(x_0, y_0, z_0) = (-1, 1, 1)$ on (a)xy plane, (b)xz plane, and (c)yz plane.

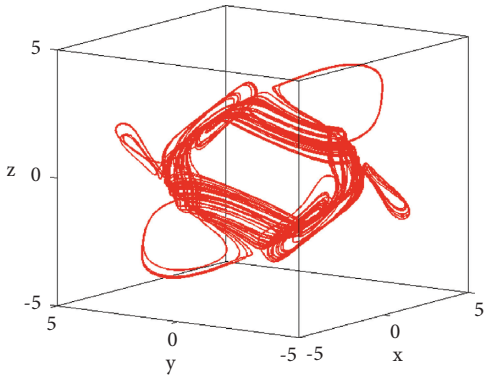


FIGURE 9: Chaotic attractor of system (4) on 3D projections when $q_1 = 0.89$ and $q_2 = q_3 = 1$ according to the system parameters $a = 1$ and $b = -9$ and the IC $(x_0, y_0, z_0) = (-1, 1, 1)$.

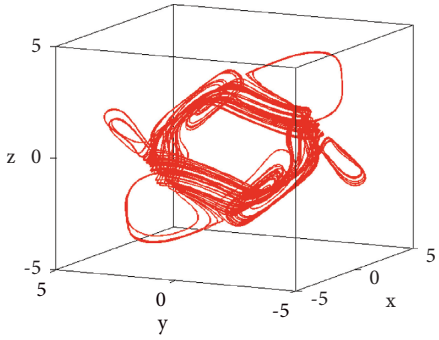


FIGURE 10: Chaotic attractor of system (4) on 3D projections when $q_2 = 0.90$ and $q_1 = q_3 = 1$ according to the system parameters $a = 1$ and $b = -9$ and the IC $(x_0, y_0, z_0) = (-1, 1, 1)$.

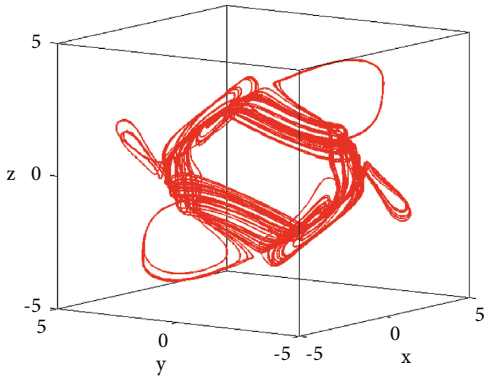


FIGURE 11: Chaotic attractor of system (4) on 3D projections when $q_3 = 0.87$ and $q_1 = q_2 = 1$ according to the system parameters $a = 1$ and $b = -9$ and the IC $(x_0, y_0, z_0) = (-1, 1, 1)$.

to our study, we will assume that the system parameters are $a = 1$ and $b = -9$ and we will select two ICs as $(x_0, y_0, z_0) = (1, 1, 1)$ for the red trajectory and $(x_0, y_0, z_0) = (-1, -1, -1)$ for the blue trajectory. In light of these values, we observe that system (4) will, e.g., generate a symmetric pair of coexisting limit cycles when $[q_1, q_2, q_3] = [0.70, 1, 1]$ (see Figure 14(a)), while it will generate a symmetric pair of coexisting periodic attractors when $[q_1, q_2, q_3] = [0.80, 1, 1]$ (see Figure 14(b)), and moreover, it will generate a symmetric pair of coexisting chaotic attractors when $[q_1, q_2, q_3] = [0.89, 1, 1]$ (see Figure 14(c)).

On the other hand, the bistability property, which is deemed one of the most recent dynamic phenomena of a system, has attracted many researchers in recent years.

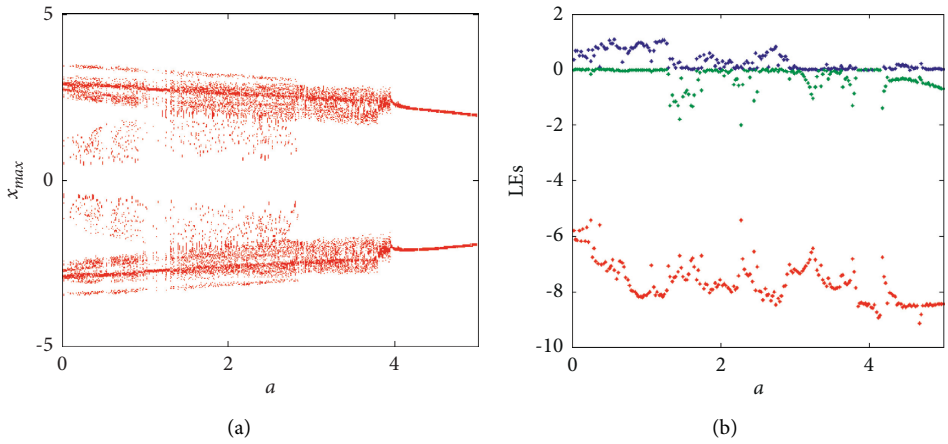


FIGURE 12: (a) The diagram of bifurcation. (b) Lyapunov exponents of system (4) when $a \in (0, 5)$ according to the other system parameter $b = -9$, the incommensurate fractional order $[q_1, q_2, q_3] = [0.8, 0.90, 0.91]$, and the IC $(x_0, y_0, z_0) = (-1, 1, 1)$.

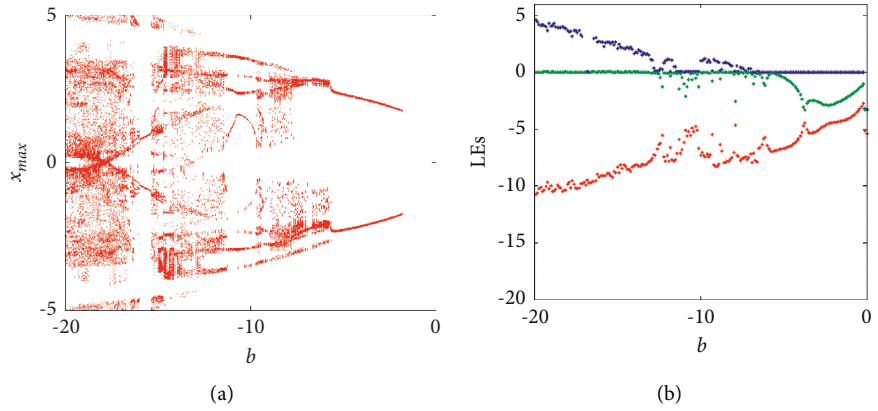


FIGURE 13: (a) The diagram of bifurcation. (b) Lyapunov exponents of system (4) when $b \in (-20, 0)$ according to the other system parameter $a = 1$, the incommensurate fractional order $[q_1, q_2, q_3] = [0.8, 0.90, 0.91]$, and the IC $(x_0, y_0, z_0) = (-1, 1, 1)$.

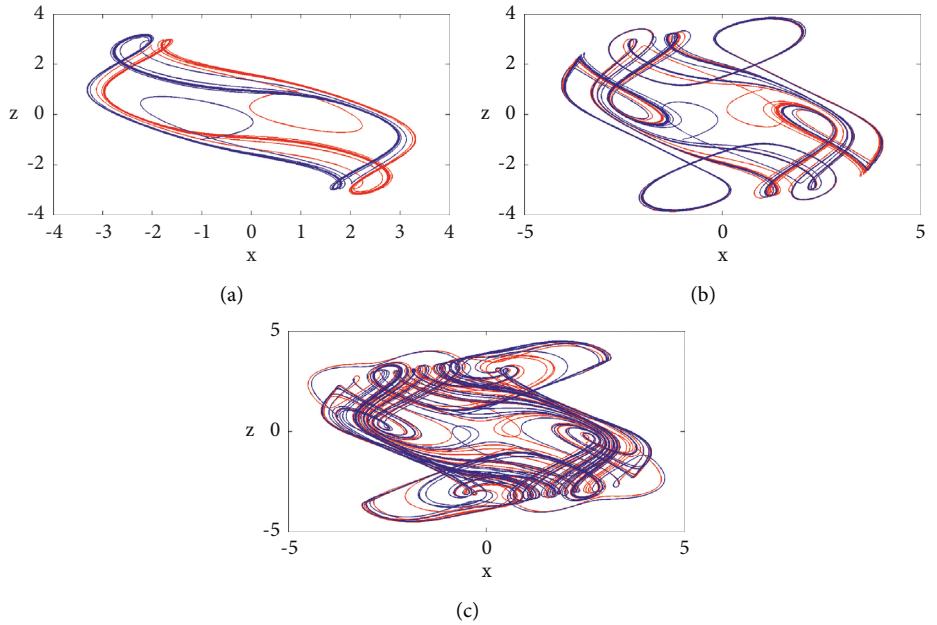


FIGURE 14: Phase portraits of coexisting symmetric attractors according to the system parameters $a = 1$ and $b = -9$, the IC $(x_0, y_0, z_0) = (1, 1, 1)$ for the red line, and the IC $(x_0, y_0, z_0) = (-1, -1, -1)$ for the blue line by continuously varying q_i as (a) $[q_1, q_2, q_3] = [0.70, 1, 1]$, (b) $[q_1, q_2, q_3] = [0.80, 1, 1]$, and (c) $[q_1, q_2, q_3] = [0.89, 1, 1]$.

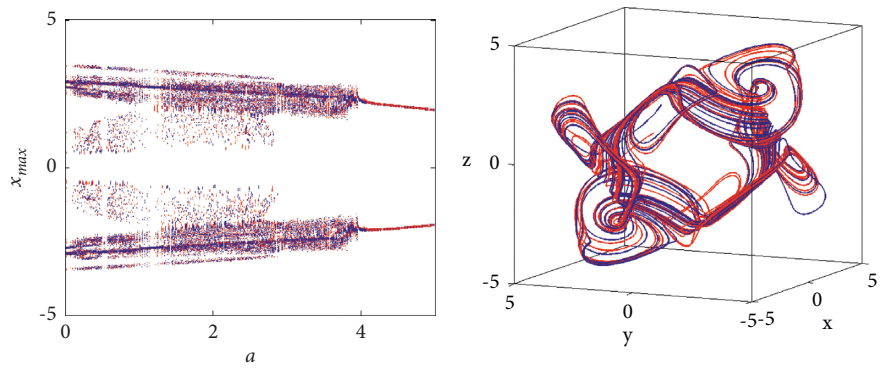


FIGURE 15: (a) Bifurcation diagram of system (4) for continuous varying $a \in (0, 5)$. (b) Coexisting chaotic attractors for $a = 1$, $b = -9$, and $[q_1, q_2, q_3] = [0.80, 0.90, 0.91]$ subject to the ICs: $(x_0, y_0, z_0) = (-1, 1, 1)$ for the red plot and $(x_0, y_0, z_0) = (0.1, 1.2, 2)$ for the blue plot.

For more clarifications about this property, the reader may refer to [26]. However, in order to show the bistability within system (4), we plot the bifurcation diagrams in Figure 15(a) when $[q_1, q_2, q_3] = [0.8, 0.90, 0.91]$, $a = 1$, and $b = -9$. It is worth noting that two sets of ICs are considered to perform the simulation of this figure. The first one is $(x_0, y_0, z_0) = (-1, 1, 1)$ which has been taken into account for the red plot, while the second one is $(x_0, y_0, z_0) = (0.1, 1.2, 2)$ which has been taken into account for the blue plot. These two plots show certainly that system (4) will exhibit bistability phenomenon if the incommensurate fractional-order value is increased. For $a = 1$, $b = -9$, and $[q_1, q_2, q_3] = [0.80, 0.90, 0.91]$, both coexisting attractors of this system are drawn in Figure 15(b) according to the two ICs, $(x_0, y_0, z_0) = (-1, 1, 1)$ for the red plot and $(x_0, y_0, z_0) = (0.1, 1.2, 2)$ for the blue plot. Furthermore, the basin of attraction of system (4) can be shown in Figure 16 which corresponds to Figures 15(a) and 15(b), where the ICs shown in yellow and blue regions lead to the limit cycles and to the chaotic attractors, respectively.

4. Variable-Boostable Attractors of the Incommensurate Fractional-Order Model

With the aim of accomplishing the complete range of the signal's linear transformations, the offset boosting might be joined with amplitude control. It was reported in [27] that insertion of a new developed boosting controller may destroy the symmetry of the variable-boostable model. From this perspective, we will add to the system states x, y, and z three additional controlled scalers m, n, and k, respectively. In view of this addition, system (4) will be turned into the following form:

$$\begin{cases} D^{q_1} x(t) = -(x(t) + m) + 2\sin(x(t) + m) + \sin(y(t) + n) + b\sin(z(t) + k), \\ D^{q_2} y(t) = -(y(t) + n) + b\sin(x(t) + m) + 2\sin(y(t) + n) + a\sin(z(t) + k), \\ D^{q_3} z(t) = -(z(t) + k) + a\sin(x(t) + m) + b\sin(y(t) + n) + 2\sin(z(t) + k). \end{cases} \tag{6}$$

In the following sections, we intend to address system (6) in light of three different cases for the system parameters $a = 1$ and $b = -9$. Besides, we will select the incommensurate fractional-order values as $[q_1, q_2, q_3] = [0.80, 0.90, 0.91]$. As a remark about the choice of initial conditions, it should be noted that when the system has unbounded solutions, the variable boosting should be accompanied with a modification of the initial conditions, while the initial conditions can be ignored in the systems with global attraction.

4.1. A Line of Variable Attractors. A variable chaotic attractor can be scattered along the 1D line especially when we are carrying out a certain control so that the offset boosting parameters take the following cases:

(i) When the parameter m is frequently varied and the other two parameters are kept as $n = k = 0$, we will gain several variable chaotic attractors scattered on the x-axis as exhibited in Figure 17(a)

(ii) When the parameter n is frequently varied and the other two parameters are kept as $m = k = 0$, we will gain several variable chaotic attractors scattered on the y-axis as exhibited in Figure 17(b)

(iii) When the parameter k is frequently varied and the other two parameters are kept as $m = n = 0$, we will gain several variable chaotic attractors scattered on the z-axis as exhibited in Figure 17(c)

4.2. A Lattice of Variable Attractors. Herein, two controlled parameters will be simultaneously adjusted and the other

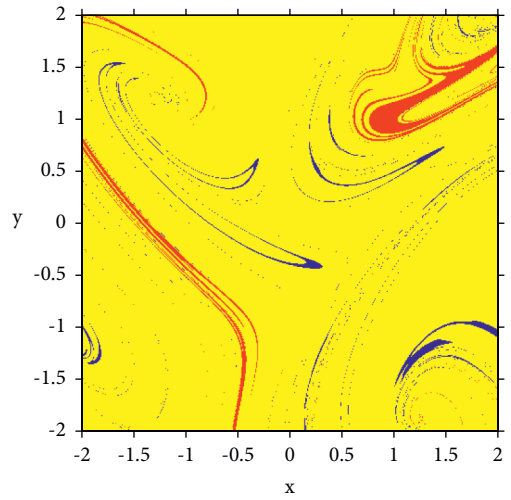


FIGURE 16: Basin of attraction of system (4) when $[q_1, q_2, q_3] = [0.80, 0.90, 0.91]$ according to the system parameters $a = 1$ and $b = -9$ and to the initial condition of the third state variable $z = 0$.

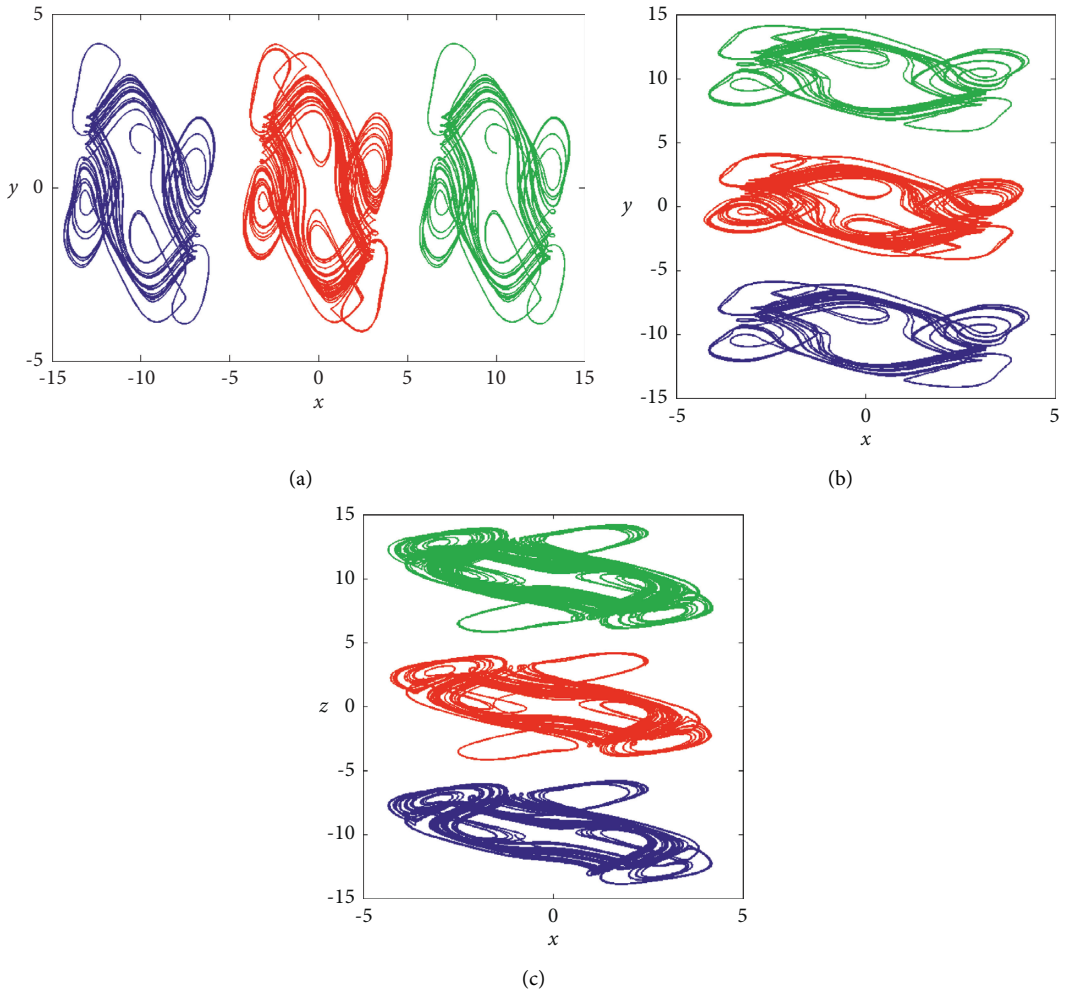


FIGURE 17: Scattering of the variable chaotic attractor on a 1D line for $a = 1$, $b = -9$, and $[q_1, q_2, q_3] = [0.80, 0.90, 0.91]$. (a) x-line when $m = \pm 10$ and $m = 0$, (b) y-line when $n = \pm 10$ and $n = 0$, and (c) z-line when $k = \pm 10$ and $k = 0$.

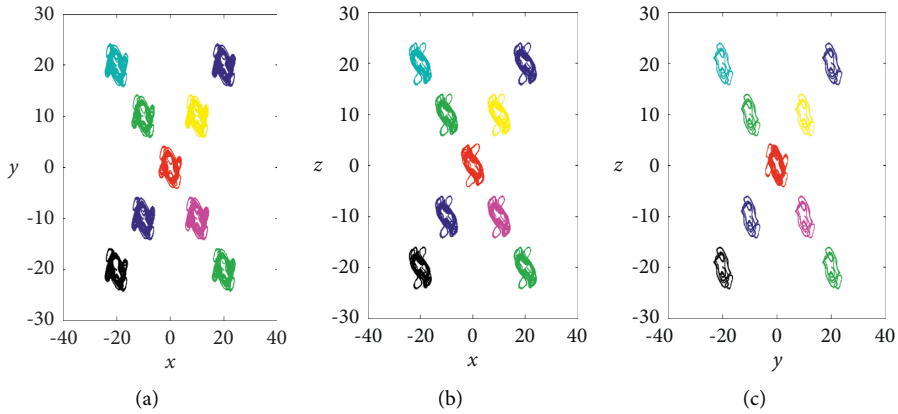


Figure 18: A 2D lattice of a variable attractor for incommensurate fractional order $[q_1, q_2, q_3] = [0.80, 0.90, 0.91]$ with $a = 1$ and $b = -9$. (a)xy lattice for $(m, n) = (0, 0)$, $(10, 10)$, $(10, -10)$, $(-10, 10)$, $(-10, -10)$, $(20, 20)$, $(20, -20)$, $(-20, 20)$, $(-20, -20)$; (b)xz lattice for $(m, k) = (1, 1)$, $(10, 10)$, $(10, -10)$, $(-10, 10)$, $(-10, -10)$, $(20, 20)$, $(20, -20)$, $(-20, 20)$, $(-20, -20)$; and (c)yz lattice for $(n, k) = (1, 1)$, $(10, 10)$, $(10, -10)$, $(-10, 10)$, $(-10, -10)$, $(20, 20)$, $(20, -20)$, $(-20, 20)$, $(-20, -20)$.

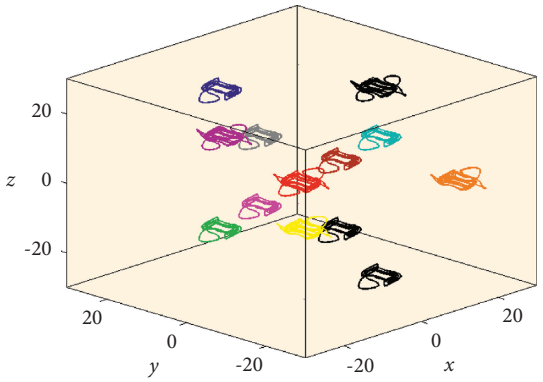


Figure 19: A 3D grid of a variable periodic and chaotic attractors when $[q_1, q_2, q_3] = [0.80, 0.90, 0.91]$ with $a = 1$ and $b = -9$ for $(m, n, k) = (0, 0, 0)$, $(0, 10, 10)$, $(0, 10, -10)$, $(0, -10, 10)$, $(0, -10, -10)$, $(0, 20, 20)$, $(0, 20, -20)$, $(0, -20, 20)$, $(20, 20, -20)$, $(-20, 0, -20)$, $(20, 0, -20)$, $(20, 20, 0)$, $(-20, 20, 0)$.

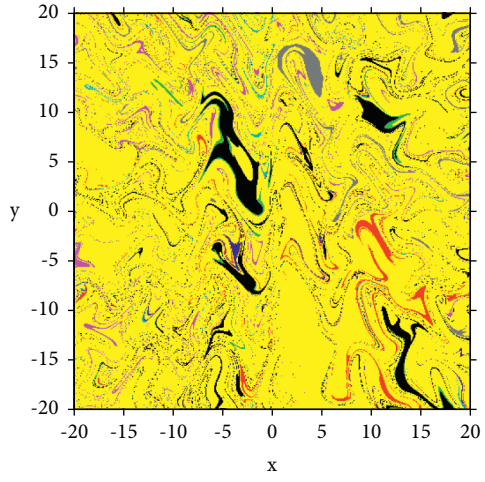


Figure 20: Basin of attraction of system (4) when $[q_1, q_2, q_3] = [0.80, 0.90, 0.91]$ according to the system parameters $a = 1$ and $b = -9$ and to the initial condition of the third state variable $z = 0$. The colors in this figure associate with the colors of the attractors given in Figure 19.

parameter will be kept at zero. This would yield a 2D lattice of variable chaotic and periodic attractors. In particular, one might consider the following three cases:

(i) When the two parameters m and n are frequently varied and the third parameter is kept as $k = 0$, we will gain several variable chaotic attractors scattered on the xy lattice as exhibited in Figure 18(a)

(ii) When the two parameters m and k are frequently varied and the third parameter is kept as $n = 0$, we will gain several variable chaotic attractors scattered on the xz lattice as exhibited in Figure 18(b)

(iii) When the two parameters n and k are frequently varied and the third parameter is kept as $m = 0$, we will gain several variable chaotic attractors scattered on the yz lattice as exhibited in Figure 18(c)

4.3. A 3D Grid of Variable Attractors. In this section, certain values of the three controlled parameters m, n, and k will be simultaneously adapted. This would yield, after taking the fractional-order values as, e.g., $[q_1, q_2, q_3] = [0.80, 0.90, 0.91]$, a 3D grid of variable chaotic and periodic attractors as illustrated in Figure 19. Furthermore, the basin of attraction of system (4) is plotted and shown in Figure 20 which corresponds to Figure 19, where the ICs shown in yellow and blue regions lead to the limit cycles and to the two chaotic attractors, respectively.

5. Conclusion

This work has formulated a novel version of Hopfield neural network models with incommensurate fractional orders using the Caputo differential operator. Through continuous variation of the values of the system parameters as well as the fractional-order derivative values, the stability of the proposed model has been analyzed numerically, and many rich complex dynamics, including symmetry, bistability, and coexisting chaotic attractors,

have been generated. It turned out that through adapting certain additional controlled constants the proposed model possesses the offset boosting of three variables. In addition, it has been shown that the resultant periodic and chaotic attractors generated from such models can be distributed in several forms, including 1D line, 2D lattice, and 3D grid and even in an arbitrary location of the phase space.

Data Availability

The data that support the findings of this study are available upon request to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The research was supported by the Taif University Researchers Supporting Project Number TURSP-2020/229, Taif University, Taif, Saudi Arabia.

References

[1] J. Fleck, "Development and establishment in artificial intelligence," *The Question of Artificial Intelligence*, vol. 16, pp. 106–164, 2018.

[2] S. Hayman, "The McCulloch-Pitts model,"vol. 6, pp. 4438-4439, in *Proceedings of the 1999 International Joint Conference on Neural Networks (IJCNN'99)*, vol. 6, pp. 4438-4439, IEEE, Washington, DC, USA, 1999.

[3] S. Zhang, Y. Yu, and H. Wang, "Mittag-Leffler stability of fractional-order Hopfield neural networks," *Nonlinear Analysis: Hybrid Systems*, vol. 16, pp. 104–121, 2014.

[4] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of the National Academy of Sciences*, vol. 79, no. 8, pp. 2554–2558, 1982.

[5] I. M. Batiha, R. B. Albadarneh, S. Momani, and I. H. Jebril, "Dynamics analysis of fractional-order Hopfield neural networks," *International Journal of Biomathematics*, vol. 13, no. 8, Article ID 2050083, 2020.

[6] H.-P. Hu, J.-K. Wang, and F.-L. Xie, "Dynamics analysis of a new fractional-order Hopfield neural network with delay and its generalized projective synchronization," *Entropy*, vol. 21, no. 1, p. 1, 2019.

[7] H. Wang, Y. Yu, G. Wen, S. Zhang, and J. Yu, "Global stability analysis of fractional-order Hopfield neural networks with time delay," *Neurocomputing*, vol. 154, pp. 15–23, 2015.

[8] A. Boroomand and M. B. Menhaj, "Fractional-order Hopfield neural networks," *Lecture Notes in Computer Science*, Springer, vol. 5506, pp. 883–890, , Berlin, Germany, 2009.

[9] E. Kaslik and S. Sivasundaram, "Nonlinear dynamics and chaos in fractional-order neural networks," *Neural Networks*, vol. 32, pp. 245–256, 2012.

[10] B. N. Lundstrom, M. H. Higgs, W. J. Spain, and A. L. Fairhall, "Fractional differentiation by neocortical pyramidal neurons," *Nature Neuroscience*, vol. 11, pp. 1335–1342, 2008.

[11] P. Arena, L. Fortuna, and D. Porto, "Chaotic behavior in noninteger-order cellular neural networks," *Physical Review E*, vol. 61, no. 1, pp. 776–781, 2000.

[12] T. Matsuzaki and M. Nakagawa, "A chaos neuron model with fractional differential equation," *Journal of the Physical Society of Japan*, vol. 72, no. 10, pp. 2678–2684, 2003.

[13] I. Petras, "A note on the fractional-order cellular neural networks," in *Proceedings of the 2006 IEEE International Joint Conference on Neural Network*, pp. 1021–1024, Vancouver, Canada, 2006.

[14] H. Zhu, S. Zhou, and W. Zhang, "Chaos and synchronization of time delayed fractional neuron network system," in *Proceedings of the 2008 9th International Conference for Young Computer Scientists*, pp. 2937–2941, Changsha, China, 2008.

[15] S. Zhou, H. Li, and Z. Zhu, "Chaos control and synchronization in a fractional neuron network system," *Chaos, Solitons & Fractals*, vol. 36, no. 4, pp. 973–984, 2008.

[16] S. Zhou, P. Hu, and H. Li, "Chaotic synchronization of a fractional neuron network system with time-varying delays," in *Proceedings of the 2009 International Conference on Communications, Circuits and Systems*, pp. 863–867, Milpitas, CA, USA, 2009.

[17] R.-C. Wu, X.-D. Hei, and L.-P. Chen, "Finite-time stability of fractional-order neural networks with delay," *Communications in Theoretical Physics*, vol. 60, no. 2, pp. 189–193, 2013.

[18] J. Yu, C. Hu, and H. Jiang, "α-stability and α-synchronization for fractional-order neural networks," *Neural Networks*, vol. 5, pp. 82–87, 2012.

[19] L. Chen, Y. Chai, and R. Wu, "Linear matrix inequality criteria for robust synchronization of uncertain fractional-order chaotic systems," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 21, no. 4, Article ID 043107, 2011.

[20] I. Podlubny, *Fractional Differential Equations*, Academic Press, New York, NY, USA, 1999.

[21] K. Diethelm, N. J. Ford, and A. D. Freed, "A predictor-corrector approach for the numerical solution of fractional differential equations," *Nonlinear Dynamics*, vol. 29, pp. 3–22, 2002.

[22] D. Baleanu, K. Diethelm, E. Scalas, and J. Trujillo, *Fractional Calculus Models and Numerical Methods*, World Scientific Publishing Company, Singapore, 2012.

[23] K. Diethelm, "An algorithm for the numerical solution of differential equations of fractional order," *Electronic Transactions on Numerical Analysis*, vol. 5, pp. 1–6, 1997.

[24] K. Diethelm and N. J. Ford, "Analysis of fractional differential equations," *Journal of Mathematical Analysis and Applications*, vol. 265, no. 2, pp. 229–248, 2002.

[25] C. Li, J. C. Sprott, and H. Xing, "Constructing chaotic systems with conditional symmetry," *Nonlinear Dynamics*, vol. 87, no. 2, pp. 1351–1358, 2017.

[26] N. Debbouche, S. Momani, A. Ouannas et al., "Generating multidirectional variable hidden attractors via newly commensurate and incommensurate non-equilibrium fractional-order chaotic systems," *Entropy*, vol. 23, no. 3, p. 261, 2021.

[27] C. Li and J. C. Sprott, "Variable-boostable chaotic flows," *Optik*, vol. 127, no. 22, pp. 10389–10398, 2016.

Hindawi
Complexity
Volume 2021, Article ID 8667868, 10 pages
https://doi.org/10.1155/2021/8667868

Research Article

Classification of Long-Tailed Data Based on Bilateral-Branch Generative Network with Time-Supervised Strategy

Yalin Huang, Yan-Hui Zhu, Zeng Zhigao, Yangkang Ou, and Lingwei Kong

School of Computer Science, Hunan University of Technology, Zhuzhou, Hunan 412008, China

Correspondence should be addressed to Yan-Hui Zhu; swayhzhu@163.com

Received 4 September 2021; Accepted 20 October 2021; Published 2 November 2021

Academic Editor: Hou-Sheng Su

In the face of the long-tailed data distribution that widely exists in real-world datasets, this paper proposes a bilateral-branch generative network model. The data of the second branch is constructed by resampling the generative network training method to improve the data quality. A bilateral-branch network model is used to curb the risk of gradient explosion and to avoid over-fitting and under-fitting with the combined effect of different data branches. Meanwhile, Time-supervised strategy is introduced to improve the model's operational efficiency and ability to cope with extreme conditions by supervising and collaboratively controlling of the bilateral-branch generative network with time-invariant parameters. Time supervised strategy could ensure the accuracy of the model while reducing the number of iterations. Experimental results on two publicly available datasets, CIFAR10 and CIFAR100, show that the proposed method effectively improves the performance of long-tail data classification.

1. Introduction

With the rapid development of convolution neural network algorithms in recent years, there has been a very impressive improvement in the performance of image classification. Undoubtedly, the success is inextricably linked to the available high-quality large-scale datasets, such as ImageNet ILSVRC 2012 [1], MS COCO [2] and Places Database [3]. However, compared to these high-quality datasets, real-world datasets are always biased and it is difficult to ensure a uniform distribution of data, and more often than not, certain classes of data are very abundant while certain remaining classes are very scarce, which leads to a long-tail distribution of data [4, 5] and affects the performance of image classification. From the reviewed related materials, class re-balancing strategies are currently often used when faced with uneven data distribution. Class re-balancing strategies are further divided into two categories, namely re-sampling strategies [6–13] and re-weighting strategies [14–17]. The re-sampling strategy means that the source data is resampled according to the desired frequency distribution to obtain a new datasets by copying the minority data [6, 8, 12, 18] and reduce the majority data [7, 12, 13]. This strategy is able to reduce error rate caused by unbalanced data during training. Although resampling can show better results, this strategy still has a negative effect on the model. For example, the SMOTE algorithm [11] merely repeats and abandons the original data in the process of resampling. Although it changes the data distribution, it cannot bring more classified information to the deep learning model. Therefore classical resampling tends to make the tail data more prone to over-fitting conditions, while the head data is also tend to under-fitting conditions.

And this can be effectively avoided by a re-weighting method that adds a regularization term to the loss. Where the loss of the regularization parameter can often be expressed in

$$\text{Loss} = \text{loss}_1 + \lambda * \text{loss}_2 \quad (0 \le \lambda \le 1). \tag{1}$$

The over-fitting and under-fitting of the model is suppressed by introducing a new loss function as a constraint. However, the regularization method also has its drawbacks. Due to the introduction of the regularization parameter as a restriction in the loss, it sometimes makes the model parameters fail to converge, and in extreme cases, it even results in gradient explosion. Therefore some scholars proposed many

ways to prevent the over-fiting and under-fiting with data-dependent regularizer [19–22]. For example, the algorithms proposed by Zhou et al. [23] based on bilateral-branch network model has good accuracy in the classification problem of long-tailed data. The algorithm obtained the first place in iNaturalist2019 with an error rate of 30.38% on the iNaturalist2018 public datasets. The model of Boyan Zhou re-weights the loss function by a data-driven approach. This data-driven approach not only effectively curbs the risk of gradient explosion, the model will converge more easily under the combined effect of different data branches while avoiding the over-fitting phenomenon on tail data [4]. However, this method requires very strict data quality. Since the regularization term itself is data-driven approach, if the data quality of the second branch is poor, it will lead to loss_2 adverse effect on the direction of the overall model learning. Therefore how to construct high quality second branch with different distributions becomes a new problem.

Many works focus on minority samples, the re-balance the data by augmentation of minority samples [11, 24, 25]. while some other works provide a Major-to-minor translation to re-balance the data distribution [26–28]. the M2m algorithm proposed by Kim et al. [29], which is different from the above methods and gives a solution from the perspective of resampling. M2m algorithm is a method of resampling data through data generation by generation against network [26, 30, 31]. This method solves the problem of unbalanced data distribution by constructing fewer classes of data through multiple classes of data, while bringing more classification information to the data and improving the data quality. This method can effectively increase the number of minority class samples while greatly reducing the over-fitting of the minority class data. However, this method often requires the use of an already pre-trained network, and generating data during training still requires a large number of iterations, which will reduce the efficiency of model learning.

The main contributions of this paper are as follows.

(1) Incorporating the respective advantages of the Bilateral-branch model and the generative network model, this paper proposes a bilateral-branch generative network model. The data of the second branch is constructed by resampling and generating by the generative network training method to improve the data quality. The bilateral-branch network model is used to curb the risk of gradient explosion, and the model is made to avoid over-fitting and under-fitting phenomena under the combined effect of different data branches to improve the effect of long-tail data classification.

(2) Since the generative network model adds a large number of iterations in the process of generating data, which affects the efficiency of the model, this paper introduces a time-supervised strategy, which supervises and limits the number of iterations of the generative network through time-variant parameters, improves the operational efficiency of the model and its ability to cope with extreme conditions, and ensures the accuracy of the model while reducing the number of iterations.

(3) The accuracy of the algorithm was tested under two publicly available datasets, CIFAR10 and CIFAR100 [32], for different distributions, and the method used in this paper, was higher than both Boyan Zhou' algorithm and M2m algorithm.

2. Construction of a Bilateral-Branch Generative Network Model

The overall framework of the bilateral-branch generative network model proposed in this paper is shown in Figure 1, and the model is abbreviated as BBGN (Bilateral-Branch Generative Network).

As shown in the figure the BBGN model first constructs the inverse distribution data branches by data resampling. We refer to the source data branch as the first branch data, and the inverse distribution data constructed by the resampling method as the second branch. After the data of the second branch is generated by the re-sampling method, the first branch data is used as the data source to generate new high-quality data for replacing the data of the second branch by the generating network (GN) module to perform data augmentation on the data of the second branch. In the GN module and later in the network model we introduce respectively the time-supervised parameters $(1 - \alpha)$ with α for coordinate and control these two functional modules.

After the process of data augmentation is completed, the original two data branches are feature extracted using a pyramid-shaped multiple layer feature perceptron to form separate sets of features. Through the time-supervised parameterα of control, the two feature sets are fused to form a one-class feature set. The fused features are averaged pooled (GAP layer in the figure) and pushed into the fully connected classifier for classification. The model loss is calculated and the model parameter weights are updated backwards based on the classification results and the loss function.

Through the coordination and control of the time-supervision strategy, in the early stage of BBGN model learning, the BBN network is influenced by the time-supervision parameter, which makes the data of the second branch play almost no role in the learning of the model, while the GN network is also subject to the synergy of this parameter and does almost no data enhancement to the data of the second branch. At this point the data of the second branch plays a weak role in the learning of the model only by resampling the data, and the learning process of the model is dominated by the source data of the first branch. Over time, under the influence of the time-supervised strategy, the influence of the BBN model's bilateral-branch data on model learning gradually changes. The role of the influence of the data of the second branch gradually increases. At the same time the GN network is gradually activated by the coordination of the temporally supervised parameters, and the resampled data of the second branch starts to play its role in the model learning process after being enhanced by the GN

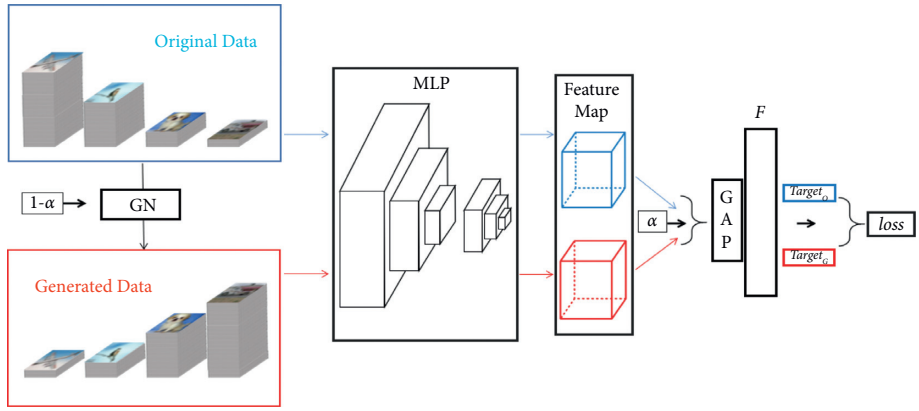


Figure 1: Overall framework of the bilateral-branch generative network model.

module. The introduction of the time-supervised strategy enables our algorithm to maintain accuracy while reducing the number of iterations and greatly improving the computational efficiency.

2.1. Bilateral-Branch Network (BBN) Model. The BBN model is data-driven and uses a loss function weighting method to avoid over- and under-fitting of the data.

The BBN model starts from a regularized weighting perspective, but unlike the usual method of constructing regularized expressions, this model achieves regularized weighting by using branches of data from different distributions to learn and fuse the losses generated by multiple branches. The pseudocode of the BBN model algorithm is as follows.

At the row 14 of Algorithm 1, we get the $\text{Target}_{\text{pred}}$ of model by mixing the feature_1 and the feature_2. This combiner function canbe expressed by the following equation (6). And then, at the row 18 of Algorithm 1 the BBN model Combiner2 function can be expressed by the following equation (2).

$$\text{Loss} = \alpha * \text{loss}_1 + (1 - \alpha) * \text{loss}_2 \quad (0 \le \alpha \le 1), \tag{2}$$

where loss_1 is the error generated in the learning process for the first branch data in the BBN model, and loss_2 is the error generated by the data of the second branch in the learning process. At the row 16 and 17 of Algorithm 1, the $\mathscr{L}$ function can be expressed by the following equation (8), and Loss is the total error generated by the BBN model in the learning process. Loss affects the learning process of the model through the synthesis of the error components of the two data.

2.2. Generating Network Models. The pseudocode of our proposed algorithm for generating models in BBGN is shown below.

At the first row of the Algorithm 2, the Bernoulli distribution function used to select the class to be generated is shown in .

$$P_G = p^{N_{\text{Target}_G}/N_0} * (1 - p)^{1-\left(N_{\text{Target}_G}/N0\right)}, \tag{3}$$

where Target_G denotes the class to be generated, and N_{Target_G} denotes the total number of samples of the class to be generated, and N_0 denotes the total sample size of the class corresponding to the class with the highest frequency, and P_G denotes the probability that such class is selected as the class to be generated.

At the row 3 fo Algorithm 2.

At the row 5 of Algorithm 2, the generation source class Target_O are selected using the distribution function as shown in .

$$P_O = 1 - \beta^{N_{\text{Target}_O} - N_{\text{Target}_G}}, \tag{4}$$

where $\beta \in (0, 1)$ is a fixed parameter, and N_{Target_O} denotes the total number of samples of the generated source class, and N_{Target_G} denotes the total number of samples of the class to be generated, and P_O denotes the probability that such class is selected as the generating source class.

And then we need to select the generation source based on the class to be generated. Since this data augmentation algorithm constructs minority class samples by extracting class-independent features from the majority class and class-related features from the minority class. Therefore the method has certain requirements on the data quality of the generation source data, and the class-irrelevant feature classes contained in the data should be as rich as possible, so as to improve the generalization ability of the model for minority class learning.

In this Algorithm 2, the class to be generated Target_G and the generated source class Target_O are selected by resampling the class label with the probability distribution. This approach can meet our requirements on the quality of the generated source data. At the same time, the possibility of generating fewer classes of data with fewer classes of data is preserved, and the choice of generating source class data is expanded.

The framework diagram of the generative network used for Algorithm 2 is shown in Figure 2.

For the selected image I, input the to-be--trained model F to obtain the I's classification result Label_F, and compute

Inputs: training datasets I_1, $\lambda, \gamma, \eta, L > 0, \beta \in [0, 1), \alpha$
Output: BBN network model F (input; weight)

$I_2 \leftarrow I_1$
for $k = 2$ to K do
  $\Delta \leftarrow (N_1 - N_k)$
  for $i = 1$ to Δ do
    $x \leftarrow$ A random sample of class k in I_2
    $I_2 = I_2 \cup x$
  end for
end for
$I_2 =$ Algorithm 2 $(I1, I2, \lambda, \gamma, \eta, L > 0, \beta \in [0, 1), \alpha)$
for $i = 0$ to N_{I_1} do
  $j \leftarrow$ A random sample of I_2
  $\text{feature}_1 = \digamma(I_{1i}, W^T)$
  $\text{feature}_2 = \digamma(I_{2i}, W^T)$
  $\text{Target}_{\text{pred}} = \text{Combiner1}(\text{feature}_1, \text{feature}_2, \alpha)$
  $\text{loss}_1 = \mathscr{L}(\text{Target}_{1i}, \text{Target}_{\text{pred}})$
  $\text{loss}_1 = \mathscr{L}(\text{Target}_{2j}, \text{Target}_{\text{pred}})$
  $\text{Loss} = \text{Combiner2}(\text{loss}_1, \text{loss}_2, \alpha)$
  $W \leftarrow W - \text{Loss}$
end for
$F \leftarrow W$

Algorithm 1: BBN algorithm.

Input: An unbalanced data branch I_1. A resampled data branch I_2. A network f. A pre-trained generative network G. $\lambda, \gamma, \eta, L > 0, \beta \in [0, 1)$
Output: A augmented balanced data branch O

$\text{Target}_G \leftarrow$ A random sample of class K in I_2 with P_G
for $k =$ in Target_G do
  $\Delta \leftarrow (N_1 - N_k) * T(t, 1 - \alpha)$
  for $i = 1$ to Δ do
    $\text{Target}_O \leftarrow$ A random sample of class k in I_1 with P_O
    $I \leftarrow$ A random sample of class Target_O in I_1
    Add some noise σ to the I
    for $t = 1$ to L do
    $\text{Loss}_G = \mathscr{L}(G; I; k)$
    $\text{Loss}_O = f_{\text{Target}_O}(I)$
    $\text{loss}_{GN} \leftarrow \nabla_R[\text{Loss}_G + \lambda * \text{Loss}_O]$
    $I \leftarrow I - \eta * (\text{loss}_{GN}/\text{loss}_{GN2})$
    end for
    $P \sim \text{Bernoulli}(\beta^{(N_{k_0} - N_k)^+})$
    If $\mathscr{L}(g; I; k) > \gamma$ or $P = 1$ **then**
    $I_2^* \leftarrow$ A random sample of class k in I_2
    end if
    $I_2^* \leftarrow I$
  end for
end for

Algorithm 2: Data generation algorithm.

this result with Label_F and the generated source class label Target_O. The distance between this result and the generated source class label Target_O is used as the regularization term loss_O in loss function. Classify I using the feature information provided by the pre-trained model G to obtain the result Label_G. Use the loss function to calculate the loss_g by Label_G and the target to be generated Target_G. calculate the total loss loss_{GN} using equation (5), and backward update the pixel of selected image I based on the total error.

$$\text{loss}_{GN} = \text{Loss}_G + \lambda * \text{Loss}_O. \tag{5}$$

In the updating process, the goal of loss_g is to cause the selected image I to be classified as the class Target_G with the pre-trained network G. Since the pre-trained network G

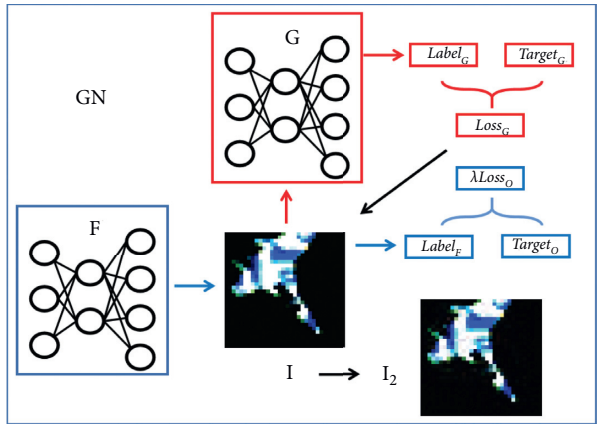


Figure 2: Framework diagram of the generation network.

contains the class-related feature information of each class, the $loss_g$ guided by the class-related feature information of the network G, could gradually eliminate the class-related features of the class $Target_O$ in the image I and add the class-related features of the class $Target_G$. The result image generated by this generative network algorithm consists of the class-independent features of the $classTarget_O$ and the class-irrelevant features of the class $Target_G$. By this way we could generated the new image of the class $Target_G$.

In the unbalanced datasets, the majority class has a very rich set of class-irrelevant features in addition to such class-relevant features. While the minority class has only class-related features with a small number of class-irrelevant features. Therefore the class irrelevant features carried in the majority classes are combined with the class relevant features carried in the minority classes in the pre-trained network G through the generative network to generate a large number of images of the minority classes. New classification information is introduced while balancing the data distribution.

Also, in order to prevent the situation that the gradient disappears during the training process and the target image cannot be generated, the regularization term can be set to eliminate this situation. The above $loss_O$ represents the prediction result of image I in the network F to be trained $Label_F$ with the image to be generated in the source $classTarget_O$ of the difference. Before the classification is performed, the $Label_F$ is in the form of a 0-1 vector, and the respective values on the vector indicate the probability of the image being of this class. By setting this difference as a regularization term, the gradient disappearance can be effectively prevented, allowing the input image I to be generated more easily as an image of other classes.

3. Time Supervision Strategy

The variation of the time-supervised parameters with training time is shown in Figure 3.

The temporal supervision strategy proposed in this paper acts on both the BBN model and the data generation module. The time-supervised strategy coordinates and controls the overall model learning process by setting the time-supervised parameters α.

3.1. Time-Supervision Strategies for BBN Models. In our BBGN model, the degree of influence of different branches on the learning process of the model can be adjusted by introducing time-supervised parameters in the computation process of the fusion and loss of the two branches. The first branch data is the original real data, while the data of the second branch enhanced by the generative network is the non-real data generated by data features. Therefore, in the early stage of model training, the real data should be the main focus, and the real features in the data should be learned as much as possible. And when the minority class gradually starts to over-fit, then gradually start to increase the degree of influence of the non-real data generated by the features on the model learning to prevent the model from over-fitting on the minority class, and at the same time enhance the generalization ability of the model on the minority class to improve the accuracy of the model.

The fused feature is obtained during the fusion of model features by .

$$\text{Feature} = \alpha * (\text{feature}_1) \cup (1-\alpha) * (\text{feature}_2). \tag{6}$$

It is important to note that when we use the results obtained from fused feature with classifier when calculating the loss function, it needs to be compared the result label to the two objectives $Target_O$ with the $Target_G$ together to calculate the loss. Since the feature fusion is performed with the time-supervised parameter α, the loss should also be calculated with the time-supervised param-eter α to determine the final loss. The loss calculation in this paper is shown in equation (7).

$$\text{Loss} = \alpha * \mathscr{L}(\text{Label}, \text{Target}_O) + (1-\alpha) * \mathscr{L}(\text{Label}, \text{Target}_G), \tag{7}$$

where the loss function $\mathscr{L}(L, T)$ is calculated as shown in .

$$\mathscr{L}[L,T] = -\sum_i T_i * \ln(L_i), \tag{8}$$

L_i denotes the ith vector in Label, representing the probability that the input is of class i. T_i denotes the ith vector in Target, where only one vector is 1 indicating that the input is of class i and the rest of the vectors are 0.

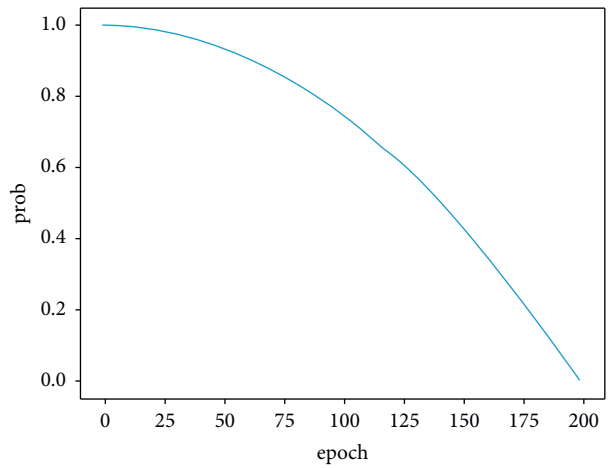


Figure 3: Curve of the change in the parameters of time supervision.

3.2. Time-Supervision Strategies in Generative Network Models. The generative network model improves the performance of the model by improving the quality of the data of the second branch. However in the early stages, the data of the second branch has little impact on the learning process of model learning, and the large amount of generated data does not contribute to the model learning process. This would generate a waste of resources. So we use equation (9) and introduce a time-supervised strategy to limit the size of generated data in the early stage.

$$L = N_{\text{Target}_G} * (1 - \alpha), \tag{9}$$

where L is the number of samples to be generated, and N_{Target_G} is the total number of classes to be generated, and α is the time supervision parameter.

By the introduction of the time-supervised strategy, the amount of generation of the generative network is limited in the early stage of model training. And with the increase of the number of iterations, the size of generated data L of the generative network is gradually activated by the time-supervised strategy, when the main component of the data of the second branch is changed from resampled data to generated data, and the data quality is gradually improved. In this way, the number of iterations can be greatly reduced and the learning efficiency can be improved while ensuring the accuracy of the model.

4. Experimentation and Analysis

4.1. Experimental Data Set. In this paper, we use online public datasets CIFAR10 and CIFAR100 [32], both of which contain 60,000 RGB color images of $32 * 32$ size. There are 50,000 images in these images for training and 10,000 images for testing. CIFAR10 and CIFAR100 have 10 classes and 100 classes of data respectively. The source datasets is a uniform datasets. And in this paper, we resample the long-tail datasets with class imbalance of CIFAR10 and CIFAR100 by setting the imbalance ratio parameter Ratio during the experiment. Among Ratio = $N_{\max}/N_{\min}$, in this paper, Ratio contains three values of 10,50,100 respectively.

4.2. Introduction of Experimentation. Our experimentation are running in the ubuntu20.04 operation system. Traning the model need about 16G RAM. The machine language of the experimentation is python. In the experimentation, at first we will train our model with pretrained model and unbalanced trainning datasets. Then we test the result of our model in the testing dataset. Finally we visualize our result and analyze it.

4.3. Analysis of Experimental Results. In this paper, the accuracy of the optimal solution on the test set is used to compare the classification ability of different methods by comparing the experimental results. Also, we track the accuracy of the classification results of different models on the test set during the training process and plot it as a line graph, which can visualize the difference of different models during the training process.

4.3.1. Results of the CIFAR10 Experiment. From Table 1 and Figure 4, we can find that the accuracy of BBN algorithm and our algorithm decreases when the number of training is between 50 and 75, while M2m algorithm can quickly improve to a higher level and stabilize at that level in the early stage. However, as the number of training gradually increases, the improvement of M2m accuracy gradually becomes slower, while the accuracy of BBN algorithm and our algorithm improves rapidly. Considering that both the BBN algorithm and our algorithm contain a bilateral-branch structure, the decrease in the accuracy of the algorithm is related to the gradual increase in the influence of the second branch. At the beginning, the accuracy drops significantly due to the small influence of the data in the second branch of the model, but as the number of training sessions increases, the accuracy of the BBN algorithm and our algorithm always steadily exceeds that of the M2m algorithm with the effect of the data in the second branch. This indicates that the pure data generation model can achieve a high accuracy rate within a smaller number of training sessions, but the final accuracy ceiling of the model is not high. Compared to the BBN algorithm, the lead is weaker on the CIFAR10 datasets

Table 1: The accuracy of resnet-32 on long-tailed cifar10.

Methods	CIFAR100		
	$N1/N100 = 10$	$N1/N100 = 50$	$N1/N100 = 100$
M2M	57.55	45.63	42.16
BBN	59.47	45.90	42.47
Methods (ours)	60.14	47.06	42.64

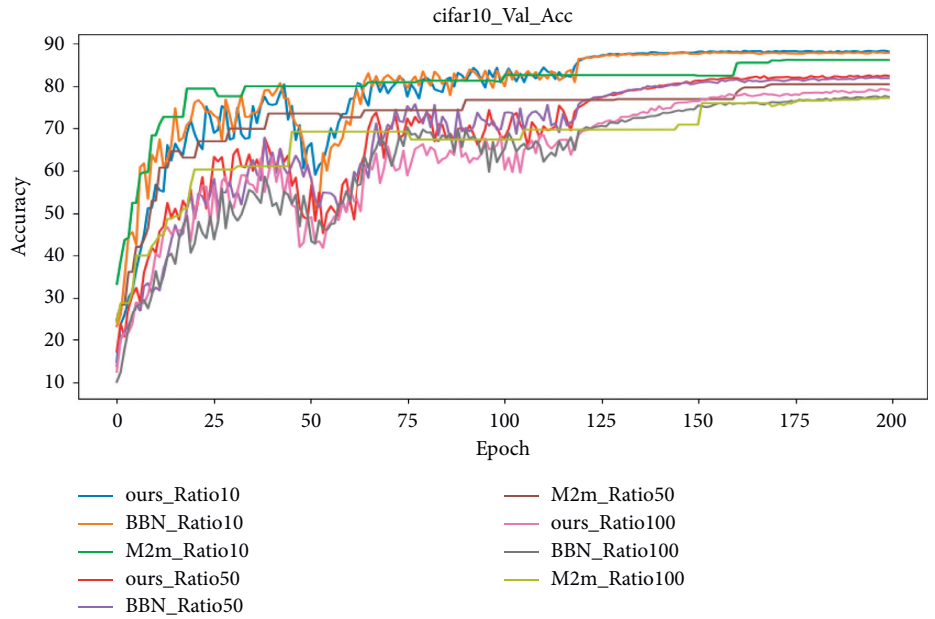


Figure 4: Accuracy variation curve of cifar10 long-tail datasets.

when the imbalance ratio is 10 versus 50, although our method outperforms the BBN algorithm. But when faced with an extreme imbalance situation, i.e., when the imbalance ratio is 100, the advantage of our algorithm is significantly improved. This shows that when the realistic environment is not complex enough, the data augmentation by resampling alone can achieve better results, but in the extreme environment, the data quality of the second branch will be very poor, and thus better results can be obtained by introducing data generation to improve the data quality.

4.3.2. Results of the CIFAR100 Experiment. The results of the CIFAR100 experiments are shown in Table 2 and Figure 5. It can be found that the curves are roughly the same as cifar10, the M2m accuracy stabilizes to a high value very quickly, while the BBN algorithm with our algorithm takes the lead in decreasing and then improving as the number of training increases. The CIFAR100 datasets is 10 times more classified and has less datasets in each class than CIFAR10, so obtaining a higher accuracy is very difficult. In this extreme condition, our algorithm can take advantage of the data enhancement and lead the BBN and M2m algorithms. However, surprisingly, in the most extreme case of Ratio = 100 for CIFAR100, the accuracy rates of all three are comparable and lower, and the accuracy curves are almost completely indistinguishable. This is perhaps due to the fact that in the too extreme case, the ResNet32 network leads to this result due to the limitations of its network structure.

To test this conjecture, we did a set of comparison experiments on the CIFAR100 datasets with Ratio = 100, using the ResNet50 network.

4.3.3. Results of the ResNet50 Experiment. The experimental results are shown in Figure 6. From the accuracy curves, we can find that when the model has more parameters and the network structure is more complex, the accuracy results of our method and the BBN method produce very obvious differences. Benefiting from the role of the data generation network, the BBGN network model constructed by the generation method with the time-supervision strategy can also significantly improve the accuracy of the model in the case of very extreme data distribution.

Table 2: The accuracy of resnet-32 on long-tailed cifar100.

Methods	CIFAR100		
	$N1/N100 = 10$	$N1/N100 = 50$	$N1/N100 = 100$
M2M	57.55	45.63	42.16
BBN	59.47	45.90	42.47
Methods (ours)	60.14	47.06	42.64

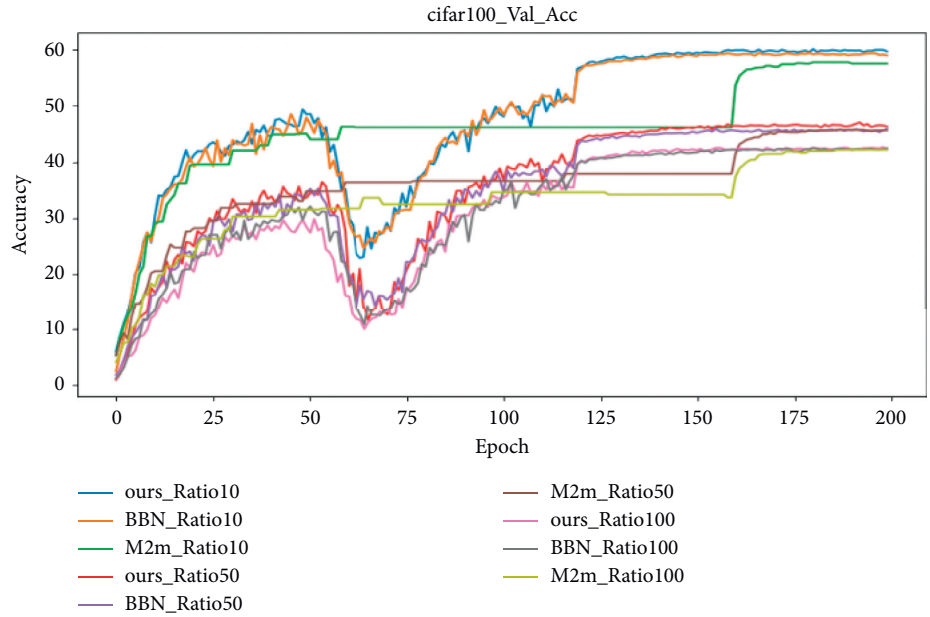


Figure 5: Accuracy variation curve of cifar100 long-tail datasets.

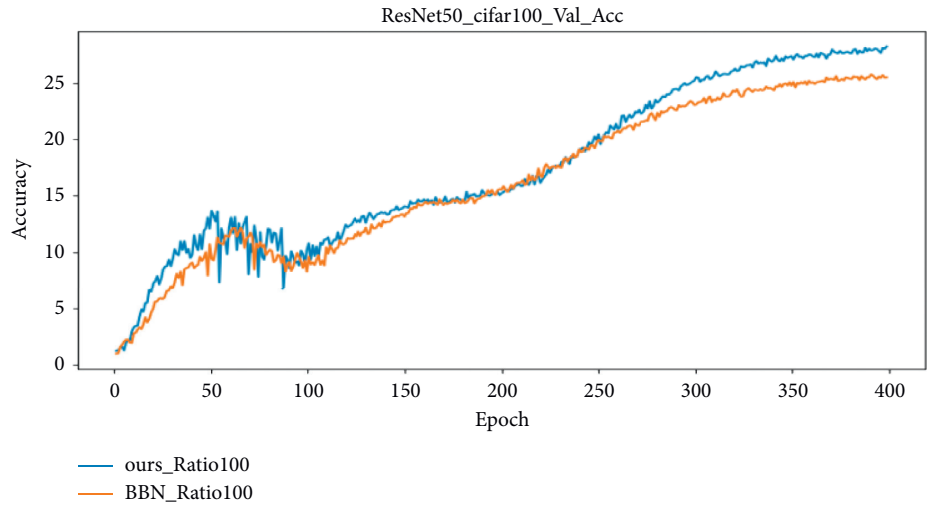


Figure 6: Accuracy curve of resnet50 network structure on a long-tailed datasets with cifar100 imbalance ratio of 100.

5. Summary and Outlook

In this paper, two classical solutions are introduced to address the characteristics of unbalanced data distribution of long-tailed data sets: resampling methods and weighting methods. Two types of algorithms for processing long-tailed data, BBN and M2m. we studied and analyzed these two methods and proposed bilateral-branch generative network model based on them. This model improves the accuracy of classifying long-tailed data by data-driven re-weighting

methods successfully and data enhancement methods based on generative networks. Also in this paper, a time-supervised strategy is introduced in this model to coordinate and control GN and BBN modules to reduce the number of algorithm iterations while maintaining a high accuracy rate. Compared with BBN and M2m algorithm, this algorithm can obtain higher accuracy rate stably.

The following shortcomings still exist in this paper: the time-supervised strategy proposed in this paper, although it reduces a large number of iterations and improves the operation efficiency, making the model run significantly more efficiently than the M2m algorithm, there are still differences compared to the BBN algorithm.

Data Availability

We used CIFAR10 and CIFAR100 public datasets to support the findings of the research. All data used in the study can be download at http://www.cs.toronto.edu/~kriz/cifar.html.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of the study.

Acknowledgments

This paper was partially supported by National Key Research and Development Program of China Grant (2018AAA0100400), the Natural Science Foundation of Hunan Province (2020JJ6089) and the Key Project of the Department of Education in Hunan Province (19A133).

References

[1] O. Russakovsky, J. Deng, H. Su et al., "ImageNet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.

[2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, June 2016.

[3] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba, "Places: a 10 million image database for scene recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 6, pp. 1452–1464, 2018.

[4] G. Van Horn and P. Perona, "The devil is in the tails: fine-grained classification in the wild," 2017, https://arxiv.org/abs/1709.01450.

[5] M. G. Kendall, A. Stuart, J. Keith Ord, S. F. Arnold, O. 'H. Anthony, and J. Forster, *Kendall's Advanced Theory of Statistics*, Vol. 1, Wiley, New York, NY, USA, 1987.

[6] L. Shen, Z. Lin, and Q. Huang, "Relay backpropagation for effective learning of deep convolutional neural networks," in *Proceedings of the Computer Vision - ECCV 2016*, Amsterdam, The Netherlands, October 2016.

[7] E. A. Haibo He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.

[8] J. Byrd and Z. Lipton, "What is the effect of importance weighting in deep learning?" in *Proceedings of the 36th International Conference on Machine Learning, PMLR*, vol. 97, pp. 872–881, Long Beach, CA, USA, 2019.

[9] C. Drummond and R. C. Holte, "C4.5, class imbalance, and cost sensitivity: why under-sampling beats oversampling," *Workshop on Learning From Imbalanced Datasets II*, vol. 11, pp. 1–8, 2003.

[10] A. More, "Survey of resampling techniques for improving classification performance in unbalanced datasets," 2016, https://arxiv.org/abs/1608.06048.

[11] N. V. Chawla, K. W. Bowyer, L. O. Hall, W. P. Philip, and S. M. O. T. E. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.

[12] M. Buda, A. Maki, and M. A. Mazurowski, "A systematic study of the class imbalance problem in convolutional neural networks," *Neural Networks*, vol. 106, pp. 249–259, 2018.

[13] N. Japkowicz and S. Stephen, "The class imbalance problem: a systematic study," *Intelligent Data Analysis*, vol. 6, no. 5, pp. 429–449, 2002.

[14] C. Huang, Y. Li, C. C. Loy, and X. Tang, "Learning deep representation for imbalanced classification," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5375–5384, CVPR, Las Vegas, NV, USA, June 2016.

[15] Y.-X. Wang, D. Ramanan, and M. Hebert, "Learning to model the tail," in *Proceedings of the 2016 NeurIPS*, pp. 7029–7039, Barcelona, Spain, December 2017.

[16] Y. Cui, M. Jia, T.-Y. Lin, Y. Song, and S. Belongie, "Class-balanced loss based on effective number of samples," in *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9268–9277, CVPR, Long Beach, CA, USA, June 2019.

[17] M. Ren, W. Zeng, B. Yang, and R. Urtasun, "Learning to reweight examples for robust deep learning," 2018, https://arxiv.org/abs/1803.09050.

[18] Y. Cui, Y. Song, C. Sun, A. Howard, and S. Belongie, "Large scale fine-grained categorization and domain-specific transfer learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, June 2018.

[19] S. Khan, M. Hayat, S. W. Zamir, J. Shen, and L. Shao, "Striking the right balance with uncertainty," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, June 2019.

[20] K. Cao, C. Wei, A. Gaidon, N. Arechiga, and T. Ma, "Learning imbalanced datasets with labeldistribution-aware margin loss," in *Proceedings of the Advances in Neural Information Processing Systems (NeurIPS)*, Vancouver, Canada, 2019.

[21] X. Zhang, Z. Fang, Y. Wen, Z. Li, and Y. Qiao, "Range loss for deep face recognition with long tailed training data," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Venice, Italy, October 2017.

[22] D. Qi, S. Gong, and X. Zhu, "Imbalanced deep learning by minority class incremental rectification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 6, pp. 1367–1381, 2018.

[23] B. Zhou, Q. Cui, X. S. Wei, and Z. M. Chen, "BBN: bilateral-branch network with cumulative learning for long-tailed visual recognition," in *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, December 2019.

[24] S. S. Mullick, S. Datta, and S. Das, "Generative adversarial minority oversampling," in *Proceedings of the IEEE*

International Conference on Computer Vision (ICCV), Seoul, South Korea, October 2019.

[25] Z. Liu, Z. Miao, X. Zhan, J. Wang, B. Gong, and X. Yu Stella, "Large-scale long-tailed recognition in an open world," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, June 2019.

[26] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycleconsistent adversarial networks," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, vol. 2, pp. 2223–2232, Venice, Italy, 2017.

[27] Y. Choi, M. Choi, M. Kim, J.-W. Ha, S. Kim, and J. Choo, "Stargan: unified generative adversarial networks for multi-domain image-to-image translation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, pp. 8789–8797, Salt Lake City, UT, USA, June 2018.

[28] S. Mo, M. Cho, and J. Shin, "InstaGAN: instance-aware image-to-image translation," in *Proceedings of the International Conference on Learning Representations (ICLR)*, New Orleans, LO, USA, 2019.

[29] J. Kim, J. Jeong, and J. Shin, "M2m: imbalanced classification via major-to-minor translation," in *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[30] P. Branco, L. Torgo, and R. P. Ribeiro, "A survey of predictive modeling on imbalanced domains," *ACM Computing Surveys*, vol. 49, no. 2, p. 31, 2016.

[31] S. Mo, M. Cho, and J. Shin, "InstaGAN: instance-aware image-to-image translation," in *Proceedings of the International Conference on Learning Representations (ICLR)*, New Orleans, LA, USA, May 2019.

[32] A. Krizhevsky, "Learning multiple layers of features from tiny images," Technical report, Department of Computer Science, University of Toronto, Toronto, Canada, 2009.

Hindawi
Complexity
Volume 2021, Article ID 8342610, 11 pages
https://doi.org/10.1155/2021/8342610

WILEY | Hindawi

Research Article

Distributed Optimization for Mobile Robots under Mobile Edge Computing Environment

Hui Luo and Quan Yin

School of Artificial Intelligent and Automation, Huazhong University of Science and Technology, Wuhan, China

Correspondence should be addressed to Hui Luo; keyluo@hust.edu.cn

Received 3 June 2021; Accepted 3 September 2021; Published 21 October 2021

Academic Editor: Xiao Ling Wang

Driven by the development of the Internet industry, mobile robots (MRs) technology has become increasingly mature and widely used in all walks of life. Since MRs are densely distributed in the network system, how to establish a reliable communication architecture to achieve good cooperation and resource sharing between MRs has become a research hotspot. In this respect, mobile edge computing (MEC) technology and millimeter wave (mmW) technology can provide powerful support. This paper proposes a mmW communication network architecture for distributed MRs in MEC environment. The mmW base station provides reliable communication services for MRs under the coverage of information cloud (IC). We design a joint resource and power allocation strategy aimed at minimizing network energy consumption. First, we use the Lyapunov optimization technique to transform the original infinite horizon Markov decision process (MDP) problem. Then, a semidistributed algorithm is introduced to solve the distributed optimization problem in the mmW network. By improving the autonomous decision-making ability of the mmW base station, the signaling overheads caused by information interaction are reduced, and information leakage is effectively avoided. Finally, the global optimal solution is obtained. Simulation results demonstrate the superiority of the proposed strategy.

1. Introduction

With the rapid development of the Internet industry, a variety of intelligent terminals emerge in an endless stream. Among them, intelligent MRs can set environmental perception, dynamic decision-making and planning, behavior control and execution, and other functions in one, widely used in the industry, agriculture, medical, service, and other industries, and has aroused universal attention all over the world. In order to realize good cooperation and resource sharing among multiple MRs, so that they can assist or replace human work to the greatest extent, we need to establish a communication framework among MRs. In the current research, the multirobot wireless network system can realize flexible cooperation through a central server [1]. Based on TCP and UDP in the local area network (LAN), the distributed communication method proposed in [2] allows robots to subscribe to target topics independently to achieve multirobot cross-platform communication. Each robot in [3] can establish communication paths to maintain a consistent communication with the cloud to effectively realize cooperative communication and link prediction. The authors in [4] proposed a scheme of mobile robot object recognition in Internet of Things (IOT) and used edge nodes to coordinate the data from robot vision.

In the traditional distributed communication network architecture, a centralized management scheme centered on the core network is usually adopted. The core network collects global network information and makes global decisions. With the gradual expansion and complexity of network structure, centralized management scheme will inevitably lead to huge signaling overheads, time delay, and privacy leakage. Among the existing distributed algorithms, the authors in [5, 6] proposed a distributed downlink resource allocation and multidomain interference management algorithm to alleviate the co-channel interference in

small cellular networks and maximize the total utility of the whole system. The authors in [7] proposed a distributed algorithm based on the machine learning framework of echo state networks, aiming at the problems of user association and spectrum allocation in heterogeneous small cellular networks, which allowed small base stations to independently complete the optimal allocation of resources under the condition of limited system state information and greatly reduced the wireless network information exchange. The authors in [8] proposed an adaptive semidistributed algorithm that can jointly allocate power and spectrum resources on licensed and unlicensed bands, which can improve the independent adaptability of the base station and maximize the global spectrum efficiency. Based on the research above, we need to develop a distributed network resource management scheme which is more suitable for MRs network and has stronger autonomy and flexibility, higher security, and cost performance.

In this regard, Mobile Edge Computing (MEC) technology can sink elastic resources such as computing, network, and storage in cloud data center to the edge of distributed network, effectively relieving the pressure on core network, improving computing efficiency, and reducing service latency [9–11]. MEC deeply integrates mobile access network with Internet service, which is very suitable for the MRs communication system with distributed structure. Meanwhile, to meet the development needs of high speed, low delay, and large capacity of mobile communication network in the future, human beings are exploring more abundant spectrum resources in higher frequency band. Among them, millimeter wave (mmW) not only has a wide spectrum, which can greatly improve the channel capacity, but also has a good direction, which can effectively reduce channel interference and protect user privacy [12]. Applying mmW technology to MRs communication networks can further improve network performance and user experience, which is an important trend of distributed MRs network development. In the existing work, the authors in [13] combined the computational shunt technology in MEC with mmW communication to conduct joint optimization of computing and communication resources, forming an energy-saving and efficient computing system. Aiming at distributed small cellular networks, a green load management and resource allocation scheme based on MEC is proposed in [14, 15], which is constrained by network cost and communication quality to achieve green computing shunt and reduce the total energy consumption of users to the greatest extent.

Based on the current research, we propose a distributed MRs communication network architecture in MEC environment. The mmW access points (APs) are randomly distributed under the coverage of the information cloud (IC) to provide communication services for the densely deployed MRs, in which the IC is responsible for information exchange, storage, and interaction. To minimize the energy consumption of the MRs network and achieve the optimal network performance, we designed a joint resource and power allocation strategy based on MEC under the condition of guaranteeing the network traffic delay, limiting rate, interference, and power. First, we use the Lyapunov optimization technique to transform the infinite horizon problem in the objective function into a single frame problem that can be processed. Second, we introduce a semidistributed algorithm for distributed optimization issues on complex mmW MECs. Compared with traditional distributed programs, this scheme improves the independent decision-making capacity of mmW APs and only needs to exchange a small amount of Lagrangian parameters through IC without sharing global information, which greatly reduces signaling overheads and protects the users' privacy.

The rest of the paper is organized as follows. In Section 2, we describe the system model of the distributed MRs communication network. Section 3 formulates and transforms the optimization problem. The semidistributed resource and power allocation are proposed in Section 4. Section 5 shows the simulation evaluation, and we draw the conclusion in Section 6.

Notations: $\|\cdot\|$ is the Euclidean norm. Let $\mathrm{diag}(x_1, \ldots, x_K)$ be the diagonal matrix. I denotes the identity matrix. $\mathbf{A} \otimes \mathbf{B}$ means the Kronecker product of matrixes $\mathbf{A}$ and $\mathbf{B}$. $\mathbb{E}[\cdot]$ represents the statistical expectation.

2. System Model

In the system, under the coverage of the IC that introduces MEC technology and is responsible for data calculation, storage, and interaction, there are I mmW APs and each of them provides services to J MRs within its coverage as shown in Figure 1. The set of mmW APs and MRs is denoted as $\mathcal{I} = \{W_1, W_2, \ldots, W_I\}$ and $\mathcal{J} = \{R_1, R_2, \ldots, R_J\}$. Each mmW AP equips with mmW transceivers to overcome the significantly high path loss though highly directional transmission, and each MR introduces MEC technology and equips with a mmW beam steerable array to support transmission over IC and mmW network.

We have each MR implementing the analog beamforming, allowing one transmission link each time, and each mmW AP implementing the hybrid beamforming, allowing multiple transmission links each time. The mmW AP W_i equipped with C_i radio frequency (RF) chains can serve most C_i MRs. Moreover, we assume that the proposed network transmits in a slotted system, which is represented as $t = \{0, 1, 2, \ldots\}$. Denote $x_{i,j}(t) = 1$ if the j-th MR is in the coverage of the mmW AP W_i at time slot t, and $x_{i,j}(t) = 0$ otherwise. We assume each MR is attached to not more than one mmW AP, i.e., $\sum_{i=1}^{I} x_{i,j}(t) \le 1$, $\forall j \in \mathcal{J}$. We further assume that the total number of MRs attached to the mmW AP does not exceed the maximum chain number, i.e., $\sum_{j=1}^{J} x_{i,j}(t) \le C_i$, $\forall i \in \mathcal{I}$.

2.1. mmW Transmission Model. The mmW network is susceptible to interference and has severe channel attenuation. By applying beamforming technology, directional transmission can be formed between mmW APs and MRs to increase data rate. According to [16], we build the model of mmW network transmission as shown in Figure 2. Let $\theta_i^a(t)$

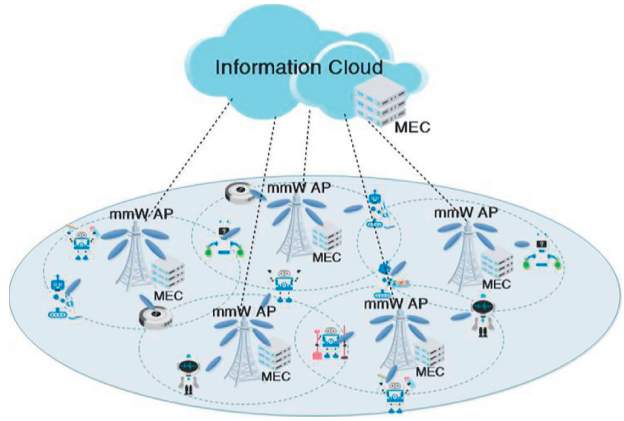


Figure 1: The network framework for mobile robots in MEC environment.

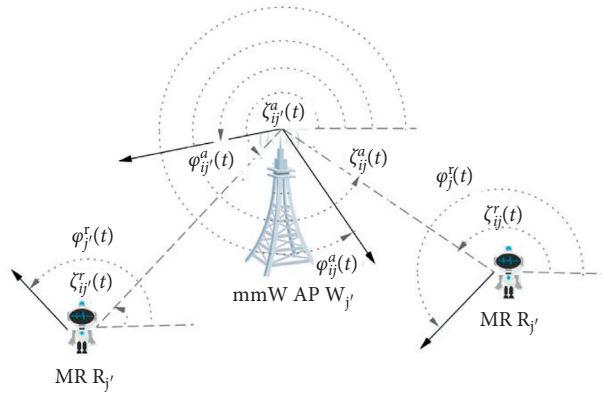


Figure 2: Illustration of the angles between the mmW AP and MRs in the transmission.

and $\theta^r_j(t)$ be the operating beamwidths of mmW AP W_i and MR R_j, respectively. Let $\zeta^a_{ij}(t)$ be the angle between the positive x-axis and the direction in which the mmW AP W_i sees MR R_j, and let $\zeta^r_{ij}(t)$ be similarly defined by changing the roles of the mmW AP W_i and MR R_j. The values of $\zeta^a_{ij}(t)$ and $\zeta^r_{ij}(t)$ satisfy $|\zeta^a_{ij}(t) - \zeta^r_{ij}(t)| = \pi$. We further denote $\varphi^a_{ij}(t)$ and $\varphi^r_j(t)$ as the boresight angles of the mmW AP W_i and MR R_j relative to the positive x-axis. Let $g^a_{ij,ji}(t)$ be the transmission gain between the beam from the mmW AP W_i to MR R_j. and $g^r_{ji,ij}(t)$ be the reception gain between the beam from MR R_j to the mmW AP W_i, which can be written as follows.

$$g^a_{ij,ji}(t) = \begin{cases} \varepsilon, & \text{if } \dfrac{\theta^a_i(t)}{2} < \left|\varphi^a_{ij}(t) - \zeta^a_{ij}(t)\right| < 2\pi - \dfrac{\theta^a_i(t)}{2}, \\ \dfrac{2\pi - (2\pi - \theta^a_i(t))\varepsilon}{\theta^a_i(t)}, & \text{otherwise}, \end{cases}$$

$$g^r_{ji,ij}(t) = \begin{cases} \varepsilon, & \text{if } \dfrac{\theta^r_i(t)}{2} < \left|\varphi^r_j(t) - \zeta^r_{ij}(t)\right| < 2\pi - \dfrac{\theta^r_i(t)}{2}, \\ \dfrac{2\pi - (2\pi - \theta^r_i(t))\varepsilon}{\theta^r_i(t)}, & \text{otherwise}. \end{cases} \quad (1)$$

2.2. Interference Management. According to [16], there are two types of interference in the proposed mmW network that should be managed. The intracell interference exists among devices within a cell and can be mitigated by using proper scheduling and beamforming design. The intercell interference exists among different cells and can be

significantly reduced by resource blocks scheduling and omnidirectional communication. Therefore, to ensure that the intercell interference in the mmW network is small enough to be ignored, the following constraints can be introduced:

$$\sum_{\substack{i' \in \mathcal{F} \\ i' \neq i}} \sum_{j' \in \mathcal{J}} p^a_{i'j'}(t) g^a_{i'j',ji}(t) g^c_{i'j}(t) g^r_{j'i',ij}(t) \leq I^T, \tag{2}$$

where I^T is the intercell interference threshold, $p^a_{ij}(t)$ denotes the transmit power from the mmW AP W_i to MR R_j, and $g^c_{ij}(t)$ refers the channel gain between the mmW AP W_i to MR R_j.

2.3. Data Transmission. Let $Q_j(t)$ be the data queue length of MR R_j at time slot t. The data entering the queue of MR R_j at time slot t are denoted as $A_j(t)$, whose maximum value is $A_{j,\max}(t)$. The traffic buffer queue $Q_j(t)$ of MR R_j evolves as

$$Q_j(t+1) = \left[Q_j(t) - R^a_j(t)\tau\right]^+ + A_j(t)\tau, \quad \forall j \in \mathcal{J}, t, \tag{3}$$

where $[x]^+ = \max\{x, 0\}$ is on account that the actual served packets cannot be larger than the current queue size. τ is the duration of each time slot. We further let $Q_j(0) = 0$, $\forall j \in \mathcal{J}$ for simplicity.

The achievable data rate of MR R_j served by the mmW AP W_i at time slot t can be written as

$$R^a_{ij}(t) = x_{i,j}(t) B^a \log_2\left(1 + \frac{p^a_{ij}(t) g^a_{ij,ji}(t) g^c_{ij}(t) g^r_{ji,ij}(t)}{B^a N_0}\right), \quad \forall j \in \mathcal{J}, i \in \mathcal{I}, t, \tag{4}$$

where B^a is the mmW bandwidth and N_0 is the noise power spectral density. Based on (4), the achievable data rate of MR R_j served by the mmW network at time slot t is given by

$$R^a_j(t) = \sum_{i=1}^{I} R^a_{ij}(t), \quad \forall j \in \mathcal{J}, t. \tag{5}$$

3. Problem Formulation and Transformation

3.1. Problem Formulation. Our goal in this work is to design a resource and power allocation strategy to minimize the energy consumption of devices and guarantee power and rate limitation and traffic latency of MRs as well as ensuring performance of the mmW network. The energy aware optimization problem can be formulated as

$$\text{P1:} \min_{\{\mathbf{p}^a\}} \limsup_{T \longrightarrow +\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\left[\sum_{j \in \mathcal{J}} p^a_j(t)\tau\right], \tag{6}$$

subject to (2),

$$\limsup_{T \longrightarrow +\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\left[Q_j(t)\right] < \overline{Q}_j, \quad \forall j \in \mathcal{J}, t, \tag{7a}$$

$$R^a_j(t) \geq R^T_j, \quad \forall j \in \mathcal{J}, t, \tag{7b}$$

$$p^a_j(t) \leq P^T_j, \quad \forall j \in \mathcal{J}, t, \tag{7c}$$

$$p^a_{ij}(t) \geq 0, \quad \forall j \in \mathcal{J}, i \in \mathcal{I}, t, \tag{7d}$$

where $p^a_j(t) = \sum_{i \in \mathcal{I}} p^a_{ij}(t)$, $\overline{Q}_j$ is the predefined average queue backlog tolerance of MR R_j, and $R^T_j(t)$ and P^T_j are the minimum data rate requirement and the maximum power consumption of MR R_j at time slot t. Constraint (7a) ensures the time-averaged latency of MR, which is proportional to the average queue backlog length according to Little's theorem. Constraint (7d) is the limitation of parameters.

3.2. Lyapunov Drift Penalty. Since the optimization function P1 is a long-term average network energy consumption problem, it is an infinite horizon Markov decision process (MDP) problem [17]. To tackle the infinite horizon problem P1 with lower computational complexity and less memory consumption, we introduce the Lyapunov drift penalty technique to transform P1 into a tractable single frame problem [18]. We define a virtual queue Z_n for constraint (7a) as

$$Z_j(t+1) = \left[Z_j(t) + Q_j(t+1) - \overline{Q}_j\right]^+, \quad \forall j \in \mathcal{J}, t. \tag{8}$$

For each slot t, the Lyapunov function can be formulated as a sum of virtual queue length squares, given by

$$L(\mathbf{Z}(t)) = \frac{1}{2}\left(\sum_{j=1}^{J} Z^2_j(t)\right), \quad \forall t, \tag{9}$$

where $\mathbf{Z}(t) = [Z_1(t), Z_2(t), \ldots, Z_J(t)]$.

Moreover, the Lyapunov drift function is expressed as

$$\triangle \mathbf{Z}(t) = \mathbb{E}[L(\mathbf{Z}(t+1)) - L(\mathbf{Z}(t)) | \mathbf{Z}(t)], \quad \forall t. \tag{10}$$

Then, the Lyapunov drift penalty function is defined to minimize the drift-plus-penalty expression boundary, which can be formulated as

$$\begin{aligned} \Delta_\nu \mathbf{Z}(t) = {} & \mathbb{E}\left[\frac{1}{2}\sum_{j \in \mathcal{J}}\left(Z^2_j(t+1) - Z^2_j(t)\right) | \mathbf{Z}(t)\right] \\ & + V\mathbb{E}\left[\sum_{j \in \mathcal{J}} p^a_j(t)\tau | \mathbf{Z}(t)\right], \quad \forall t, \end{aligned} \tag{11}$$

where V is the nonnegative control parameter, which represents the weight of the objective function.

Lemma 1. *The upper bound for the Lyapunov drift penalty (11) can be written as*

$$\triangle_{\gamma}\mathbf{Z}(t) \leq \sum_{j\in\mathcal{J}} \frac{Z_{j1}+Z_{j2}}{2} + V\mathbb{E}\left[\sum_{j\in\mathcal{J}} p_j^a(t)\tau|\mathbf{Z}(t)\right] - \mathbb{E}\left[\sum_{j\in\mathcal{J}} \left(Q_j(t)+Z_j(t)\right)R_j^a(t)\tau|\mathbf{Z}(t)\right], \quad \forall t. \tag{12}$$

Here, $Z_{j1} = (Q_{j,\max}+A_{j,\max})^2 + (Q_{j,\max})^2 + (\overline{Q}_j)^2, Z_{j2} = 2Z_{j,\max}(Q_{j,\max}+A_{j,\max})$, *and* $Q_{j,\max}$ *and* $Z_{j,\max}$ *are upper bounds for* $Q_j(t)$ *and* $Z_j(t)$.

Proof. See Appendix A.

From Lemma 1, P1 can be decoupled into the following auxiliary problem at each single SP t:

$$\text{P2: } \min_{\{\mathbf{p}^a\}} V\sum_{j\in\mathcal{J}} p_j^a(t) - \sum_{j\in\mathcal{J}} \left(Q_j(t)+Z_j(t)\right)R_j^a(t), \tag{13}$$

subject to (2) and (7b)–(7d). For more precise transformation, we introduce Theorem 1 to prove the relationship between the solutions of P1 and P2. □

Theorem 1. *S^{β} is denoted as the optimal average energy consumption under* P2*, which is limited by an upper bound.*

$$S^{\beta} \leq \frac{\sum_{j\in\mathcal{J}}\left(Z_{j1}+Z_{j2}\right)}{2V} + S^{\alpha}, \tag{14}$$

where S^{α} is the optimal average energy consumption under P1*. The optimal value of S^{β} can be arbitrarily close to the optimal value S^{α} by selecting asymptotically large V.*

Proof. See Appendix B. □

3.3. Distributed Optimization Problem. In the following, we omit the time slot index t for ease of exposition. Based on P2, we introduce a block of slack variables $c = \{c_1, c_2, \ldots, c_J\}$ and $d = \{d_1, d_2, \ldots, d_J\}$ and get the decentralized mmW network problem as follows:

$$\min_{\{\mathbf{p}^a\}} V\sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{J}} p_{ij}^a - \sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{J}} (Q_i+Z_i)R_{ij}^a, \tag{15}$$

subject to (2) and (7d),

$$R_{ij}^a \geq x_{i,j}\left(R_j^T - c_j\right), \quad \forall j\in\mathcal{J}, i\in\mathcal{J}, \tag{16a}$$

$$p_{ij}^a \leq x_{i,j}\left(P_j^T - d_j\right), \quad \forall j\in\mathcal{J}, i\in\mathcal{J}. \tag{16b}$$

To solve the decentralized network problem, the mmW APs are modeled as agents which can make decisions independently. Based on the initialization-free distributed method in [17, 19], problem (15) can be rewritten as

$$\min_{\mathbf{X}_i, i\in\mathcal{J}} \sum_{i\in\mathcal{J}} f_i(\mathbf{X}_i), \tag{17}$$

subject to

$$\sum_{i\in\mathcal{J}} \mathbf{G}_i\mathbf{X}_i = \sum_{i\in\mathcal{J}} \mathbf{I}, \tag{18a}$$

$$\mathbf{X}_i \in \Omega_i, \quad \forall i\in\mathcal{J}. \tag{18b}$$

where $f_i(\mathbf{X}_i)$ is a convex function which is equivalent to (15) and $\mathbf{X}_i = [p_{1i}^a, \ldots, p_{Ji}^a]^T$. Equation (18a) is the global constraint related to (2), where $\mathbf{I} = [I^T, \ldots, I^T]^T$, $\mathbf{G}_i = [\mathbf{g}_{i1}, \ldots, \mathbf{g}_{iJ}]$, $\mathbf{g}_{ij} = [g_{ij,1}, \ldots, 0, \ldots, g_{ij,J}]^T$, and $g_{ij,k} = \sum_{n\in\mathcal{J}} x_{n,j} g_{ij,kn}^a g_{ji,nk}^u g_{jn}^c$. Equation (18b) is the local feasibility constraint of agent i associated with (7c), (16a), and (16b).

The local objective function $f_i(\mathbf{X}_i)$, the interference matrix $\mathbf{G}_i$ and closed convex set Ω_i are the private information not shared with others of agent i. Equation (17) is a distributed optimization problem. In our framework, under the coverage of information cloud, the mmW APs can carry out data calculation and information interaction through the IC. Hence, we need to design a semidistributed joint resource and power allocation strategy for the agents, in which each agent interacts with its neighbor agents through the IC and collaborates to find its own optimal scheme.

4. Semidistributed Resource and Power Allocation

4.1. The Proposed Semidistributed Algorithm. First, we defined the Lagrangian duality of (17) with Lagrange multiplier λ as

$$\max_{\lambda} q(\boldsymbol{\lambda}) = \sum_{i\in\mathcal{J}} q_i(\boldsymbol{\lambda}) = \sum_{i\in\mathcal{J}} \inf_{\mathbf{X}_i\in\Omega_i} \left\{f_i(\mathbf{X}_i) + \boldsymbol{\lambda}^T\left(\mathbf{G}_i - \mathbf{I}\right)\right\}, \tag{19}$$

where λ can be updated via IC. Due to the global multiplier λ, we need to solve the subproblems to calculate the gradients in the IC. To avoid this shortage, we can introduce the Laplacian matrix $\mathbf{L}$ and the multiplier matrix $\Lambda = \mathrm{col}(\lambda_1, \ldots, \lambda_I)$ and reformulate the following constrained optimization problem [15].

$$\max_{\Lambda} Q(\boldsymbol{\Lambda}) = \sum_{i\in\mathcal{J}} q_i(\boldsymbol{\lambda}_i), \tag{20}$$

subject to

$$(\mathbf{L}\otimes\mathbf{I}_i)\boldsymbol{\Lambda} = 0, \tag{21a}$$

where $\otimes$ denotes the Kronecker product operation and $\mathbf{I}_i$ is an identity matrix.

Before further defining the Laplacian matrix $\mathbf{L}$, we need to introduce the graph theory which can describe the information exchange among the agents through a graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$. According to the graph theory [20], the edge set is defined as $\mathcal{E} \subset \mathcal{N}\times\mathcal{N}$. If agent k can receive information from agent n, then agent n belongs to agent k's neighbor set, denoted as $\mathcal{N}_k = \{n|\ (n,k)\in\mathcal{E}\}$. In this work, we need to assume that the information exchange graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ is

(1) $\dot{\mathbf{X}}_i(\ell+1) = P_{\Omega_i}(\mathbf{X}_i(\ell) - \nabla f_i(\mathbf{X}_i(\ell)) + \lambda_i(\ell)\mathbf{G}_i) - \mathbf{X}_i(\ell)$;
(2) $\dot{\lambda}_i\ (\ell+1) = -\sum_{k\in\mathcal{N}_i}(\lambda_i(\ell) - \lambda_k(\ell)) - \sum_{k\in\mathcal{N}_k}(\mathbf{z}_i(\ell) - \mathbf{z}_k(\ell)) + ((\mathbf{I}/K) - \mathbf{X}_i(\ell)\mathbf{G}_i)$;
(3) $\dot{\mathbf{z}}_i(\ell+1) = \sum_{k\in\mathcal{N}_i}(\dot{\lambda}_i(\ell) - \lambda_k(\ell))$;

Algorithm 1: Distributed algorithm for each agent i, $i \in \mathcal{I}$.

(1) Initialize $\ell = 1, \mathbf{X}(1), \Lambda(1), \mathbf{Z}(1)$
(2) Set the neighbor sets for all the mmW APs as $\mathcal{N}_i, \forall i \in \mathcal{I}$
(3) Let $\sigma \longrightarrow 0$
(4) while $(\|\dot{X}(\ell)\|_2^2 + \|\dot{\Lambda}(\ell)\|_2^2 + \|\dot{Z}(\ell)\|_2^2) \geq \sigma$
(5) mmW APs obtain the information of $\Lambda_i(\ell)$ and $\mathbf{Z}_i(\ell)$, $\forall i \in \mathcal{I}$ from the information cloud;
(6) Each mmW AP i, $\forall i \in \mathcal{I}$ updates its power allocation and multiplier parameters independently based on (23);
(7) mmW APs i feed back the information of $\Lambda_i(\ell)$ and $\mathbf{Z}_i(\ell)$, $\forall i \in \mathcal{I}$ to the information cloud;
(8) set $\ell = \ell + 1$.
(9) Endwhile

Algorithm 2: Semidistributed power allocation scheme.

Table 1: Parameter setting.

Parameters	Values
Minimum data rate R^T	0.5 to 30 mbps
Maximum power consumption P^T	27 dBm
Queue backlog tolerance Q^j	1 to 10 ms
Noise power spectral density N^0	−134 dBm/MHz
The mmW bandwith B^a	2 GHz
Operating beamwidths between mmW AP and MR θ	10°
The intercell interference threshold I^T	10^{-10} W
The duration of each time slot τ	1 ms

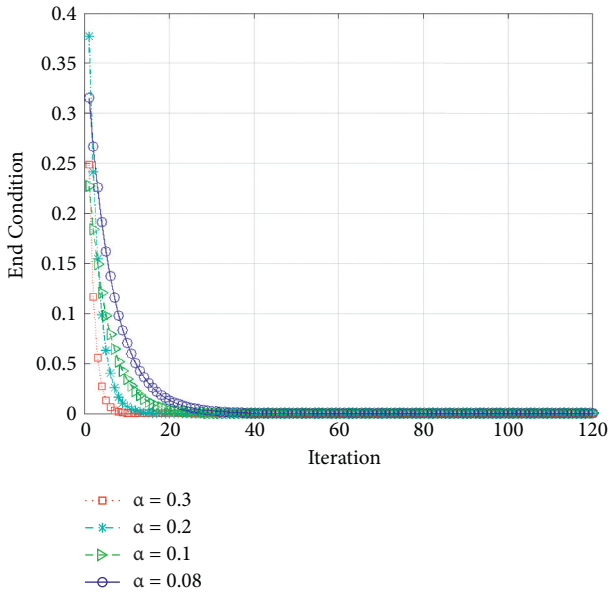


Figure 3: Convergence performance of Algorithm 2 under asynchronous integration step.

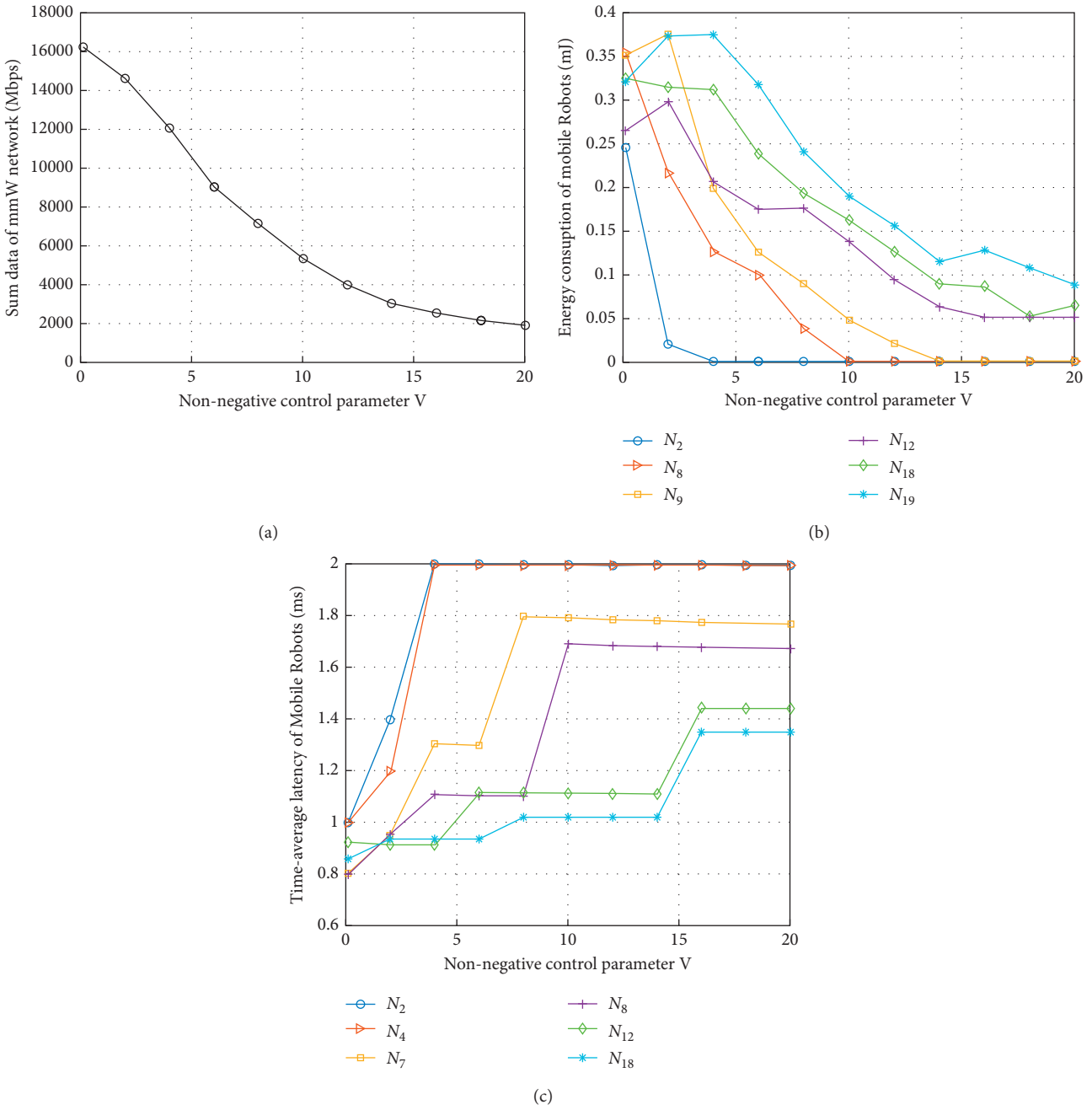


FIGURE 4: The effects of nonnegative control parameter on mmW network performance: (a) data rate; (b) energy consumption; (c) average latency.

undirected and connected to guarantee that any agent's information can reach any other agents.

According to the above, the Laplacian matrix can be defined as $\mathbf{L} = \mathbf{De}\,\mathbf{g} - \mathbf{A}$, where$\mathbf{A} = [a_{kn}]$ is the adjacency matrix of graph $\mathcal{G}$. We denote $a_{kn} = 1$ if n belongs to agent k's neighbor set, and $a_{kn} = 0$ otherwise. $\mathbf{De}\,\mathbf{g} = \mathrm{diag}\,(\sum_{n=1}^{I} a_{1n}, \ldots, \sum_{n=1}^{I} a_{In})$ is the degree matrix. Constraint (21a) is defined to guarantee $\lambda_1 = \lambda_2 = \cdots = \lambda_I$, which insure the consensus among agents i is reached.

Based on the above conversion, the original distributed optimization problem has been decoupled and each agent can decide its own resource and power allocation strategy.

By introducing the Lagrangian multiplier $\mathbf{Z} = \mathrm{col}\,(\mathbf{z}_1, \ldots, \mathbf{z}_I)$, we denote the augmented Lagrangian duality of (20) as

$$\min_{\mathbf{Z}} \max_{\mathbf{\Lambda}} Q(\mathbf{\Lambda}, \mathbf{Z}) = \sum_{i \epsilon, \mathcal{F}} q_i(\boldsymbol{\lambda}_i) - \mathbf{Z}^T(\mathbf{L} \otimes \mathbf{I}_i)\mathbf{\Lambda} - \frac{1}{2}\mathbf{\Lambda}^T(\mathbf{L} \otimes \mathbf{I}_i)\mathbf{\Lambda}. \tag{22}$$

According to [19], (22) can be solved by applying the gradient flow and the distributed algorithm for agent i based on projection dynamics given in Algorithm 1.

In Algorithm 1, $\dot{f}(\cdot)$ represents gradient calculation of f, ℓ is the update index, $P_{\Omega_i}(x) = \text{argmin}_{y\in\Omega_i}\|x - y\|$ is the projection operation of x onto the closed convex set Ω_i, ∇f means the gradient of f, and $\mathcal{N}_i$ is the neighbor set of agent i. Based on the distributed algorithm, both the power allocation and multiplier parameters can be updated by

$$\begin{aligned}\mathbf{p}_i^{\mathbf{a}}(\ell+1) &= \mathbf{p}_i^{\mathbf{a}}(\ell) + \alpha\dot{\mathbf{p}}_i^{\mathbf{a}}\ (\ell+1),\\ \boldsymbol{\lambda}_i(\ell+1) &= \boldsymbol{\lambda}_i(\ell) + \alpha\dot{\boldsymbol{\lambda}}_i\ (\ell+1),\\ \mathbf{z}_i(\ell+1) &= \mathbf{z}_i(\ell) + \alpha\dot{\mathbf{z}}_i\ (\ell+1),\end{aligned} \tag{23}$$

where α is the integration step.

Algorithm 2 summarizes the detailed procedures of the semidistributed scheme. Specifically, mmW APs achieve the information of Lagrangian multipliers from the information cloud, then calculatetheir power allocation strategy independently, and feed back $\lambda_i(\ell)$ and $\mathbf{z}_i(\ell)$, $\forall i \in \mathcal{I}$, to the information cloud.

4.2. The Performance Analysis of the Semidistributed Scheme. The first and most important advantage of Algorithm 2 is that it does not sacrifice the optimality. We can obtain the global optimal solution for the distributed mmW networks to maximize the utilization efficiency of wireless resource. We introduce the following theorem to prove.

Theorem 2. *The proposed semidistributed algorithm can converge to an optimal power allocation solution.*

Proof. See Appendix C.

The second merit is initialization-free, which means that the semidistributed algorithm will converge to the global optimal solution under any initial condition due to the projection operation. □

5. Simulation and Evaluation

In this section, different contrast graphs were obtained by adjusting various parameters. There are multiple line segments with different parameters in each experimental result image. We draw images with lines of different colors and shapes for easy differentiation [5]. Under the coverage of IC, we set 4 mmW APs, each equipped with 6 RF chains. And there are 24 densely deployed MRs served by the mmW network. We denote the set of MRs as $\mathcal{N} = \{N_1, N_2, \ldots, N_{24}\}$, where the minimum data rate of the MRs is set from 0.5 Mbps to 30 Mbps. The remaining parameters are shown in Table 1.

Figure 3 shows the convergence of Algorithm 2 and the effect of integration step on the convergence rate. From the figure, we can observe that the semidistributed power allocation algorithm converges to the optimal solution through iteration, and the larger the integration step setting is, the faster the convergence rate will be. Therefore, we select the appropriate integration step on the premise of ensuring the performance of the algorithm.

Figure 4 shows the variation of data rate, energy consumption, and average latency with the nonnegative control parameters V, which represents the weight of the objective function in the conversion of Lyapunov drift penalty. From Figure 4(C), we can clearly observe that the time-average latency increases gradually with the gradual increase of parameter V. In addition, mobile robots with higher data rates usually result in higher energy consumption and lower time latency.

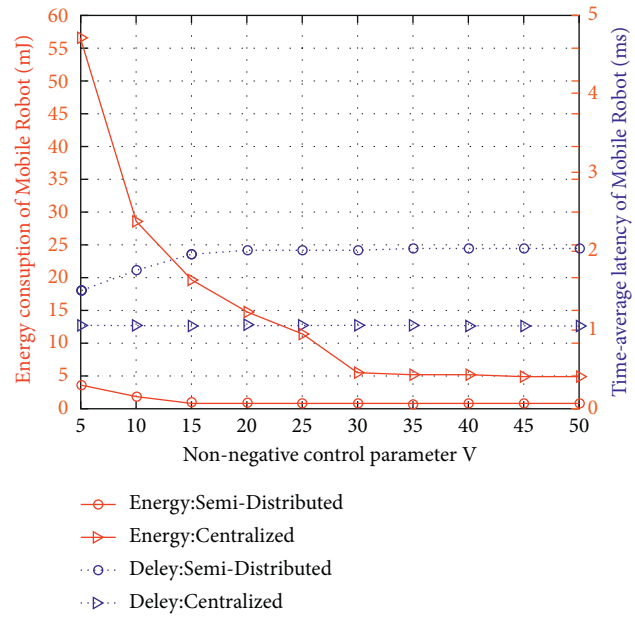


FIGURE 5: The performance between two different solutions.

Figure 5 shows the performance of two different solutions. The semidistributed method refers to the proposed resource and power allocation scheme in this paper, and the centralized method refers to solve optimization problems directly. From the picture, we can observe that the semidistributed method has significantly lower energy consumption but slightly higher time-average latency than the centralized method.

6. Conclusion

This paper proposes a mmW communication network for distributed MRs in a MEC environment. While minimizing network energy consumption, a joint resource and power distribution scheme is designed. First, Lyapunov optimization is processed by the target function as a single frame problem and then introduced into the semidistributed algorithm which further solves distributed optimization issues on mmW MECs. Experiments prove that the program can effectively reduce network energy consumption, and the semidistributed algorithm can greatly reduce signaling overheads, prevent privacy leakage, and further enhance the performance of distributed MRs communication networks.

Appendix

A

Proof. of Lemma 1.

Denote $\Delta Z_j(t) = \mathbb{E}[(1/2)Z_j(t+1)^2 - (1/2)Z_j(t)^2 | Z_j(t)]$. Without expectation, $\Delta Z_j(t)$ can be written as

$$\Delta Z_j(t) = \frac{1}{2}Z_j(t+1)^2 - \frac{1}{2}Z_j(t)^2, \quad \forall j \in \mathcal{J}. \tag{A.1}$$

Further, we can derive that

$$\begin{aligned}
\Delta Z_j(t) &\overset{(a)}{\leq} \frac{1}{2}\Big[Q_j^2(t+1) + 2Q_j(t+1)Z_j(t) + \big(\overline{Q}_j\big)^2 \\
&\quad -2\overline{Q}_j\big(Q_j(t+1) + Z_j(t)\big)\Big], \\
&\overset{(b)}{\leq} \frac{1}{2}\Big[Q_j^2(t) + \big(\overline{Q}_j\big)^2 + \big(R_j^a(t)\tau\big)^2 - 2\tau Q_j(t)R_j^a(t) \\
&\quad + \big(A_j(t)\tau\big)^2 + 2\tau Q_j(t)A_j(t) + +2Q_j(t+1)Zj(t)\Big], \\
&\overset{(c)}{\leq} \frac{1}{2}\Big[2Q_j^2(t) + \big(\overline{Q}_j\big)^2 - 2\tau Q_j(t)R_j^a(t) + \big(A_j(t)\tau\big)^2 \\
&\quad + 2\tau Q_j(t)A_j(t) + 2Z_j(t)\big(Q_j(t) - R_j^a(t)\tau + A_j(t)\tau\big)\Big] \\
&\leq \frac{1}{2}\big(Z_{j1} + Z_{j2}\big) - R_j^a(t)\tau\big(Q_j(t) + Z_j(t)\big), \quad \forall j \in \mathcal{J}.
\end{aligned} \tag{A.2}$$

Here, (a) is with the fact of $([x]^+)^2 \leq x^2$, and (b) is proved by the following equation:

$$\begin{aligned}
Q_j^2(t+1) &= \Big(\big[Q_j(t) - R_j^a(t)\tau\big]^+ + A_j(t)\tau\Big)^2 \\
&\leq \Big(\big[Q_j(t) - R_j^a(t)\tau\big]^+\Big)^2 + \big(A_j(t)\tau\big)^2 \\
&\quad + 2A_j(t)\tau Q_j(t).
\end{aligned} \tag{A.3}$$

We assume that the queue data that have not been backlogged and need not to be transmitted; inequation (c) has the causality constraint of

$$R_j^a(t)\tau \leq Q_j(t) < +\infty, \quad \forall j \in \mathcal{J}. \tag{A.4}$$

According to [18], constraint (A.4) can be ignored. Substituting (A.2) into (11), we can obtain the upper bound of $\Delta_v \mathbf{Z}(t)$, which ends the proof. □

B

Proof. of Theorem 1.

Substituting (11) into (12), summing up time slot t from 0 to $T-1$, and dividing T, we have

$$\begin{aligned}
&\frac{1}{T}\mathbb{E}\left[\frac{1}{2}\sum_{t=0}^{T-1}\sum_{j\in\mathcal{J}}\big(Z_j^2(t+1) - Z_j^2(t)\big)\right] + V\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[\sum_{i\in\mathcal{J}} p_i^a(t)\tau\right] \\
&= \frac{1}{T}\mathbb{E}\left[\frac{1}{2}\sum_{j\in\mathcal{J}}\big(Z_j^2(T) - Z_j^2(0)\big)\right] + V\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[\sum_{i\in\mathcal{J}} p_i^a(t)\tau\right] \\
&\overset{(a)}{=} \frac{1}{T}\mathbb{E}\left[\sum_{j\in\mathcal{J}}\frac{Z_j(T)^2}{2}\right] + V\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[\sum_{i\in\mathcal{J}} p_i^a(t)\tau\right] \\
&\overset{(b)}{=} \frac{1}{T}\mathbb{E}[L(Z(T))] + V\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[\sum_{i\in\mathcal{J}} p_i^a(t)\tau\right] \\
&\leq \sum_{i\in\mathcal{J}}\frac{Z_{j1} + Z_{j2}}{2} + VE(\varpi) \\
&\quad - \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[\sum_{j\in\mathcal{J}}\big(Q_j(t) + Z_j(t)\big)R_j^a(t)\tau|\mathbf{Z}(t)\right],
\end{aligned} \tag{B.1}$$

where (a) is due to the assumption of $\mathbf{Z}(0) = 0$, and (b) is based on (9). Moreover, $E(\varpi)$ is an arbitrary solution from the Lyapunov function. Let $T \longrightarrow +\infty$, (B.1) can be rewritten as

$$\begin{aligned}
&\lim_{T\longrightarrow+\infty}\frac{1}{T}\mathbb{E}[L(\mathbf{Z}(T))] + V\lim_{T\longrightarrow+\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[\sum_{i\in\mathcal{J}} p_i^a(t)\tau\right] \\
&\leq \sum_{i\in\mathcal{J}}\frac{Z_{j1} + Z_{j2}}{2} + VE(\varpi) \\
&\quad - \lim_{T\longrightarrow+\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[\sum_{j\in\mathcal{J}}\big(Q_j(t) + Z_j(t)\big)R_j^a(t)\tau|\mathbf{Z}(t)\right],
\end{aligned} \tag{B.2}$$

As a stable network, it satisfies$\lim_{T\longrightarrow+\infty}(L(Z(T))/T) = 0$. Hence, (B.2) can be reduced to

$$\begin{aligned}
S^\beta &= \lim_{T\longrightarrow+\infty}\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[\sum_{i\in\mathcal{J}} p_i^a(t)\tau\right] \leq \sum_{j\in\mathcal{J}}\frac{Z_{j1} + Z_{j2}}{2V} + E(\varpi) \\
&\quad - \lim_{T\longrightarrow+\infty}\frac{1}{TV}\sum_{t=0}^{T-1}\mathbb{E}\left[\sum_{j\in\mathcal{J}}\big(Q_j(t) + Z_j(t)\big)R_j^a(t)\tau|\mathbf{Z}(t)\right] \\
&\leq \sum_{j\in\mathcal{J}}\frac{Z_{j1} + Z_{j2}}{2V} + E(\varpi).
\end{aligned} \tag{B.3}$$

Thereafter, let $\varpi \longrightarrow \varpi^*$, we can obtain

$$S^{\beta} \leq \sum_{j \in \mathcal{J}} \frac{Z_{j1} + Z_{j2}}{2V} + S^{\alpha}. \tag{B.4}$$

This ends the proof of (14). □

C

Proof. of Theorem 2.

Without loss of generation, we assume that each agent has only one variable, i.e., $\mathbf{X}_k \in \mathbf{R}$. A new vector $\mathbf{S} = \mathrm{col}(\mathbf{X}, \Lambda, \mathbf{Z})$ is defined, and its feasible region is $\overline{\Omega} = \Omega \times \mathbf{R}^{\mathbf{K}} \times \mathbf{R}^{\mathbf{K}}$. By applying the gradient flow, the updating rules of all agents can be expressed as

$$\dot{S} = \mathbf{P}_{\overline{\Omega}}(\mathbf{S} - \mathbf{F}(\mathbf{S})) - \mathbf{S}, \tag{C.1}$$

where the vector function $F(\mathbf{S})$ is denoted as

$$F(\mathbf{S}) = \begin{pmatrix} \nabla f(\mathbf{X}) - \mathbf{G}\boldsymbol{\Lambda} \\ \mathbf{L}\boldsymbol{\Lambda} + \mathbf{L}\mathbf{Z} - (\mathbf{I} - \mathbf{G}\mathbf{X}) \\ -\mathbf{L}\boldsymbol{\Lambda} \end{pmatrix}. \tag{C.2}$$

To prove the convergence of the proposed algorithm, the Lyapunov stability can be applied. First, we introduce a Lyapunov function to analyze the Lyapunov stability, which is given by

$$V_g = -\langle F(\mathbf{S}), H(\mathbf{S}) - \mathbf{S} \rangle - \frac{1}{2}\|H(\mathbf{S}) - \mathbf{S}\|_2^2 + \frac{1}{2}\|\mathbf{S} - \mathbf{S}^*\|_2^2, \tag{C.3}$$

where $H(\mathbf{S}) = P_{\overline{\Omega}}(\mathbf{S} - F(\mathbf{S}))$, $\mathbf{S}^* = \mathrm{col}(\mathbf{X}^*, \Lambda^*, \mathbf{Z}^*)$, $\mathbf{X}^*$ is the optimal solution, and Λ^* and Z^* are the optimal multipliers. To guarantee the Lyapunov stability of the equilibrium point, V_g must be constant positive or constant negative. Due to

$$\begin{aligned} &-\langle F(\mathbf{S}), H(\mathbf{S}) - \mathbf{S} \rangle - \frac{1}{2}\langle H(\mathbf{S}) - \mathbf{S}, H(\mathbf{S}) - \mathbf{S} \rangle \\ &\geq \frac{1}{2}\|\mathbf{S} - H(\mathbf{S})\|_2^2, \end{aligned} \tag{C.4}$$

we can obtain that $V_g \geq (1/2)\|\mathbf{S} - H(\mathbf{S})\|_2^2 + (1/2)\|\mathbf{S} - \mathbf{S}^*\| \geq 0$ and $V_q = 0$ if and only if $\mathbf{S} = \mathbf{S}^*$.

Since any asymmetric variational inequality can be converted to a differentiable optimization problem [21], V_g can be expressed as

$$\begin{aligned} \dot{V}_g = &\left(F(\mathbf{S}) - \left(J_F(\mathbf{S}) - I\right)(H(\mathbf{S}) - \mathbf{S}) + \mathbf{S} - \mathbf{S}^*\right)^T \\ &(H(\mathbf{S}) - \mathbf{S}), \end{aligned} \tag{C.5}$$

where $J_F(\mathbf{S})$ is the Jacobian matrix of $F(\mathbf{S})$, which is defined as

$$J_F(\mathbf{S}) = \begin{pmatrix} \nabla^2 R(\mathbf{X}) & -\mathbf{D} & 0 \\ \mathbf{D} & \mathbf{L} & \mathbf{L} \\ 0 & -\mathbf{L} & 0 \end{pmatrix}. \tag{C.6}$$

According to the basic property of projection operation, it can be given that

$$\langle \mathbf{S} - F(\mathbf{S}) - H(\mathbf{S}), H(\mathbf{S}) - \mathbf{S}^* \rangle \geq 0, \tag{C.7}$$

which is equivalent to

$$\langle \mathbf{S} - F(\mathbf{S}) - H(\mathbf{S}), H(\mathbf{S}) - \mathbf{S} + \mathbf{S} - \mathbf{S}^* \rangle \geq 0. \tag{C.8}$$

Then, we can infer

$$\begin{aligned} &\langle \mathbf{S} - \mathbf{S}^* + F(\mathbf{S}), H(\mathbf{S}) - \mathbf{S} \rangle + \|H(\mathbf{S}) - \mathbf{S}\|_2^2 \\ &\leq -\langle F(\mathbf{S}), \mathbf{S} - \mathbf{S}^* \rangle. \end{aligned} \tag{C.9}$$

With the assumption that $R(\mathbf{X}_k)$ are strictly convex and the information exchanging graph is undirected and connected, the following result can be obtained:

$$\mathbf{S}^T J_F(\overline{\mathbf{S}})\mathbf{S} = \mathbf{X}^T \nabla^2 R(\overline{\mathbf{X}})\mathbf{X} + \boldsymbol{\Lambda}^T \mathbf{L}\boldsymbol{\Lambda} > 0, \quad \forall \overline{\mathbf{S}} \in \overline{\boldsymbol{\Omega}}. \tag{C.10}$$

Based on (C.9) and (C.10), $\dot{V}_g$ can be further analyzed as follows:

$$\begin{aligned} \dot{V}_g = &-(H(\mathbf{S}) - \mathbf{S})^T J_F(\mathbf{S})(H(\mathbf{S}) - \mathbf{S}) + \|H(\mathbf{S}) - \mathbf{S}\|_2^2 \\ &+ \langle \mathbf{S} - \mathbf{S}^* + F(\mathbf{S}), H(\mathbf{S}) - \mathbf{S} \rangle \\ &\leq -\langle F(\mathbf{S}), \mathbf{S} - \mathbf{S}^* \rangle. \end{aligned} \tag{C.11}$$

Since $\langle F(\mathbf{S}), \mathbf{S} - \mathbf{S}^* \rangle \geq 0, \dot{V}_g \leq 0$, and the equilibrium point is assumed to be $\mathbf{S}^*$. Therefore, the equilibrium point is Lyapunov stable. Hence, the proposed distributed algorithm can converge to its equilibrium point, which ends the proof. □

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Y. Zhang, Q. Xia, and Y. Xu, "Multi-robot cooperation system based on wireless network," in *Proceedings of the 2019 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced*

Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), pp. 162–165, Leicester, UK, February 2019.

[2] Y. Li, F. Du, P. Wang, and Y. Yin, "A multi-robot topic communication method based on tcp and udp," in *Proceedings of the 2020 Chinese Automation Congress (CAC)*, pp. 668–672, Shanghai, China, November 2020.

[3] J. Karjee, S. Behera, H. Kumar Rath, and A. Simha, "Distributed cooperative communication and link prediction in cloud robotics,," in *Proceedings of the 2017 IEEE International Conference on Sensing, Communication and Networking (SECON Workshops)*, pp. 1–7, San Diego, CA, USA, June 2017.

[4] M. Fu, S. Sun, K. Ni, and X. Hou, "Mobile robot object recognition in the internet of things based on fog computing," in *Proceedings of the 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 1838–1842, Lanzhou, China, November 2019.

[5] F. Ahmed, A. A. Dowhuszko, and O. Tirkkonen, "Distributed algorithm for downlink resource allocation in multicarrier small cell networks,," in *Proceedings of the 2012 IEEE International Conference on Communications (ICC)*, pp. 6802–6808, Ottawa, ON, Canada, June 2012.

[6] J. Xiao, C. Yang, A. Anpalagan, Q. Ni, and M. Guizani, "Joint interference management in ultra-dense small-cell networks: a multi-domain coordination perspective," *IEEE Transactions on Communications*, vol. 66, no. 11, pp. 5470–5481, 2018.

[7] M. Chen, W. Saad, and C. Yin, "Echo state networks for self-organizing resource allocation in LTE-U with uplink-downlink decoupling," *IEEE Transactions on Wireless Communications*, vol. 16, no. 1, pp. 3–16, 2017.

[8] R. Yin, S. Liu, G. Yu, Y. Zhang, and Q. Chen, "Semi-distributed joint power and spectrum allocation for laa based small cell networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 6, pp. 4141–4153, 2020.

[9] F. Vhora and J. Gandhi, "A comprehensive survey on mobile edge computing: challenges, tools, applications," in *Proceedings of the 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)*, pp. 49–55, Erode, India, April 2020.

[10] M. Satyanarayanan, "The emergence of edge computing," *Computer*, vol. 50, no. 1, pp. 30–39, 2017.

[11] M. McClellan, C. Cervelló-Pastor, and S. Sallent, "Deep learning at the mobile edge: opportunities for 5g networks," *Applied Sciences*, vol. 10, no. 14, p. 4735, 2020.

[12] D. Wu, J. Wang, Y. Cai, and M. Guizani, "Millimeter-wave multimedia communications: challenges, methodology, and applications," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 232–238, 2015.

[13] S. Barbarossa, E. Ceci, M. Merluzzi, and E. Calvanese-Strinati, "Enabling effective mobile edge computing using millimeterwave links," in *Proceedings of the 2017 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 367–372, Paris, France, May 2017.

[14] F. Guo, L. Ma, H. Zhang, H. Ji, and X. Li, "Joint load management and resource allocation in the energy harvesting powered small cell networks with mobile edge computing," in *Proceedings of the IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOMWKSHPS)*, pp. 299–304, Piscataway, NJ, USA, April 2018.

[15] Q. Han, G. Zheng, and C. Xu, "Energy-efficient resource allocation for mobile edge computing in noma-enabled small cell networks," in *Proceedings of the 2020 IEEE 20th International Conference on Communication Technology (ICCT)*, pp. 415–419, China, Nanning, October 2020.

[16] H. Shokri-Ghadikolaei, C. Fischione, G. Fodor, P. Popovski, and M. Zorzi, "Millimeter wave cellular networks: a mac layer perspective," *IEEE Transactions on Communications*, vol. 63, no. 10, pp. 3437–3458, 2015.

[17] Q. Chen, X. Xu, H. Jiang, and X. Liu, "An energy-aware approach for industrial internet of things in 5g pervasive edge computing environment," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 7, pp. 5087–5097, 2021.

[18] M. J. Neely, "A lyapunov optimization approach to repeated stochastic games," in *Proceedings of the 2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 1082–1089, Monticello, IL, USA, October 2013.

[19] P. Yi, Y. Hong, and F. Liu, "Initialization-free distributed algorithms for optimal resource allocation with feasibility constraints and its application to economic dispatch of power systems," *Automatica*, vol. 74, 2015.

[20] M. Mesbahi and M. Egerstedt, "Graph theoretic methods in multiagent networks," *Princeton University Press*, vol. 33, 2010.

[21] M. Fukushima, "Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems," *Mathematical Programming*, vol. 53, no. 1, pp. 99–110, 1992.

Hindawi
Complexity
Volume 2021, Article ID 4646501, 12 pages
https://doi.org/10.1155/2021/4646501

Research Article

On Computation and Analysis of Topological Index-Based Invariants for Complex Coronoid Systems

Muhammad Aamer Rashid,[1] **Sarfraz Ahmad**,[1] **Muhammad Kamran Siddiqui**,[1] **and Mohammed K. A. Kaabar**[2,3,4]

[1]*Department of Mathematics, COMSATS University Islamabad,, Lahore Campus, Lahore 54000, Pakistan*
[2]*Gofa Campus, Near Gofa Industrial College and German Adebabay, Nifas Silk-Lafto, 26649, Addis Ababa, Ethiopia*
[3]*Jabalia Campus, United Nations Relief and Works Agency (UNRWA) Palestinian Refugee Camp, Gaza Strip Jabalya, State of Palestine*
[4]*Institute of Mathematical Sciences, Faculty of Science, University of Malaya, Kuala Lumpur 50603, Malaysia*

Correspondence should be addressed to Mohammed K. A. Kaabar; mohammed.kaabar@wsu.edu

Received 8 May 2021; Accepted 22 September 2021; Published 18 October 2021

Academic Editor: Miaomiao Wang

In chemical graph theory, benzenoid systems are interrogated as they exhibit the chemical compounds known as benzenoid hydrocarbons. Benzenoid schemes are circumscribed as planar connected finite graphs having no cut vertices wherein the entire internal sections are collaboratively congruent regular hexagon. The past couple of decennium has acknowledged an extravagant development regarding implementation of information theoretic framework in miscellaneous ramification of science, for instance, in social sciences, biological, physical, and engineering. Explicitly, this tremendous improvement has been outstanding in the field of soft computing, molecular biology, and information technology. The information theory, delineated by Claud Shannon, has no less importance when it was considered. Shannon put forwarded the apprehension of entropy to enumerate upper bounds in transmission rates in telephonic channels, in optical communication, and in wireless. The prestigious feature of entropy is that it entitles the amount of uncertainty in a system. The substantial participation of this paper is to explore characteristics of graph entropies and then keep moving forward to talk about the formation of coronoid polycyclic aromatic hydrocarbons. Likewise, we estimate entropies through precise topological indices established on degree of terminal nodes.

1. Introduction

Consider $\mathscr{D} = (\mathscr{D}_V, \mathscr{D}_E)$ be a graph containing $\mathscr{D}_V$ and $\mathscr{D}_E$ as the vertex set and the edge set of $\mathscr{D}$ correspondingly. The size and the order of $\mathscr{D}$ are expressed by m and n correspondingly, and $\widehat{I}(a)$ is characterized as the degree of any vertex a. Topological descriptors of a chemical structure are molecular descriptors. In QSPR/QSAR analyses, miscellaneous molecular descriptor is operated to correlate different biological and physico-chemical activities. In this study, we will talk about some degree-based indices.

To determine the unpredictability of a scheme, entropy is used [1]. This consideration was grown for analyzing the fundamental information of graphs. Laterally, it was employed substantially in graphs and chemical networks. The graph entropy [2] consideration demonstrated on the denominations of vertex orbits. Utilization of graph entropy is interdisciplinary [3].

In the literature, diverse graph entropies are estimated by means of degree of vertex, order of the graphs, eccentricity of the vertices, and characteristic polynomials [4, 5]. Over the past few years, graph entropies are estimated which are established on matchings, independent sets, and degree of vertices [6]. Mowshowits and Dehmer talked about few relations between the complexity of graphs and Hosoya entropy. We postulated that the presented degree-based entropy can be employed to assess network diversity. Equivalent entropy measures which are established on vertex-degrees to distinguish network diversity have been suggested by Tan and Wu [7].

In chemical graph theory, benzenoid structures are interrogated [8] since they exhibit the chemical compounds known as benzenoid hydrocarbons. Benzenoid schemes are circumscribed as planar connected finite graphs having no cut vertices wherein the entire internal sections are collaboratively congruent regular hexagons.

In [9], interpretation of the entropy was put forwarded for the edge weighted graph $\mathscr{D} = (\mathscr{D}_v; \mathscr{D}_E; \psi(az))$, in which $\mathscr{D}_V$, $\mathscr{D}_E$, and $\psi(az)$ symbolize the set of vertices, the set of edges, and the edge weight of the edge (az) in $\mathscr{D}$ correspondingly. Subsequently the entropy of the edge weighted graph is portrayed in the following equation. The relation between topological indices and corresponding entropy measures is presented in Table 1.

$$ENT_{\psi}(\mathscr{D}) = -\sum_{a'z' \in \mathscr{D}_e} \frac{\psi(a'z')}{\sum_{az \in \mathscr{D}_E} \psi(az)} \log\left[\frac{\psi(a'z')}{\sum_{az \in \mathscr{D}_E} \psi(az)}\right]. \tag{1}$$

2. Structure of Coronoid Polycyclic Aromatic Hydrocarbons

Coronoids are derived by these benzenoid systems by removing some interior vertices or edges. This will create a different interior region enclosed by a polygon having greater than six edges. By using the abovementioned conditions, more than one hexagon can be trimmed from the originator benzenoid structure. Coronoid is a benzenoid with a hole. It may have more than one hole. Graphical illustration of coronoid and noncoronoid systems is provided in Figures 1(a) and 1(b). It is to be noted that Figure 1(b) is noncoronoid because some of its edges do not belong to any of its hexagons [14].

tCycloarenes are macrocyclic combined compounds constituted by circumferentially connected benzene loops that enclose a hole with inner-directed carbon-hydrogen bonds. As a consequence, cycloarenes are connected with a subclass of circulenes or coronoids. The background of cycloarenes traces back to 1987; meanwhile, the main example with 12 benzene rings, categorized as kekulene, was disclosed by Staab and Diederich [15]. There have been countless hypothetical investigations focusing on the magnetic tendency, vibrational rate of occurrences of cycloarenes, and aromaticity [16]. Kekulene and cyclo decakis benzene were the merely two substantial patterns accessible for analysis [17].

Afterwards, Kumar et al. [18] synthesized another model of cycloarene, with 14 benzene rings, specifically septulene. The synthesis of kekulene and septulene has kick started numerous theoretical and experimental studies on coronoids [19]. The study of coronoids is also gaining momentum due to their superaromaticity. Superaromaticity or macrocyclic aromaticity is described as an additional thermodynamic consolidation as a consequence of macrocyclic association in tremendous-ring molecules like kekulene, and it constitutes a little contribution of universal aromaticity.

Whole coronoids investigated until now are approximately superaromatic through constructive superaromatic stabilization energies (SSEs). Intriguingly, the extent of SSE oscillates between single-layered and several-layered species. SSEs for even-layered coronoids [9, 20–22] are high, while those for odd-layered ones [5, 23, 24] are pretty low [23]. This provides an impetus for a deeper study into the properties of coronoid systems and their relationship with the underlying molecular structures. This study might be applicable to various fields of nanotechnology. As an illustration, the eradication of a unique carbon atom with a graphite framework establishes a one-atom hole referred to as a Schottky defect [25]. Individual-wall nanocones, [26] grime platelets, [27], and extended graphite layers [24] may contain vacancy hole defects involving larger (multiatom) holes which can be studied by modeling them as coronoids [28]. Graphenes are nanosized polycyclic aromatic hydrocarbons with potential uses in the fabrication of organic electronic devices [29]. The origination of coronoids by demonstrating a cavity in nanographene might be an efficacious approach to regulate their electronic and optical properties without amendments to their exterior structures. The cavities, that make an integral part of the coronoids, act as prototypes for scheming and synthesizing novel nanomaterials of significance in nano and biotechnology and the incipient field of nano-medicines. They have also been used in the design and synthesis of distinct porous and mesoporous materials grounded on calixarenes and mesoporous silica for the sequestration and complexation of toxic nuclear waste and other environmental pollutants [30] Coronoid systems are also widely examined in the study of coronoid hydrocarbons.

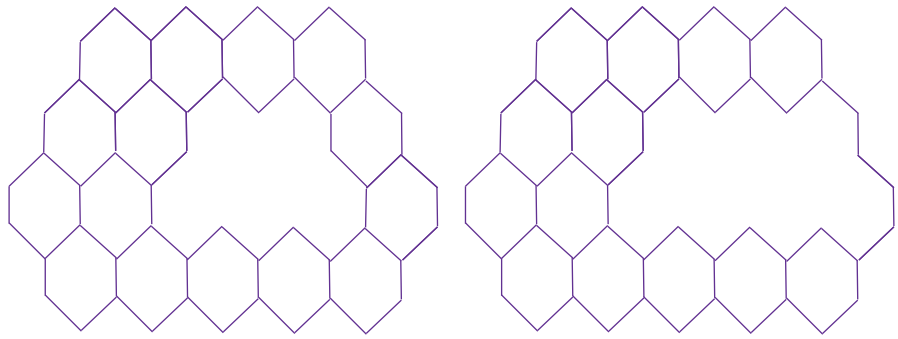
It has been proved that it is possible to compute the total π-electron energy, the resonance energy, and the enumeration of coronoid hydrocarbons accurately using the knowledge of coronoid structures [20]. The conjugate graph-theoretical circuit theory, inspired by Clar's aromatic sextet, correlates to the description of diverse enclosed consolidated cycles existing in the polycyclic aromatic compounds [31, 32]. This theory also provides combinatorial and graph-theoretical methods for efficient determination of the relative stabilities of coronoid structures, graphenes, cycloarenes, carbon nanotubes, and nanotori. For further information on the comprehensive research issued on coronoid systems by both Dias with coauthors and Cyvin with coauthors, refer to [21]. In comparison to the computationally intensive quantum chemical calculations, the graph-theoretical techniques are considerably more productive in obtaining the properties of coronoid systems. During a recent investigation, Aihara et al. [23] emphasized that the graph theory is not merely an extremely valuable mechanism in estimating topological resonance energies but additionally in uncovering significant challenges along the previous speculations of aromaticity. By employing graph-theoretical approaches, the investigation of coronoid networks has gained increased importance [22].

3. Coronoid System $C_1[u, v, w, x]$

In this fragment, we will take into consideration the single coronoid system. This system is also recognised as one hole benzenoid. It is extracted by eradicating few of the interior edges or vertices from the benzenoid system. In this

Table 1: Degree-based topological indices together with their corresponding edge weight $\psi(az)$ of the edge az.

Degree-based topological indices	Edge weight $\psi(az)$ of the edge az	Entropies
The randić index [10]	$(\widehat{I}(r) \times \widehat{I}(s))^{\alpha}$, $\alpha = 1, -1, 1/2, -1/2$	The randić entropy [11]
The atom bond connectivity index [12]	$\sqrt{\widehat{I}(r) + \widehat{I}(s) - 2/\widehat{I}(r) \times \widehat{I}(s)}$	The atom bond connectivity entropy [11]
The geometric arithmetic index [13]	$2\sqrt{\widehat{I}(r) \times \widehat{I}(s)/\widehat{I}(r) + \widehat{I}(s)}$.	The geometric arithmetic entropy [11]



(a) (b)

Figure 1: Coronoid and noncoronoid systems.

procedure, a hole is created within the system having a lowest size of two benzenes. The x-circumscribed basic coronoid system is defined as $C_k[u, v, w, x]$, $k = 1, 2$, in which $u \geq 3$, $v, w \geq 1$, and $x \geq 1$. The coronoid structure fluctuates correspondingly to its parameters u, v, w, and x bringing escalation to the particular cases. In [21, 33], some particular cases are discussed, and these exclusive models are employed to prognosticate the resonance energy of aromatic molecules. In theoretical chemistry, these models are considered as ideal models to investigate conjugation circuits of π electrons.

These special cases are used to predict the resonance energy of aromatic molecules and have attracted a great deal of interest in the field of theoretical chemistry as ideal models to explore conjugation circuits of π-electrons [34, 35]. We will represent coronoid structure as $C_1[u, v, w, x]$ in which $u \geq 3$ and $1 \leq v \leq w \leq u$. Figure 2 illustrates $C_1[u, v, w, x]$ graphically.

Table 2 portrays the edge partition of $C_1[u, v, w, x]$ on the basis of degrees of terminal vertices of each edge.

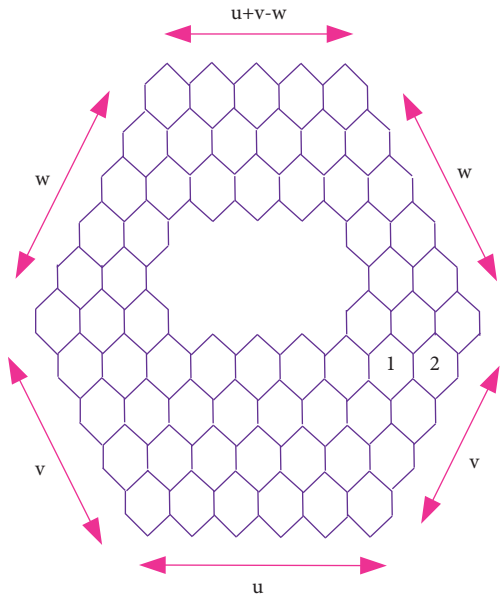


Figure 2: Coronoid system of $C_1[u, v, w, x]$ [34].

3.1. Results for Coronoid System $C_1[u, v, w, x]$

3.1.1. Randić Entropy of $C_1[u, v, w, x]$. We enumerate the Randic index for $\alpha = 1, -1, 1/2, -1/2$ with the help of Tables 1 and 2 as follows:

$$R_{\alpha}(C_1[u, v, w, x]) = (6) \times (4)^{\alpha} + (4u + 6v + 2w + 22) \times (6)^{\alpha} + \left((6x + 8)u + (9x + 12)v + (3x + 4)w + 9x^2 + 3x - 58\right) \times (9)^{\alpha}. \quad (2)$$

Therefore, equation (1) with Tables 1 and 2 is in the form as follows:

Table 2: Edge partition of the coronoid system $C_1[u,v,w,x]$ on the basis of degrees of terminal vertices of each edge.

$(\widehat{I}(a),\widehat{I}(z))$	Number of replication	Classes of edges
$(2,2)$	6	$\mathscr{D}_{E_1}$
$(2,3)$	$4u+6v+2w+22$	$\mathscr{D}_{E_2}$
$(3,3)$	$(6x+8)u+(9x+12)v+(3x+4)w+9x^2+3x-58$	$\mathscr{D}_{E_3}$

$$
\begin{aligned}
\mathrm{ENT}_{R_\alpha}(C_1[u,v,w,x]) &= \log(R_\alpha)-\frac{1}{(R_\alpha)}\sum_{i=1}^{3}\sum_{az\in\mathscr{D}_{E_i}}\left[(\widehat{I}(a)\times\widehat{I}(z))^\alpha\right]\log\left[(\widehat{I}(a)\times\widehat{I}(z))^\alpha\right],\\
\mathrm{ENT}_{R_\alpha}(C_1[u,v,w,x]) &= \log(R_\alpha)-\frac{6}{(R_\alpha)}(4^\alpha)\log[4^\alpha]-\frac{(4u+6v+2w+22)}{(R_\alpha)}(6^\alpha)\log[6^\alpha]\\
&\quad-\frac{(6x+8)u+(9x+12)v+(3x+4)w+9x^2+3x-58}{(R_\alpha)}(9^\alpha)\log[9^\alpha],\\
\mathrm{ENT}_{R_1}(C_1[u,v,w,x]) &= \log\left[(96+54x)u+(144+81x)v+(48+27x)w+81x^2+27x-366\right]\\
&\quad-\frac{24\log[4]}{\left[(96+54x)u+(144+81x)v+(48+27x)w+81x^2+27x-366\right]}\\
&\quad-\frac{6(4u+6v+2w+22)\log[6]}{\left[(96+54x)u+(144+81x)v+(48+27x)w+81x^2+27x-366\right]}\\
&\quad-\frac{9\left((6x+8)u+(9x+12)v+(3x+4)w+x^2+3x-58\right)\log[9]}{\left[(96+54x)u+(144+81x)v+(48+27x)w+81x^2+27x-366\right]},\\
\mathrm{ENT}_{R_{1/2}}(C_1[u,v,w,x]) &= \log\left[\begin{matrix}(4\sqrt{6}+18x+24)u+(6\sqrt{6}+27x+36)v\\+(2\sqrt{6}+9x+12)w+27x^2+9x+22\sqrt{6}-162\end{matrix}\right]\\
&\quad-\frac{12\log[2]}{\left[(4\sqrt{6}+18x+24)u+(6\sqrt{6}+27x+36)v+(2\sqrt{6}+9x+12)w+27x^2+9x+22\sqrt{6}-162\right]}\\
&\quad-\frac{\sqrt{6}(4u+6v+2w+22)\log[\sqrt{6}]}{\left[(4\sqrt{6}+18x+24)u+(6\sqrt{6}+27x+36)v+(2\sqrt{6}+9x+12)w+27x^2+9x+22\sqrt{6}-162\right]}\\
&\quad-\frac{3\left((6x+8)u+(9x+12)v+(3x+4)w+x^2+3x-58\right)\log[3]}{\left[(4\sqrt{6}+18x+24)u+(6\sqrt{6}+27x+36)v+(2\sqrt{6}+9x+12)w+27x^2+9x+22\sqrt{6}-162\right]},\\
\mathrm{ENT}_{R_{-1}}(C_1[u,v,w,x]) &= \log\left[\left(\frac{14}{9}+\frac{2x}{3}\right)u+\left(\frac{7}{3}+x\right)v+\left(\frac{7}{9}+\frac{x}{3}\right)w+x^2+\frac{x}{3}-\frac{23}{18}\right]\\
&\quad+\frac{(3/2)\log[4]}{\left[((14/9)+(2x/3))u+((7/3)+x)v+((7/9)+(x/3))w+x^2+(x/3)-(23/18)\right]}\\
&\quad+\frac{(1/6)(4u+6v+2w+22)\log[6]}{\left[((14/9)+(2x/3))u+((7/3)+x)v+((7/9)+(x/3))w+x^2+(x/3)-(23/18)\right]}\\
&\quad+\frac{(1/9)\left((6x+8)u+(9x+12)v+(3x+4)w+x^2+3x-58\right)\log[9]}{\left[((14/9)+(2x/3))u+((7/3)+x)v+((7/9)+(x/3))w+x^2+(x/3)-(23/18)\right]},\\
\mathrm{ENT}_{R_{(-1/2)}}(C_1[u,v,w,x]) &= \log\left[\left(\frac{4}{\sqrt{6}}+2x+\frac{8}{3}\right)u+(\sqrt{6}+3x+4)v+\left(\frac{2}{\sqrt{6}}+x+\frac{4}{3}\right)w+3x^2+x+\frac{22}{\sqrt{6}}-\frac{49}{3}\right]\\
&\quad+\frac{(3/2)\log[2]}{\left[((4/\sqrt{6})+2x+(8/3))u+(\sqrt{6}+3x+4)v+((2/\sqrt{6})+x+(4/3))w+3x^2+x+(22/\sqrt{6})-(49/3)\right]}\\
&\quad+\frac{(1/\sqrt{6})(4u+6v+2w+22)\log[\sqrt{6}]}{\left[((4/\sqrt{6})+2x+(8/3))u+(\sqrt{6}+3x+4)v+((2/\sqrt{6})+x+(4/3))w+3x^2+x+(22/\sqrt{6})-(49/3)\right]}\\
&\quad+\frac{(1/3)\left((6x+8)u+(9x+12)v+(3x+4)w+x^2+3x-58\right)\log[3]}{\left[((4/\sqrt{6})+2x+(8/3))u+(\sqrt{6}+3x+4)v+((2/\sqrt{6})+x+(4/3))w+3x^2+x+(22/\sqrt{6})-(49/3)\right]}.
\end{aligned}
\tag{3}
$$

3.1.2. The Atom Bond Connectivity Entropy of $C_1[u,v,w,x]$. Simple calculations with Tables 1 and 2 yield the atom bond connectivity index as follows:

$$ABC(C_1[u,v,w,x]) = \left(2\sqrt{2} + 4x + \frac{16}{3}\right)u + (3\sqrt{2} + 6x + 8)v + \left(\sqrt{2} + 2x + \frac{8}{3}\right)w + 6x^2 + 2x + 14\sqrt{2} - \frac{116}{3}. \tag{4}$$

Therefore, equation (1) with Tables 1 and 2 is in the form as follows:

$$\mathrm{ENT}_{ABC}(C_1[u,v,w,x]) = \log(ABC) - \frac{1}{(ABC)}\sum_{i=1}^{3}\sum_{az\in\mathscr{D}_{E_i}}\left[\sqrt{\frac{\widehat{I}(a)+\widehat{I}(z)-2}{\widehat{I}(a)\times\widehat{I}(z)}}\right]\log\left[\sqrt{\frac{\widehat{I}(a)+\widehat{I}(z)-2}{\widehat{I}(a)\times\widehat{I}(z)}}\right],$$

$$\mathrm{ENT}_{ABC}(C_1[u,v,w,x]) = \log\left[\left(2\sqrt{2} + 4x + \frac{16}{3}\right)u + (3\sqrt{2} + 6x + 8)v + \left(\sqrt{2} + 2x + \frac{8}{3}\right)w + 6x^2 + 2x + 14\sqrt{2} - \frac{116}{3}\right] - \frac{6[1/\sqrt{2}]\log[1/\sqrt{2}]}{\left[(2\sqrt{2} + 4x + (16/3))u + (3\sqrt{2} + 6x + 8)v + (\sqrt{2} + 2x + (8/3))w + 6x^2 + 2x + 14\sqrt{2} - (116/3)\right]} - \frac{(4u+6v+2w+22)[1/\sqrt{2}]\log[1/\sqrt{2}]}{\left[(2\sqrt{2} + 4x + (16/3))u + (3\sqrt{2} + 6x + 8)v + (\sqrt{2} + 2x + (8/3))w + 6x^2 + 2x + 14\sqrt{2} - (116/3)\right]} - \frac{\left((6x+8)u + (9x+12)v + (3x+4)w + x^2 + 3x - 58\right)[2/3]\log[2/3]}{\left[(2\sqrt{2} + 4x + (16/3))u + (3\sqrt{2} + 6x + 8)v + (\sqrt{2} + 2x + (8/3))w + 6x^2 + 2x + 14\sqrt{2} - (116/3)\right]}. \tag{5}$$

3.1.3. The Geometric Arithmetic Entropy $C_1[u,v,w,x]$. Simple calculations with Tables 1 and 2 yield the geometric arithmetic index as follows:

$$\mathrm{GA}(G) = \left(\frac{8\sqrt{6}}{5} + 6x + 8\right)u + \left(\frac{12\sqrt{6}}{5} + 9x + 12\right)v + \left(\frac{4\sqrt{6}}{5} + 3x + 4\right)w + 9x^2 + 3x - 52 + \frac{44\sqrt{6}}{5}. \tag{6}$$

Therefore, equation (1) with Tables 1 and 2 is in the form as follows:

$$\mathrm{ENT}_{\mathrm{GA}}(C_1[u,v,w,x]) = \log(\mathrm{GA}) - \frac{1}{(\mathrm{GA})}\sum_{i=1}^{3}\sum_{az\in\mathscr{D}_{E_i}}\left[\frac{2\sqrt{\widehat{I}(a)\times\widehat{I}(z)}}{\widehat{I}(a)+\widehat{I}(z)}\right]\log\left[\frac{2\sqrt{\widehat{I}(a)\times\widehat{I}(z)}}{\widehat{I}(a)+\widehat{I}(z)}\right],$$

$$\mathrm{ENT}_{\mathrm{GA}}(C_1[u,v,w,x]) = \log\left(\left(\frac{8\sqrt{6}}{5} + 6x + 8\right)u + \left(\frac{12\sqrt{6}}{5} + 9x + 12\right)v + \left(\frac{4\sqrt{6}}{5} + 3x + 4\right)w + 9x^2 + 3x - 52 + \frac{44\sqrt{6}}{5}\right) - \frac{(4u+6v+2w+22)[2\sqrt{6}/5]\log[2\sqrt{6}/5]}{\left(((8\sqrt{6}/5) + 6x + 8)u + ((12\sqrt{6}/5) + 9x + 12)v + ((4\sqrt{6}/5) + 3x + 4)w + 9x^2 + 3x - 52 + (44\sqrt{6}/5)\right)}. \tag{7}$$

4. Coronoid System $C_1[u, v, w, x]$

The coronoid system of $\mathscr{D}$ is delineated as $C_2[n, p, q, r]$, in which $x \geq 1$, $u \geq 3$, and $1 \leq v < w \leq u$. Figure 3 shows the coronoid system of $C_2[u, v, w, x]$ [34, 35].

The approach we will use here is to form the partitions of the edges of $C_2[u, v, w, x]$ of terminal vertices of each edge (see Table 3).

4.1. Results for Coronoid System $C_2[u, v, w, x]$

4.1.1. Randic Entropy of $C_2[u, v, w, x]$. We enumerate the Randic index for $\alpha = 1, -1, 1/2, -1/2$ with the help of Tables 3 and 1 as follows:

$$R_\alpha(C_2[u, v, w, x]) = (6) \times (4)^\alpha + (4u + 4v + 4w + 24) \times (6)^\alpha + \left((6u + 6v + 6w)x + 9x^2 - 9x - 40\right) \times (9)^\alpha. \tag{8}$$

As a consequence, equation (1) with Tables 1 and 3 is embodied in the form as follows:

$$\begin{aligned}
\mathrm{ENT}_{R_\alpha}(C_2[u, v, w, x]) &= \log(R_\alpha) - \frac{1}{(R_\alpha)} \sum_{i=1}^{3} \sum_{az \in \mathscr{D}_{E_i}} \left[(\hat{I}(a) \times \hat{I}(z))^\alpha \right] \log\left[(\hat{I}(a) \times \hat{I}(z))^\alpha \right], \\
\mathrm{ENT}_{R_1}(C_2[u, v, w, x]) &= \log\left((24 + 54x)(u + v + w) + 81x^2 - 81x - 192\right) \\
&\quad - \frac{24 \log[4]}{\left((24 + 54x)(u + v + w) + 81x^2 - 81x - 192\right)} \\
&\quad - \frac{6(4u + 4v + 4w + 24)\log[6]}{\left((24 + 54x)(u + v + w) + 81x^2 - 81x - 192\right)} - \frac{9\left((6u + 6v + 6w)r + 9x^2 - 9x - 40\right)\log[9]}{\left((24 + 54x)(u + v + w) + 81x^2 - 81x - 192\right)}, \\
\mathrm{ENT}_{R_{-1}}(C_2[u, v, w, x]) &= \log\left(\frac{2}{3}(1 + x)(u + v + w)x^2 - x + \frac{19}{18}\right) + \frac{(3/2)\log[4]}{\left((2/3)(1 + x)(u + v + w)x^2 - x + (19/18)\right)} \\
&\quad + \frac{(1/6)(4u + 4v + 4w + 24)\log[6]}{\left((2/3)(1 + x)(u + v + w)x^2 - x + (19/18)\right)} + \frac{(1/9)\left((6u + 6v + 6w)x + 9x^2 - 9x - 40\right)\log[9]}{\left((2/3)(1 + x)(u + v + w)x^2 - x + (19/18)\right)}, \\
\mathrm{ENT}_{R_{1/2}}(K_2[n, p, q, r]) &= \log\left((4\sqrt{6} + 18)(n + p + q) + 27r^2 - 27r + 24\sqrt{6} - 108\right) \\
&\quad - \frac{12 \log[2]}{\left((4\sqrt{6} + 18)(n + p + q) + 27r^2 - 27r + 24\sqrt{6} - 108\right)} \\
&\quad - \frac{\sqrt{6}(4u + 4v + 4w + 24)\log[\sqrt{6}]}{\left((4\sqrt{6} + 18)(n + p + q) + 27r^2 - 27r + 24\sqrt{6} - 108\right)} \\
&\quad - \frac{3\left((6u + 6v + 6w)x + 9x^2 - 9x - 40\right)\log[3]}{\left((4\sqrt{6} + 18)(n + p + q) + 27r^2 - 27r + 24\sqrt{6} - 108\right)}, \\
\mathrm{ENT}_{R_{(-1/2)}}(C_2[u, v, w, x]) &= \log\left(\left(\frac{4}{\sqrt{6}} + 2x\right)(u + v + w) + 3x^2 - 3x + 4\sqrt{6} - \frac{31}{3}\right) \\
&\quad + \frac{(3/2)\log[2]}{\left(((4/\sqrt{6}) + 2x)(u + v + w) + 3x^2 - 3x + 4\sqrt{6} - (31/3)\right)} \\
&\quad + \frac{\sqrt{6}(4u + 4v + 4w + 24)\log[\sqrt{6}]}{(((4/\sqrt{6}) + 2x))(u + v + w) + 3x^2 - 3x + 4\sqrt{6} - (31/3)} \\
&\quad + \frac{(1/3)\left((6u + 6v + 6w)x + 9x^2 - 9x - 40\right)\log[3]}{\left(((4/\sqrt{6}) + 2x)(u + v + w)t + n3qx^2h -_3 xx7 + C4; \sqrt{6} - (31/3)\right)}.
\end{aligned} \tag{9}$$

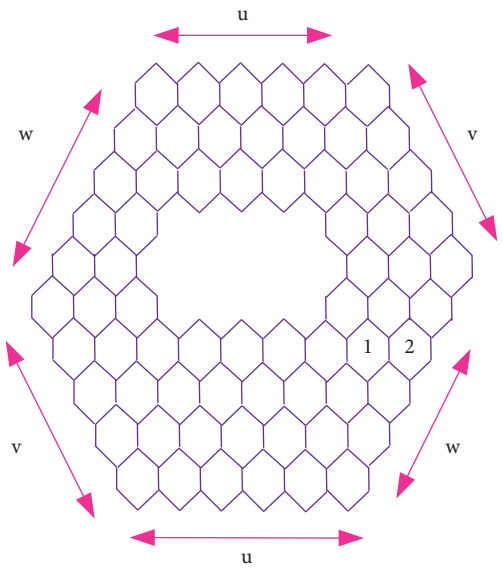


Figure 3: Coronoid system of $C_2[u, v, w, x]$ [35].

Table 3: Edge partition of $C_2[u, v, w, x]$ established on degrees of terminal vertices of each edge.

$(\hat{I}(a), \hat{I}(z))$	Number of replication	Classes of edges
$(2, 2)$	6	$\mathcal{G}_{E_1}$
$(2, 3)$	$4u + 4v + 4w + 24$	$\mathcal{G}_{E_2}$
$(3, 3)$	$(6u + 6v + 6w)x + 9x^2 - 9x - 40$	$\mathcal{G}_{E_3}$

4.1.2. The Atom Bond Connectivity Entropy of $C_2[u, v, w, x]$. Simple calculations with Tables 1 and 3 yield the atom bond connectivity index as follows:

$$ABC(C_2[u, v, w, x]) = (2\sqrt{2} + 4x)(u + v + w) + 6x^2 - 6x + 5\sqrt{2} - \frac{80}{3}. \tag{10}$$

Therefore, equation (1) with Tables 1 and 3 is in the form as follows:

$$\mathrm{ENT}_{ABC}(C_2[u, v, w, x]) = \log(ABC) - \frac{1}{(ABC)} \sum_{i=1}^{3} \sum_{az \in \mathcal{D}_{E_i}} \left[\sqrt{\frac{\hat{I}(a) + \hat{I}(z) - 2}{\hat{I}(a) \times \hat{I}(z)}}\right] \log\left[\sqrt{\frac{\hat{I}(a) + \hat{I}(z) - 2}{\hat{I}(a) \times \hat{I}(z)}}\right],$$

$$\mathrm{ENT}_{ABC}(C_2[u, v, w, x]) = \log(ABC) - \frac{[6/\sqrt{2}]}{(ABC)} \log\left[\frac{1}{\sqrt{2}}\right] - \frac{(4u + 4v + 4w + 24)}{(ABC)} \left[\frac{1}{\sqrt{2}}\right] \log\left[\frac{1}{\sqrt{2}}\right] - \frac{(6u + 6v + 6w)x + 9x^2 - 9x - 40}{(ABC)} \left[\frac{2}{3}\right] \log\left[\frac{2}{3}\right]. \tag{11}$$

4.1.3. The Geometric Arithmetic Entropy of $C_2[u, v, w, x]$. Simple calculations with Tables 1 and 3 yield the geometric arithmetic index as follows:

$$\mathrm{GA}(G) = \left(\frac{8\sqrt{6}}{5} + 6x\right)(u + v + w) + 9x^2 - 9x - 34 + \frac{48\sqrt{6}}{5}. \tag{12}$$

Table 4: Comparison of Randic entropies for $C_1[u,1,1,x]$.

$[u,x]$	ENT_{R_1}	$\mathrm{ENT}_{R_{-1}}$	$\mathrm{ENT}_{R_{1/2}}$	$\mathrm{ENT}_{R_{-1/2}}$
[3, 1]	0.1132	1.8445	1.8542	1.8540
[4, 2]	0.9964	2.1801	2.1906	2.0910
[5, 3]	1.3732	2.4213	2.4296	2.2841
[6, 4]	1.6202	2.6094	2.6157	2.4560
[7, 5]	1.8065	2.7634	2.7684	2.6219
[8, 6]	1.9571	2.8939	2.8980	2.7511
[9, 7]	2.0839	3.0072	3.0305	2.8709
[10, 8]	2.1938	3.1072	3.1701	3.0134
[11, 9]	2.2907	3.1969	3.2793	3.1253
[12, 10]	2.3776	3.2781	3.3802	3.2018

Table 5: Comparison of $\mathrm{ENT}_{\mathrm{ABC}}$ and $\mathrm{ENT}_{\mathrm{GA}}$ entropies for $C_1[u,v,w,x]$.

$[u,s]$	$\mathrm{ENT}_{\mathrm{ABC}}$	$\mathrm{ENT}_{\mathrm{GA}}$
[3, 1]	1.8572	1.8360
[4, 2]	2.203	2.2801
[5, 3]	2.4412	2.5220
[6, 4]	2.6269	2.7097
[7, 5]	2.7893	2.8834
[8, 6]	2.8987	2.9998
[9, 7]	3.0111	3.0967
[10, 8]	3.0806	3.1667
[11, 9]	3.1697	3.2563
[12, 10]	3.2605	3.3574

Therefore, equation (1) with Tables 1 and 3 is in the form as follows:

$$\mathrm{ENT}_{\mathrm{GA}}\left(C_2[u,v,w,x]\right) = \log(\mathrm{GA}) - \frac{1}{(\mathrm{GA})}\sum_{i=1}^{3}\sum_{az\in\mathcal{D}_{E_i}}\left[\frac{2\sqrt{\widehat{I}(a)\times\widehat{I}(z)}}{\widehat{I}(a)+\widehat{I}(z)}\right]\log\left[\frac{2\sqrt{\widehat{I}(a)\times\widehat{I}(z)}}{\widehat{I}(a)+\widehat{I}(z)}\right],$$

$$\mathrm{ENT}_{\mathrm{GA}}\left(C_2[u,v,w,x]\right) = \log\left(\left(\frac{8\sqrt{6}}{5}+6x\right)(u+v+w)+9x^2-9x-34+\frac{48\sqrt{6}}{5}\right) \tag{13}$$

$$-\frac{(4u+4v+4w+24)[2\sqrt{6}/5]\log[2\sqrt{6}/5]}{\left(((8\sqrt{6}/5)+6x)(u+v+w)t+n9qx^2h-_9\,xx7-C34;+(48\sqrt{6}/5)\right)}.$$

5. Comparisons and Discussion for $C_1[u,v,w,x]$

In QSPR/QSAR discussions, topological indices are applied to connect the biological functions of the anatomies with their corporeal properties like distortion, strain energy, stability, and melting point. These determinations can be accomplished by employing degree-based indices as these indices have clarity of decision and rapidity. In this research, we asseverated some degree-based entropies. We proposed a new approach to estimate the entropy by estimating its topological indices. The degree-based entropy can also be exerted to structural chemistry, ecological networks, biology, national security, social network, and so on. Additionally, to investigate structural symmetry and asymmetry in real networks, the values of entropies have a significant role. Entropy is monotonically increasing function as in all situations. Here, we estimated mathematically all degree-based entropies for diverse values of u and x keeping $p = 1$ and $q = 1$ for $C_1[u,v,w,x]$. Besides, we construct Tables 4 and 5 for tiny values of u, x, $p = 1$ and $q = 1$ for degree-based entropy to numerical comparison for the structure of $C_1[u,v,w,x]$. Now, numerical comparison is represented in Tables 4 and 5. Also, the graphical comparison is depicted in Figures 4–6.

6. Comparisons and Discussion for $C_2[u,v,w,x]$

Multiple exploitations of complex networks deserted on the entropy corresponding with structural information are discussed. In [36–39], many methods were suggested to

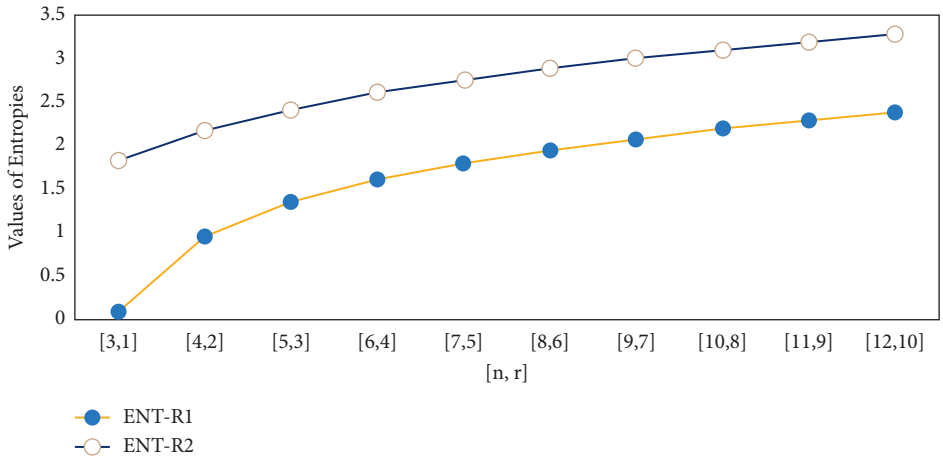


Figure 4: Comparison of Randic entropies for $\alpha = 1$ and for $\alpha = -1$.

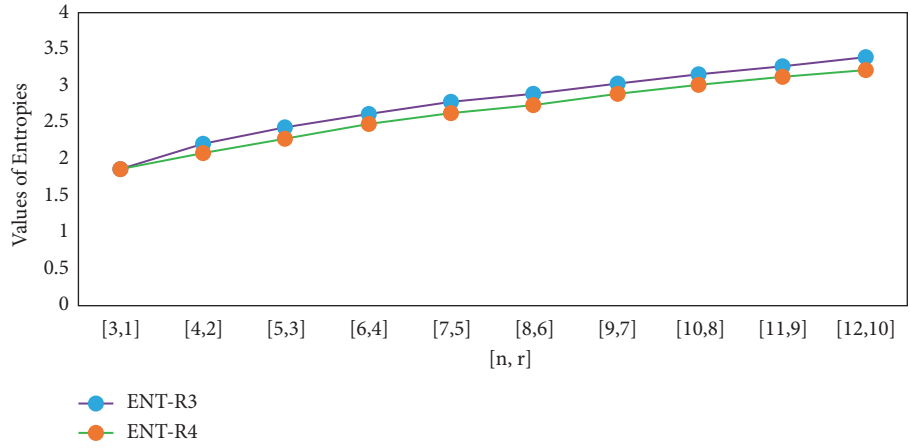


Figure 5: Comparison of Randic entropies for $\alpha = 1/2$ and for $\alpha = -1/2$.

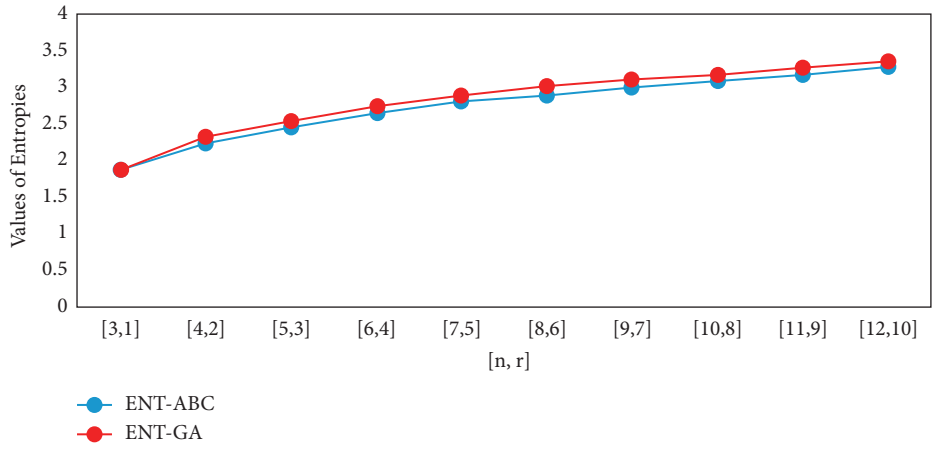


Figure 6: Comparison of atom bond connectivity entropy and geometric arithmetic entropy.

check the structural complexity. However, the entropy approach is considered to be the most significant approach to discriminate the structural information of the complex networks. Besides, degree-based indices have theoretical role on account of having the capability of computing pharmaceutical properties. Therefore, we have listed mathematically some degree-based entropies for little considerations of parameters u and x keeping $v = 1$ and $w = 2$ for $C_2[u, v, w, x]$ as depicted in Tables 6 and 7. Also, the graphical comparison is depicted in Figures 7–9.

Table 6: Comparison of Randic entropies for $C_2[u, v, w, x]$.

$[u, x]$	ENT_{R_1}	$ENT_{R_{-1}}$	$ENT_{R_{1/2}}$	$ENT_{R_{-1/2}}$
[3, 1]	1.8378	1.8380	1.8433	2.4669
[4, 2]	2.1715	2.2689	2.1791	2.5181
[5, 3]	2.4143	2.5208	2.4195	2.6359
[6, 4]	2.6032	2.7401	2.7070	2.7594
[7, 5]	2.7579	2.9152	2.7608	2.8754
[8, 6]	2.8890	3.0123	2.8912	2.9814
[9, 7]	3.0026	3.1307	3.0045	3.078
[10, 8]	3.1030	3.2114	3.1046	3.1661
[11, 9]	3.1930	3.2915	3.1943	3.2469
[12, 10]	3.2744	3.3732	3.2756	3.3213

Table 7: Comparison of ENT_{ABC} and ENT_{GA} entropies for $C_2[u, v, w, x]$.

$[u, x]$	ENT_{ABC}	ENT_{GA}
[3, 1]	1.7596	1.8451
[4, 2]	2.0944	2.1818
[5, 3]	2.3402	2.4216
[6, 4]	2.5347	2.6085
[7, 5]	2.7122	2.7819
[8, 6]	2.8249	2.8921
[9, 7]	2.9497	3.0052
[10, 8]	3.0583	3.1008
[11, 9]	3.1512	3.1912
[12, 10]	3.2531	3.2731

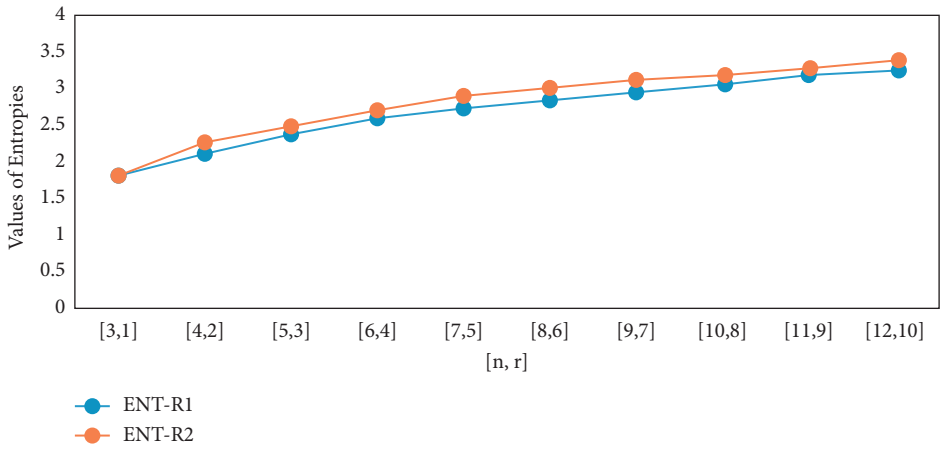


Figure 7: Comparison of Randic entropies for $\alpha = 1$ and for $\alpha = -1$ for the coronoid system $C_2[u, v, w, x]$.

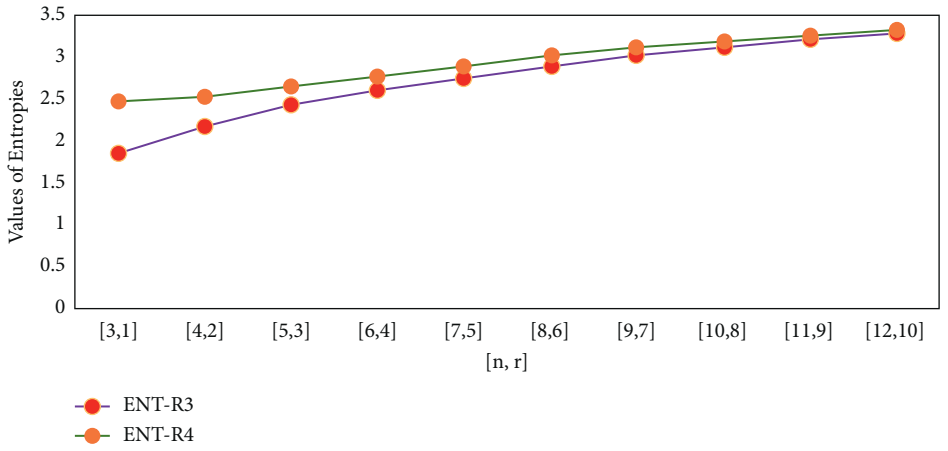


Figure 8: Comparison of Randic entropies for $\alpha = 1/2$ and for $\alpha = -1/2$ for the coronoid system $C_2[u, v, w, x]$.

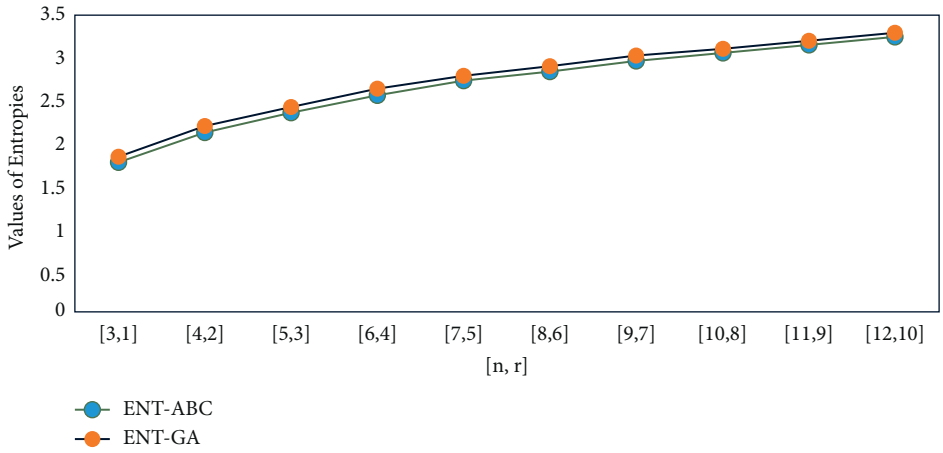


Figure 9: Comparison of Randic entropies for $\alpha = 1/2$ and for $\alpha = -1/2$ for the coronoid system $C_2[u, v, w, x]$.

7. Conclusion

In this research, we have appropriated some degree-based indices for the characterization of the unambiguous graph-theoretical system of biochemical concern. We talked about topological indices, for instance, general Randic index, atomic bond connectivity index, and geometric arithmetic index for coronoid polycyclic aromatic hydrocarbons. Besides, we enumerated the corresponding entropies. The enumerated results link individual physico-chemical characteristics like distortion, stability, melting points, and strain energy of chemical compounds. The mathematical findings for these coronoid systems are helpful for the chemist to understand the biochemical applications of these coronoid systems $C_1[u, v, w, x]$ and $C_2[u, v, w, x]$.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Authors' Contributions

All authors contributed equally to this work.

References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[2] N. Rashevsky, "Life, information theory, and topology," *Bulletin of Mathematical Biophysics*, vol. 17, no. 3, pp. 229–235, 1955.

[3] R. E. Ulanowicz, "Quantitative methods for ecological network analysis," *Computational Biology and Chemistry*, vol. 28, no. 5-6, pp. 321–339, 2004.

[4] M. Dehmer, "Information processing in complex networks: graph entropy and information functionals," *Applied Mathematics and Computation*, vol. 201, no. 1–2, pp. 82–94, 2008.

[5] M. Dehmer and A. Mowshowitz, "A history of graph entropy measures," *Information Sciences*, vol. 181, no. 1, pp. 57–78, 2011.

[6] S. Manzoor, M. K. Siddiqui, and S. Ahmad, "On entropy measures of molecular graphs using topological indices," *Arabian Journal of Chemistry*, vol. 13, no. 8, pp. 6285–6298, 2020.

[7] Y. J. Tan and J. Wu, "Network structure entropy and its application to scale-free networks," *Systems Engineering-Theory&Practice*, vol. 6, pp. 1–3, 2004.

[8] M. Randić, "Aromaticity of polycyclic conjugated hydrocarbons," *Chemical Reviews*, vol. 103, no. 9, pp. 3449–3605, 2003.

[9] Z. Chen, M. Dehmer, and Y. Shi, "A note on distance-based graph entropies," *Entropy*, vol. 16, no. 10, pp. 5416–5427, 2014.

[10] D. Amić, D. Bešlo, B. Lucić, S. Nikolić, and N. Trinajstić, "The vertex-connectivity index revisited," *Journal of Chemical Information and Computer Sciences*, vol. 38, no. 5, pp. 819–822, 1998.

[11] S. Manzoor, M. K. Siddiqui, and S. Ahmad, "Degree-based entropy of molecular structure of hyaluronic acid–curcumin conjugates," *The European Physical Journal Plus*, vol. 136, no. 1, pp. 1–21, 2021.

[12] E. Estrada, L. Torres, L. Rodriguez, and I. Gutman, "An atom-bond connectivity index: modelling the enthalpy of formation of alkanes," *Indian Journal of Chemistry*, vol. 37A, pp. 849–855, 1998.

[13] D. Vukičević and B. Furtula, "Topological index based on the ratios of geometrical and arithmetical means of end-vertex degrees of edges," *Journal of Mathematical Chemistry*, vol. 46, no. 4, pp. 1369–1376, 2009.

[14] Y. M. Chu, M. K. Siddiqui, S. Javed, L. Sherin, and F. Kausar, "On zagreb type molecular descriptors of ceria oxide and their applications," *Journal of Cluster Science*, pp. 1–10, 2021.

[15] H. A. Staab and F. Diederich, "Cycloarenes, a new class of aromatic compounds, I. Synthesis of kekulene," *Chemische Berichte*, vol. 116, no. 10, pp. 3487–3503, 1983.

[16] H. Miyoshi, S. Nobusue, A. Shimizu, and Y. Tobe, "Non-alternant non-benzenoid kekulenes: the birth of a new kekulene family," *Chemical Society Reviews*, vol. 44, no. 18, pp. 6560–6577, 2015.

[17] D. J. H. Funhoff and H. A. Staab, "Cyclo[d.e.d.e.e.d.e.d.e.e.] decakisbenzene, a New Cycloarene," *Angewandte Chemie International Edition in English*, vol. 25, no. 8, pp. 742–744, 1986.

[18] B. Kumar, R. L. Viboh, M. C. Bonifacio et al., "Septulene: the heptagonal homologue of kekulene," *Angewandte Chemie*, vol. 124, no. 51, pp. 12967–12972, 2012.
[19] M. Randic, "Aromaticity and conjugation," *Journal of the American Chemical Society*, vol. 99, no. 2, pp. 444–450, 1977.
[20] J. Cioslowski, "Additive nodal increments for approximate calculation of the total ?-electron energy of benzenoid hydrocarbons," *Theoretica Chimica Acta*, vol. 68, no. 4, pp. 315–319, 1985.
[21] J. R. Dias, "Structure and electronic characteristics of coronoid polycyclic aromatic hydrocarbons as potential models of graphite layers with hole defects," *The Journal of Physical Chemistry A*, vol. 112, no. 47, pp. 12281–12292, 2008.
[22] J. R. Dias, "Handbook of polycyclic hydrocarbons," *Part A: Benzenoid Hydrocarbons*, John Wiley & Sons, New Jersey, USA, 1987.
[23] J.-i. Aihara, M. Makino, T. Ishida, and J. R. Dias, "Analytical study of superaromaticity in cycloarenes and related coronoid hydrocarbons," *The Journal of Physical Chemistry A*, vol. 117, no. 22, pp. 4688–4697, 2013.
[24] J. R. Dias, "On the spectacular structural isomorphism between *CnHs* monoradical and Cn + sHs + 3 diradical benzenoid hydrocarbons: from reactive intermediates to vacancy (hole) defects in graphite," *The Journal of Physical Chemistry A*, vol. 112, no. 14, pp. 3260–3274, 2008.
[25] F. Dietz, N. Tyutyulkov, G. Madjarova, and K. Müllen, "Is 2-D graphite an ultimate large hydrocarbon? II. Structure and energy spectra of polycyclic aromatic hydrocarbons with defects," *The Journal of Physical Chemistry B*, vol. 104, no. 8, pp. 1746–1761, 2000.
[26] A. T. Balaban, D. J. Klein, and X. Liu, "Graphitic cones," *Carbon*, vol. 32, no. 2, pp. 357–359, 1994.
[27] A. Giordana, A. Maranzana, G. Ghigo, M. Causà, and G. Tonachini, "Soot platelets and PAHs with an odd number of unsaturated carbon atoms and π electrons: theoretical study of their spin properties and interaction with ozone," *The Journal of Physical Chemistry A*, vol. 112, no. 5, pp. 973–982, 2008.
[28] G. G. Hall, "Molecules with holes," *Theoretica Chimica Acta*, vol. 73, no. 5-6, pp. 425–435, 1988.
[29] J. Wu, W. Pisula, and K. Müllen, "Graphenes as potential material for electronics," *Chemical Reviews*, vol. 107, no. 3, pp. 718–747, 2007.
[30] T. Parsons-Moss, S. Jones, J. Wang et al., "Reduction of plutonium in acidic solutions by mesoporous carbons," *Journal of Radioanalytical and Nuclear Chemistry*, vol. 307, no. 3, pp. 2593–2601, 2016.
[31] Z. Raza, "The expected values of arithmetic bond connectivity and geometric indices in random phenylene chains," *Heliyon*, vol. 6, no. 7, Article ID e04479, 59 pages, 2020.
[32] Z. Raza and M. E. K. Sukaiti, "M-polynomial and degree based topological indices of some nanostructures," *Symmetry*, vol. 12, no. 5, pp. 831–841, 2020.
[33] M. Arockiaraj, J. Clement, and K. Balasubramanian, "Topological indices and their applications to circumcised donut benzenoid systems, kekulenes and drugs," *Polycyclic Aromatic Compounds*, vol. 40, no. 2, pp. 280–303, 2020.
[34] K. Julietraja and P. Venugopal, "Computation of degree-based topological descriptors using M-polynomial for coronoid systems," *Polycyclic Aromatic Compounds*, vol. 1, pp. 1–24, 2020.
[35] K. Julietraja, P. Venugopal, S. Prabhu, J.-B. Liu, and M. K. Siddiqui, "M-polynomial and degree-based molecular descriptors of certain classes of benzenoid systems," *Polycyclic Aromatic Compounds*, vol. 2, pp. 1–30, 2021.
[36] K. Anand and G. Bianconi, "Entropy measures for networks: toward an information theory of complex topologies," *Physical Review E*, vol. 80, no. 4, pp. 045102–045155, 2009.
[37] G. Bianconi, "The entropy of randomized network ensembles," *Europhysics Letters*, vol. 81, no. 2, pp. 28–45, 2008.
[38] M. Cai, R. Y. K. Du Hai-Feng, and W. Marcus, "A new network structure entropy based node difference and edge difference," *Acta Physica Sinica*, vol. 60, pp. 110–115, 2011.
[39] Y.-H. Xiao, W.-T. Wu, H. Wang, M. Xiong, and W. Wang, "Symmetry-based structure entropy of complex networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 387, no. 11, pp. 2611–2619, 2008.

Hindawi
Complexity
Volume 2021, Article ID 8806874, 10 pages
https://doi.org/10.1155/2021/8806874

Research Article

Beamforming Design and Covert Performance Analysis for Full-Duplex Multiantenna System

Ling Yang,[1,2] **Weiwei Yang**,[1] **Liang Tang**,[3] **Jia Tu**,[4] **Xingbo Lu**,[1] **and Zhengyun He**[2]

[1]*College of Communications Engineering, PLA Army Engineering University, Nanjing 210007, China*
[2]*College of Railway Transportation, Hunan University of Technology, Zhuzhou 412007, China*
[3]*College of Science, Hunan University of Technology, Zhuzhou 412007, China*
[4]*College of International Studies, National University of Defense Technology, Nanjing 210039, China*

Correspondence should be addressed to Weiwei Yang; weiweiyang1981@126.com

Received 25 May 2021; Accepted 6 September 2021; Published 7 October 2021

Academic Editor: Hou-Sheng Su

In this work, a wireless covert communication system with full-duplex (FD) multiantenna receiver is considered. In order to improve the convert performance of the wireless communication system in the FD mode, a scheme based on selection combining/zero forcing beamforming (SC/ZFB) is proposed. More specifically, a covert message receiver with a FD multiantenna uses the zero forcing beamforming method to transmit randomly varying noise power to the adversary while receiving covert information from the sender. Firstly, we derive the optimal detection threshold and the corresponding closed expression of the minimum detection error rate of the warden. Secondly, the transmission interruption probability is explored to measure the communication reliability between the sender and the receiver of the covert message. Finally, the throughput performance of the covert communication system is analyzed under random geometry. Our analysis shows that the proposed SC/ZFB scheme can achieve the positive effective convert rate while interfering with the detection of the warden as much as possible. It is worth noting that the increase of the number of antennas and the power of covert message transmission can improve the convert performance of the system.

1. Introduction

The confidentiality of network activities and the integrity of data transmission are the basic goals of wireless network system construction. Although the network system structure [1, 2] and network control protocol [3, 4] design determine the efficiency of data transmission to a certain extent, due to the wireless network uses the public electromagnetic wave as an open medium to transmit data signals, the risk of data transmission will be greatly increased. Therefore, how to ensure the security and privacy of information in wireless networks has been a hot research issue. The common methods include traditional information theory technology [5–7], encryption technology [8, 9], and emerging covert communication technology [10, 11]. Traditional methods [12, 13] focus on protecting information content from eavesdropping, while new methods focus on protecting information transmission process. The covert communication technology (also known as low probability detection communication (LPD)) appeared in the form of spread spectrum technology in the earliest stage. In recent years, the researchers have studied covert communication under additive white Gaussian noise (AWGN) channel [14, 15], binary symmetric channel (BSC) [16], and discrete memoryless channel (DMC) [17], respectively. At the same time, the basic limitations of covert communication under these channels have also been studied in these literatures, especially in the pioneering work of covert communication [14]; a pessimistic conclusion is put forward that when $n \longrightarrow \infty$, there is $O(\sqrt{n})/n \longrightarrow 0$, which means that when n is large enough, the covert rate is 0. Therefore, more and more people are devoting themselves to exploring the conditions that can ensure the positive covert rate and putting forward some new research methods; the main

purpose is to create conditions for the warden to detect errors. The basic strategy is adopted to disturb the warden detection by uncertainties of noise [18, 19], channel [20], interference power [19, 21–23], and communication time [24]. In addition, covert communication schemes with relays [25] have also been proposed successively, and the literature [26] studies covert communication with finite block length in AWGN channels. Yan et al. [27] study delay-intolerant covert communications in AWGN channels with finite block length. Recently, some researchers consider covert communication in the network environment of unmanned aerial vehicle (UAV) [28]. However, the existing covert communication systems involve multiple antennas in addition to [29–31]; most of the existing research on covert communication focuses on single-antenna scenarios, while multi-antenna covert communication scenarios [32] have not been well developed.

Full-duplex (FD) technology in wireless communication scenarios has been effectively explored in [22, 33, 34]. Shu et al. [33] send artificial interference to the adversary through the FD receiver, so as to enhance the convert performance. Wang et al. [34] use channel uncertainty to explore covert communication in relay networks operating in the FD mode. Similarly, Shahzad et al. [22] also use FD technology; when the receiver of the covert message receives the signal, it will generate artificial noise with different power, resulting in uncertainty to the opponent.

As for beamforming technology, it is rarely used. In [35], the enhancement effect of multiantenna AN nodes on covert performance is studied, and it is pointed out that directional beamforming is the optimal AN strategy. In fact, how to use multiantenna technology to improve spectrum efficiency so as to bring more security gains in covert communication systems and how to design more appropriate multiantenna scenarios to make effective covert rate higher are necessary to be considered. Therefore, the potential of multiantenna technology for covert communication in fading channels remains to be tapped, which also stimulates the current research work.

In this work, we consider covert communication in the context of multiantenna receiver networks. Specifically, the full-duplex and multiantenna receiver generates artificial noise (AN) with a randomized transmit power to deliberately confusion and affect the detection of Willie. The covert transmission selection combining/zero forcing beamforming (SC/ZFB) schemes are proposed: the destination selects the best antenna to receive the information from the source and uses all the remaining antennas to send the jamming signal to warden according to the principle of ZFB. The main contributions of this paper are summarized as follows:

(i) We demonstrate the possibility that the proposed SC/ZFB scheme can achieve covert performance in wireless communication networks; the results show that the SC/ZFB schemes can achieve positive effective covert rate.

(ii) The closed-form expressions of the minimum detection error rate and transmission outage probability are derived, respectively, and the covertness and reliability of the system are described, respectively, on this basis.

(iii) Our analysis shows that increasing the number of antennas can improve the covert performance. In other words, with the increase of the number of antennas, the detection performance and covert performance of the covert communication system will also increase.

The remainder of this paper is organized as follows. The system model is introduced in Section 2. Sections 3 and 4 analyze the detection performance and the covert throughput for the SC/ZFB schemes, respectively. The theoretical analyzes are verified by numerical results in Section 5. Finally, Section 6 concludes the paper and summarizes the key findings.

Notation: throughout this paper, the boldface uppercase letters are used to denote matrices. † is denoted as the conjugate transpose operation. $\|\cdot\|$ is defined as the Frobenius norm. $F(\cdot)$ and $f(\cdot)$ represent the cumulative distribution function (CDF) and the probability density function (PDF) of random variable, respectively. $E(\cdot)$ denotes the expectation operation. A list of the fundamental variables is provided in Table 1.

2. System Model

2.1. Considered Scenario and Adopted Assumptions. As shown in Figure 1, considering a multiantenna wireless network consisting of a source Alice, a destination Bob, and a warden Willie, Warden Willie seeks to detect any transmission by Alice, and Alice desires to deliver messages to Bob reliably while guarantees a low probability of being detected by Willie; at the same time, they know each other's existence and location. Bob is equipped with N_B antennas, and all the other nodes, including Alice and Willie, are single-antenna devices. In addition to receiving covert messages, Bob uses additional antenna for transmission of AN in order to confuse Willie, and FD operations are assumed at Bob.

We assume that the wireless channels are subject to independent quasi-static Rayleigh fading, where the channel coefficient remains constant over one communication slot of n channel uses and changes independently from slot to slot. The channels from Alice-to-Bob, Alice-to-Willie, and Bob-to-Willie, are denoted by $\mathbf{h}_{AB}$, h_{AW}, and $\mathbf{h}_{BW}$, respectively, while the self-interference channel of Bob is denoted by $\mathbf{h}_{BB}$. We assume that Bob knows $\mathbf{h}_{AB}$, while Willie knows h_{AW} and $\mathbf{h}_{BW}$. Considering Bob knows about Willie, so it is assumed that Bob also knows $\mathbf{h}_{BW}$.

Based on FD operation, Bob first selects the best antenna based on the channel state information (CSI) of the channel between Alice and Bob to receive the covert messages and utilizes the remaining $N_B - 1$ antennas to send a weighted AN signal to disturb Willie's detection. We adopt the ZFB algorithm to avoid the undesirable AN signals at Alice. Thus, the optimal weight vector $\mathbf{W}_{ZF}$ is the solution of the following optimization problem:

Table 1: List of fundamental variables.

Symbol	Description
N_B	Antenna number of Bob
$h_{a,b}$	The channel between a and b who both consist a single antenna
$\mathbf{h}_{a,b}$	The channel vector between a and b consisting one and multiple antennas, respectively
$\mathbf{w}_{a,b}$	The received weighting vector on link $a-b$
$T^{\perp}$	The projection idempotent matrix with rank N_B-2
B_j	The jth antenna of Bob
P_z	The AN power of Bob
P_X	The power send by X, $X \in$ {Alice, Bob, Willie}
σ_A^2	The variance of AWGN at node A, $A \in$ {Bob, Willie}
γ_A	The SNR on A, $A \in$ {Bob, Willie}
y_A	The received signal at A, $A \in$ {Bob, Willie}
n	The number of each channel use
$n_B[i]$	The complex additive Gaussian noise at A, $A \in$ {Bob, Willie}
T_W	The average power
τ	The predefined detection threshold
ξ	The detection error rate
ξ^*	The minimum detection error rate
P_{FA}	The false alarm rate
P_{MD}	The miss detection rate
R	The transmission rate
δ	The transmission outage probability
R_C	The effective covert rate
$\overline{R_C}$	The maximizing effective covert rate

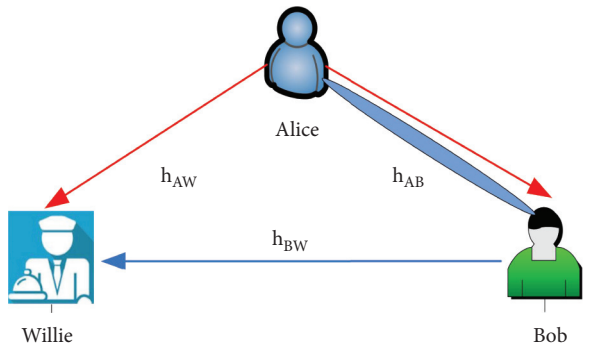


Figure 1: System model.

$$\max_{\mathbf{W}_{ZF}} \left|\mathbf{h}_{BW}^{\dagger}\mathbf{W}_{ZF}\right| \quad \text{s.t.} \quad \left|\mathbf{h}_{BA}^{\dagger}\mathbf{W}_{ZF}\right| = 0 \& \left\|\mathbf{W}_{ZF}\right\|_F = 1, \tag{1}$$

where † is the conjugate transpose operator, $\|\cdot\|_F$ denotes the Frobenius norm, and $\mathbf{h}_{BW}$ and $\mathbf{h}_{BA}$ denote the $(N_B-1)\times 1$ channel vectors between the remaining N_B-1 antennas of the Bob and the Willie and the remaining N_B-1 antennas of the Bob and the Alice, respectively. According to the theorems in [36], the solution of the optimal weight vector in formula (1) can be obtained as

$$\mathbf{W}_{ZF} = \frac{T^{\perp}\mathbf{h}_{BW}}{\left\|T^{\perp}\mathbf{h}_{BW}\right\|}, \tag{2}$$

where $T^{\perp} = (\mathbf{I} - \mathbf{h}_{BA}(\mathbf{h}_{BA}^{\dagger}\mathbf{h}_{BA})^{-1}\mathbf{h}_{BA}^{\dagger})$ is the projection idempotent matrix with rank N_B-2.

Let us define $Z_1 \triangleq P_Z|\mathbf{h}_{BW}^{\dagger}\mathbf{W}_{ZF}|^2/\sigma_W^2$ and $Z \triangleq P_Z|\mathbf{h}_{BW}^{\dagger}\mathbf{W}_{ZF}|^2$; the corresponding probability density function (PDF) and the cumulative distribution function (CDF) can be expressed as

$$f_{Z_1}(Z) = \frac{Z^{N_B-3}\exp\left(-\left(\sigma_W^2 Z/P_Z\right)\right)}{(N_B-3)!\left(P_Z/\sigma_W^2\right)^{N_B-2}}, \quad N_B \geq 3, Z \geq 0,$$
$$F_{Z_1}(Z) = 1 - \exp\left(-\frac{Z\sigma_W^2}{P_Z}\right)\sum_{l=0}^{N_B-3}\frac{1}{l!}\left(\frac{Z\sigma_W^2}{P_Z}\right)^{l}, \quad Z \geq 0, \tag{3}$$

where P_Z denotes the power of the AN signal from Bob and the σ_W^2 is noise variance of Willie.

The instantaneous SINR at Bob is given by

$$\gamma_{\mathrm{Bob}} = \frac{P_A \max_{1\leq i\leq N_B}\left(\left|h_{AB_i}\right|^2\right)}{\sigma_B^2}. \tag{4}$$

When Alice transmits, the received signal at Willie is given by

$$y_{\mathrm{Bob}}[i] = \sqrt{P_A}\max_{1\leq i\leq N_B}\left(\left|h_{AB_i}\right|^2\right)x_A[i] + n_B[i], \tag{5}$$

where h_{AB_i} denotes the channel coefficient between Alice and the ith antenna of Bob, $x_A[i]$ is the transmitted signal by Alice satisfying $E[x_A[i]x_A^{\dagger}[i]] = 1$, i is the index of each channel use, $i = 1, 2, \ldots, n$, and $n_B[i]$ is the complex additive Gaussian noise at Willie with σ_B^2 as its variance, i.e., $n_B[i] \sim \mathrm{CN}(0, \sigma_B^2)$.

2.2. Detection Metrics at Willie. Warden Willie attempts to judge whether Alice is transmitting to Bob in a communication slot. To this end, Willies should distinguish two hypotheses, namely, the null hypothesis H_0 meaning that

Alice is not transmitting covert information and the alternate hypothesis H_1 indicating an ongoing communication. The two hypotheses are detailed as below:

$$y_{\text{willie}}[i] = \begin{cases} \sqrt{P_Z}\,|h_{BW}^{\dagger} W_{ZF}|V_B[i] + n_W[i], & H_0, \\ \sqrt{P_A}\,h_{AW}x_A[i] + \sqrt{P_Z}\,|\mathbf{h}_{BW}^{\dagger}\mathbf{W}_{ZF}|V_B[i] + n_W[i], & H_1, \end{cases} \tag{6}$$

where $n_W[i]$ is the AWGN at Willie with σ_W^2 as its variance, i.e., $n_W[i] \sim \text{CN}(0, \sigma_W^2)$. Willie does not know P_{B_j} and P_Z in a communication slot, while P_A is fixed and known.

We focus on a communication slot, where Willie has to decide whether Alice transmitted covert information to Bob, or not. By applying Neyman-Pearson criterion [37] and likelihood ratio test, Willie can obtain the optimal decision rule to minimize his detection error; the rule is given by

$$T_W \underset{D_0}{\overset{D_1}{\lessgtr}} \tau, \tag{7}$$

where $T_W = 1/n \sum_{i=1}^{n} y_{\text{Willie}}[i]$ is the average power received at Willie in the slot, τ is a predefined detection threshold, D_1 and D_0 are the binary decisions that infer whether the Alice transmits covert message to Bob or not, respectively, and decision D_0 is made if $T_W < \tau$, whereas decision D_1 is made otherwise. We consider an infinite number of channel uses, i.e., $n \longrightarrow \infty$.

The average received power is given by

$$T_W = \begin{cases} P_Z\left|h_{BW}^{\dagger} W_{ZF}\right|^2 + \sigma_W^2, & H_0, \\ P_A\left|h_{AW}\right|^2 + P_Z\left|\mathbf{h}_{BW}^{\dagger}\mathbf{W}_{ZF}\right|^2 + \sigma_W^2, & H_1. \end{cases} \tag{8}$$

At the end of the communication slot, Willie must make a decision. We define the false alarm rate as $P_{\text{FA}} = P(D_1|H_0)$, and the miss detection rate is defined as $P_{\text{MD}} = P(D_0|H_1)$. Under the assumption of equal a priori probabilities of hypotheses H_0 and H_1, the detection performance of Willie is normally measured by the detection error probability, which is defined as $\xi \triangleq P_{\text{FA}} + P_{\text{MD}}$.

3. Detection Performance at Willie

In this section, we derive Willie false alarm and miss detection rates and examine how Willie optimally sets the value of τ aiming to minimize the detection error rate ξ.

3.1. False Alarm and Miss Detection Rates. Considering Rayleigh fading, the cumulative distribution function (CDF) of the channel coefficient $|h_{B_jW}|^2$ between the jth antenna of Bob and the Willie is given by $F_{|h_{B_jW}|^2}(x) = 1 - \exp(-x)$.

Theorem 1. *For SC/ZBF scheme, the false alarm and miss detection rates at Willie are derived as (10) and (11), shown at the top of the next page, where*

$$k_1 = \frac{\tau - \sigma_w^2}{(N_B - 1)P_{B_j}^{\max}}, \qquad k_2 = \frac{\tau - \sigma_w^2 - P_A\left|h_{AW}\right|^2}{(N_B - 1)P_{B_j}^{\max}}, \tag{9}$$

$$P_{\text{FA}} = \begin{cases} 1, & \tau \le \sigma_W^2, \\ xEi(-k_1) + \exp(-k_1), & \tau > \sigma_W^2, N_B = 3, \\ \exp(-k_1), & \tau > \sigma_W^2, N_B = 4, \\ \left\{1 + \sum_{l=2}^{N_B-3}\left[\frac{1}{l(l-1)}\sum_{k=0}^{l-2}\frac{(k_1)^{k+1}}{k!}\right]\right\}\exp(-k_1), & \tau > \sigma_W^2, N_B \ge 5, \end{cases} \tag{10}$$

$$P_{\text{MD}} = \begin{cases} 0, & \tau \le \sigma_W^2 + P_A\left|h_{AW}\right|^2, \\ 1 - [k_2 Ei(-k_2) + \exp(-k_2)], & \tau > \sigma_W^2 + P_A\left|h_{AW}\right|^2, N_B = 3, \\ 1 - \exp(-k_2), & \tau > \sigma_W^2 + P_A\left|h_{AW}\right|^2, N_B = 4, \\ 1 - \left\{1 + \sum_{l=2}^{N_B-3}\left[\frac{1}{l(l-1)}\sum_{k=0}^{l-2}\frac{(k_2)^{k+1}}{k!}\right]\right\}\exp(-k_2), & \tau > \sigma_W^2 + P_A\left|h_{AW}\right|^2, N_B \ge 5. \end{cases} \tag{11}$$

Proof. According to the optimal decision rule and (8), the false alarm rate is given by

$$P_{\mathrm{FA}} = P\left[P_Z\left|\mathbf{h}_{BW}^{\dagger}\mathbf{W}_{ZF}\right|^2 + \sigma_W^2 > \tau\right] = \begin{cases} 1, & \tau < \sigma_W^2, \\ 1 - P\left[\dfrac{P_Z\left|\mathbf{h}_{BW}^{\dagger}\mathbf{W}_{ZF}\right|^2}{\sigma_W^2} < \dfrac{\tau - \sigma_W^2}{\sigma_W^2}\right], & \tau \geq \sigma_W^2. \end{cases} \tag{12}$$

From (10), the P_{FA} can be obtained as

$$P_{\mathrm{FA}} = \begin{cases} 1, & \tau < \sigma_W^2, \\ 1 - \displaystyle\int_0^{(N_B-1)P_{B_j}^{\max}} \left[1 - \exp\left(-\frac{\tau - \sigma_w^2}{P_Z}\right)\sum_{l=0}^{N_B-3}\frac{1}{l!}\left(\frac{\tau - \sigma_w^2}{P_Z}\right)^l\right] \times \frac{1}{(N_B-1)P_{B_j}^{\max}}\,\mathrm{d}P_Z & \tau \geq \sigma_W^2. \end{cases} \tag{13}$$

Then, we solve the integral in (13) with the aid of equation (3.351. 2.11) and equation (3.351.4) in [34], utilizing the variable substitution (i.e., setting $t = 1/P_Z$). Similarly, the miss detection rate is given by

$$P_{\mathrm{MD}} = P\left[P_A\left|h_{AW}\right|^2 + P_Z\left|\mathbf{h}_{BW}^{\dagger}\mathbf{W}_{ZF}\right|^2 + \sigma_W^2 < \tau\right] = \begin{cases} 0, & \tau < \sigma_W^2 + P_A\left|h_{AW}\right|^2, \\ \displaystyle\int_0^{(N_B-1)P_{B_j}^{\max}} \left[1 - \exp\left(-\frac{\tau - \sigma_w^2 - P_A\left|h_{AW}\right|^2}{P_Z}\right) \times \sum_{l=0}^{N_B-3}\frac{1}{l!}\left(\frac{\tau - \sigma_w^2 - P_A\left|h_{AW}\right|^2}{P_Z}\right)^l\right] \times \frac{1}{(N_B-1)P_{B_j}^{\max}}\,\mathrm{d}P_Z & \tau \geq \sigma_W^2 + P_A\left|h_{AW}\right|^2. \end{cases} \tag{14}$$

3.2. The Optimization of Detection Threshold and Detection Error Rate at Willie

Theorem 2. *For the SC/ZBF scheme, under the assumption of a radiometer at Willie, the optimal threshold that minimizes ξ is derived as*

$$\tau^* = P_A\left|h_{AW}\right|^2 + \sigma_W^2, \tag{15}$$

and the corresponding detection error rate is derived as

$$\xi^* = \begin{cases} x^* Ei\left(-x^*\right) + \exp\left(-x^*\right), & N_B = 3, \\ \exp\left(-x^*\right), & N_B = 4, \\ \left\{1 + \displaystyle\sum_{l=2}^{N_B-3}\left[\frac{1}{l(l-1)}\sum_{k=0}^{l-2}\frac{(x^*)^{k+1}}{k!}\right]\right\}\exp\left(-x^*\right), & N_B \geq 5, \end{cases} \tag{16}$$

$$\xi = \begin{cases} 1, & \tau \le \sigma_W^2, \\ k_1 Ei(-k_1) + \exp(-k_1), & \sigma_W^2 < \tau \le \sigma_W^2 + P_A|h_{AW}|^2, N_B = 3, \\ \exp(-k_1), & \sigma_W^2 < \tau \le \sigma_W^2 + P_A|h_{AW}|^2, N_B = 4, \\ \left\{1 + \sum_{l=2}^{N_B-3}\left[\frac{1}{l(l-1)}\sum_{k=0}^{l-2}\frac{(k_1)^{k+1}}{k!}\right]\right\}\exp(-k_1), & \sigma_W^2 < \tau \le \sigma_W^2 + P_A|h_{AW}|^2, N_B \ge 5, \\ 1 + k_1 Ei(-k_1) + \exp(-k_1) - k_2 Ei(-k_2) - \exp(-k_2), & \tau > \sigma_W^2 + P_A|h_{AW}|^2, N_B = 3, \\ 1 + \exp(-k_2) - \exp(-k_2), & \tau > \sigma_W^2 + P_A|h_{AW}|^2, N_B = 4, \\ 1 + \left\{1 + \sum_{l=2}^{N_B-3}\left[\frac{1}{l(l-1)}\sum_{k=0}^{l-2}\frac{(k_1)^{k+1}}{k!}\right]\right\}\exp(-k_1) - \left\{1 + \sum_{l=2}^{N_B-3}\left[\frac{1}{l(l-1)}\sum_{k=0}^{l-2}\frac{(k_2)^{k+1}}{k!}\right]\right\}\exp(-k_2), & \tau > \sigma_W^2 + P_A|h_{AW}|^2, N_B \ge 5, \end{cases} \tag{17}$$

where

$$x^* = \frac{P_A|h_{AW}|^2}{(N_B - 1)P_{B_j}^{\max}}. \tag{18}$$

Proof. Following (12) and (13), the detection error rate can be expressed as (17), shown at the top of the next page.

For $N_B = 3$, defining $f(x) = xEi(-x) + \exp(-x)$, we derive the first derivative of $f(x)$ as $f'(x) = Ei(-x) < 0$. Similarly, we can derive the second derivative as $f''(x) = e^{-x}/x > 0, x > 0$. When $\sigma_W^2 < \tau \le \sigma_W^2 + P_{B_k}|h_{B_kW}|^2$, the first derivative of ξ with respect to τ can be derived as

$$\frac{d\xi}{d\tau} = \frac{1}{(N_B - 1)P_{B_j}^{\max}}\frac{df(k_1)}{dk_1} < 0. \tag{19}$$

When $\tau > \sigma_W^2 + P_{B_k}|h_{B_kW}|^2$, the second derivative of ξ with respect to τ can be derived as

$$\frac{d\xi}{d\tau} = \frac{1}{(N_B - 1)P_{B_j}^{\max}}\left[\frac{df(k_1)}{dk_1} - \frac{df(k_2)}{dk_2}\right] > 0. \tag{20}$$

According to $k_1 > k_2$ and the monotonicity of function $f(x)$, the minimum detection error rate ξ^* at Willie is obtained at $\tau^* = \sigma_W^2 + P_A|h_{AW}|^2$.

Using the same method, we can obtained that the minimum detection threshold and the corresponding minimum detection error rate for $N_B = 4$ and $N_B \ge 5$ are the same as those for $N_B = 3$. Hence, the optimal threshold that minimizes ξ is $\tau^* = P_A|h_{AW}|^2 + \sigma_W^2$. Substituting into (17), the desired ξ^* can be obtained. □

Corollary 1. *For the SC/ZBF scheme, the minimum detection error rate ξ^* increases with the increase of $P_{B_j}^{\max}$ and N_B.*

Proof. From (16), for $N_B \ge 3$, ξ^* is monotonic decreasing function of x^* and x^* is monotonic decreasing function of $P_{B_j}^{\max}$ and N_B, so $P_{B_j}^{\max} \longrightarrow \infty$ or $N_B \longrightarrow \infty$, $\xi^* \longrightarrow 1$. □

Remark 1. The minimum detection error rate ξ^* is a monotonic decreasing function of P_A. This is because of ξ^* is monotonic decreasing function of x^*, and x^* is monotonic increasing function of P_A, so $P_A \longrightarrow \infty$ and $\xi^* \longrightarrow 0$.

4. Performance of Covert Communication

For the SC/ZBF scheme, the maximizing the effective covert rate subject to the constraint can be expressed as

$$\begin{aligned} &\max \overline{R_C} \\ &\text{s.t.} \quad \xi^* \ge 1 - \varepsilon, \end{aligned} \tag{21}$$

where $\overline{R_C} \triangleq R(1 - \delta)$.

4.1. Transmission Outage Probability from Alice to Bob

Theorem 3. *A closed-form expression for the exact transmission outage probability δ of SC/ZFB is provided below:*

$$\delta = \left[1 - \exp\left(-\frac{(2^R - 1)\sigma_B^2}{P_A}\right)\right]^{N_B}. \tag{22}$$

Proof. If the transmission outage probability from Alice to Bob incurs, following (4), the probability can be represented as

$$\delta = P\left\{\frac{P_A \max_{1 \le i \le N_B}\left|h_{AB_i}\right|^2}{\sigma_B^2} < 2^R - 1\right\} = P\left\{\max_{1 \le i \le N_B}\left|h_{AB_i}\right|^2 < \frac{(2^R - 1)\sigma_B^2}{P_A}\right\}. \tag{23}$$

To this end, substituting the PDF of $\max_{1\le i\le N_B}(|h_{AB_i}|^2)$ into (23), the desired result can be derived. □

Corollary 2. *For the SC/ZBF scheme, the transmission outage probability δ decreases with the increase of P_A and the increase of N_B and increases with the increase of σ_B^2 and R.*

Proof. If P_A gets bigger, $1-\exp(-((2^R-1)\sigma_B^2)/P_A)$ gets smaller and $1-\exp(-((2^R-1)\sigma_B^2)/P_A)\ge 0$, so δ gets smaller. Similarly, if σ_B^2 gets bigger, δ gets smaller as R gets bigger.

Due to $1-\exp(-((2^R-1)\sigma_B^2)/P_A)<1$, δ gets smaller as N_B gets bigger. □

4.2. Optimal Transmission Power of Covert Information

Theorem 4. *For a given transmission rate R and the covertness constraint ε, the optimal value for Alice transmission power of SC/ZFB is given by*

$$P_A^* = \begin{cases} \dfrac{(N_B-1)P_{B_j}^{\max} t_{\varepsilon_1}}{|h_{AW}|^2}, & N_B = 3, \\ \dfrac{(N_B-1)P_{B_j}^{\max}}{|h_{AW}|^2}\ln\left(\dfrac{1}{1-\varepsilon}\right), & N_B = 4, \\ \dfrac{(N_B-1)P_{B_j}^{\max} t_{\varepsilon_2}}{|h_{AW}|^2}, & N_B \ge 5, \end{cases} \tag{24}$$

where t_{ε_1} is the solution of $xEi(-x)+\exp(-x)=1-\varepsilon$ and t_{ε_2} is the solution of $\{1+\sum_{l=2}^{N_B-3}[(1/l(l-1))\sum_{k=0}^{l-2}(x^{k+1}/k!)]\}\exp(-x)=1-\varepsilon$.

Proof. According to the minimum detection error rate and its proof process, it is easy to be obtained; for $N_B>3$, ξ^* is monotonically decreasing with respect to P_A. Therefore, the optimal Alice transmission power is determined by $\xi^*=1-\varepsilon$, and (24) can be obtained by jointing (16). □

4.3. Maximum Effective Covert Rate

Theorem 5. *For the SC/ZFB scheme, the achievable maximum effective covert rate $\overline{R_C^*}$ is given by*

$$\overline{R_C}^* = R - R\left\{1-\exp\left[-\frac{(2^R-1)\sigma_B^2}{P_A^*}\right]\right\}^{N_B}. \tag{25}$$

Proof. From (21), we have

$$\overline{R_C} = R(1-\delta) = R - R\left\{1-\exp\left[-\frac{(2^R-1)\sigma_B^2}{P_A}\right]\right\}^{N_B}. \tag{26}$$

Because δ decreases with the increase of P_A, so $\overline{R_C}$ gets bigger, that is, $\overline{R_C}$ is monotonic increasing function with respect to P_A. Therefore, the maximum value of P_A can maximize the value of $\overline{R_C}$, and $\xi^*\ge 1-\varepsilon$ should be satisfied. For $N_B>3$, ξ^* decreases with the increase of P_A, so the maximum value of P_A is determined by $\xi^*\ge 1-\varepsilon$, i.e., P_A^*. Substituting (24) into (26), (25) can be derived. □

Corollary 3. *For the SC/ZBF scheme, the maximum effective covert rate $\overline{R_C^*}$ increases with the increase of $P_{B_j}^{\max}$, P_A^*, N_B, and ε.*

Proof. From (24), for $N_B\ge 3$, P_A^* is the monotonic increasing function of $P_{B_j}^{\max}$, and from (25), $\overline{R_C^*}$ is the monotonic increasing function of P_A^*, so $\overline{R_C^*}$ is the monotonic increasing function of $P_{B_j}^{\max}$. Other contents can be proved similarly. □

Corollary 4. *For the SC/ZFB scheme, when the transmit power P_A at Alice increases, the maximum effective covert rate $\overline{R_{C_1}^*}$ approaches a fixed value R:*

$$\lim_{P_A^*\longrightarrow\infty}\overline{R_{C_2}^*} = R. \tag{27}$$

Proof. Following (25), $\exp\{-[(2^R-1)\sigma_B^2]/P_A^*\}\longrightarrow 1$ as $P_A\longrightarrow\infty$, and then, $\{1-\exp[-(2^R-1)\sigma_B^2/P_A^*]\}^{N_B}\longrightarrow 0$, so $\overline{R_C^*}\longrightarrow R$. □

Remark 2. The maximum effective covert rate $\overline{R_C^*}$ is the monotonic decreasing function of the channel noise σ_B^2. That is, the larger σ_B^2 is, the smaller $\overline{R_C^*}$ is.

5. Numerical Result

In this section, we present representative numerical results to verify the detection performance of Willie and the covert communication performance of the considered system. In our simulations, because the channel coefficients change independently from slot to slot, their values within a time slot remains constant; we set all the channel coefficients to 0 dB. Except for the necessary comparisons, channel noise values are also 0 dB, and set the number of antennas N_B to 3 and the transmission rate R to 1.5, respectively.

Figure 2 shows the detection error rate ξ^* at Willie versus $P_{B_j}^{\max}$ for different values of P_A and N_B for the SC/ZFB scheme. In this figure, we first observe that ξ^* is a monotonically increasing function of P_A, and it is easy to understand that ξ^* increases with N_B. This is because the increase of interference power and antenna number of single antenna makes the total interference of Willie transmission larger. The total interference directly hinders the detection of Willie, resulting in the higher error rate of Willie. We also observe that ξ^* is a monotonically decreasing function of P_A, since a higher transmission power used by Alice increases the probability of being detected by Willie. This observation is also consistent with our earlier comments in Corollary 1.

Figure 3 depicts the transmission outage probability δ with different P_A for the SC/ZFB scheme, δ decreases with the increase of P_A. This is because an increase in P_A will increase the SINR at Bob, making transmission more secure

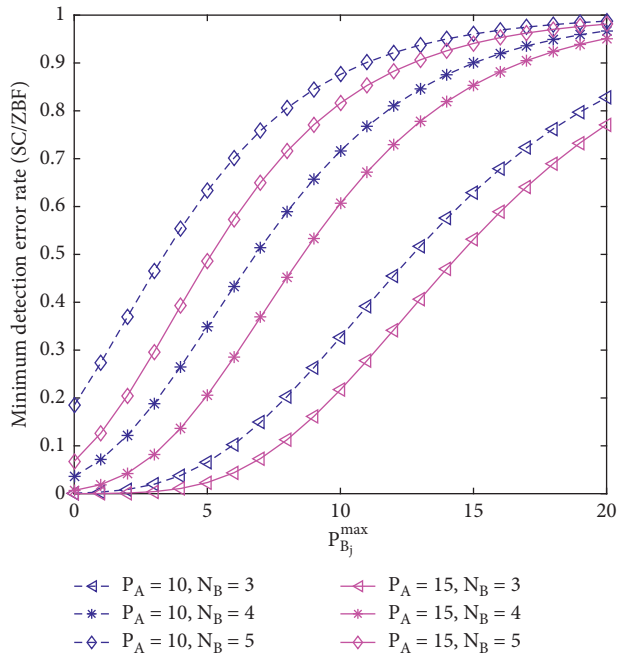


Figure 2: ξ^* versus $P_{B_j}^{\max}$ with different values of P_A and N_B.

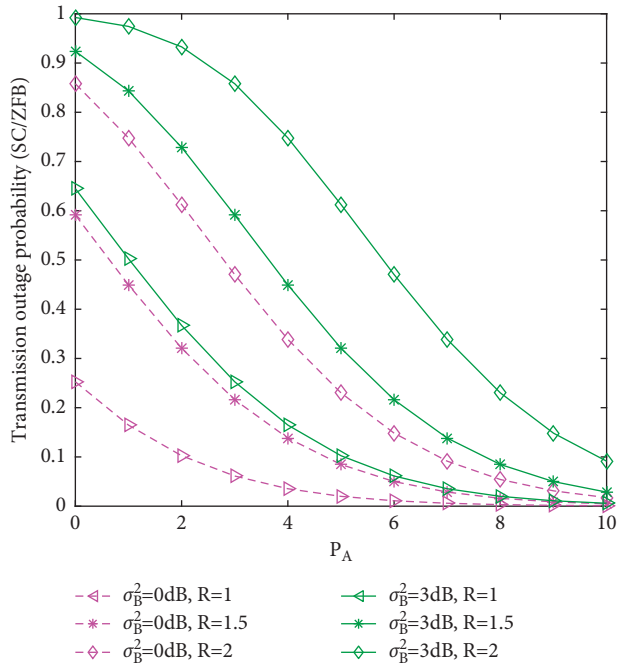


Figure 3: δ versus P_A with different values of σ_B^2 and R.

and less likely to interrupt. The relationship between δ and Bob's channel noise σ_B^2 is also described in Figure 3. δ increases with the increase of σ_B^2. This is because the SINR at Bob decreases with the increase of σ_B^2, which makes the transmission easy to interrupt. It can also be observed from Figure 3 that the higher the transmission rate, the easier the transmission to interrupt. These observations are consistent with those in Corollary 2.

It can be seen from Figure 4 that the maximum effective covert rate $\overline{R_C^*}$ increases with the increase of interference power $P_{B_j}^{\max}$ transmitted by a single antenna based on the SC/ZFB scheme. This is because the greater the interference power transmitted, the more error-prone Willie detection, which makes the effective covert rate greater. $\overline{R_C^*}$ increases with the increase of N_B and ε in SC/ZFB schemes. These are consistent with Corollary 3. In

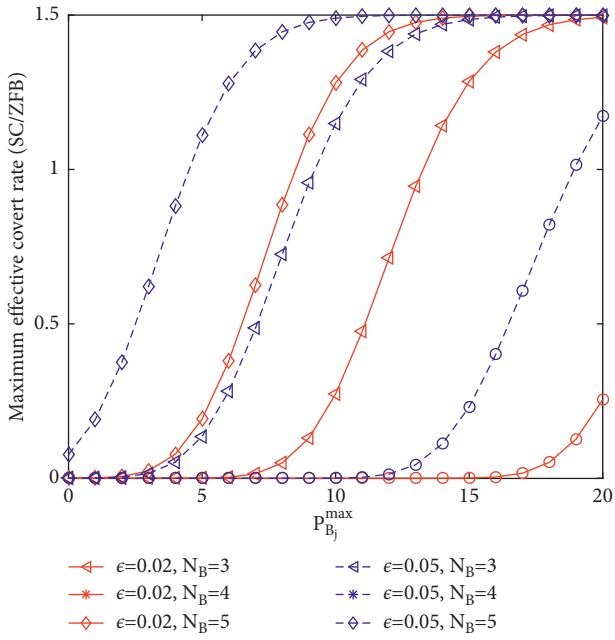


FIGURE 4: $\overline{R_C^*}$ versus $P_{B_j}^{\max}$ with different values of ε and N_B.

addition, Figure 4 shows that the maximum effective covert rate tends to a fixed value R, which verifies the correctness of Corollary 4.

6. Conclusion

In this paper, we have investigated the wireless covert communication network with the FD multiantenna receiver to achieve the convert performance by the generated randomly AN power of the selected antenna, and the analysis process of the covert performance was based on the detection performance of the warden. Our examinations showed that a higher positive effective covert rate can be ensured by controlling the number of antennas, and the AN power of the antenna is selected randomly by Bob. It can be expected that more multiantenna covert transmission schemes will be devised to achieve better covert performance in the future.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China, under Grant no. 62071486, the National Natural Science Foundation of Hunan Province, under Grant no. 2021JJ50051, and the Scientific Research Project of Education Department Hunan Province, China (nos. 19C0561 and 18C0531).

References

[1] Y. Liu and H. Su, "Some necessary and sufficient conditions for containment of second-order multi-agent systems with sampled position data," *Neurocomputing*, vol. 378, no. 32, pp. 228–237, 2020.

[2] C. Xu, H. Xu, H. Su, and C. Liu, "Disturbance-observer based consensus of linear multi-agent systems with exogenous disturbance under intermittent communication," *Neurocomputing*, vol. 404, no. 3, pp. 26–33, 2020.

[3] Y. Liu and H. Su, "Containment control of second-order multi-agent systems via intermittent sampled position data communication," *Applied Mathematics and Computation*, vol. 362, no. 362, Article ID 124522, 2019.

[4] C. Xu, B. Li, and L. Yang, "Semi-global containment of discrete-time high-order multi-agent systems with input saturation via intermittent control," *IET Control Theory & Applications*, vol. 14, no. 16, pp. 2303–2309, 2020.

[5] X. Sun, W. Yang, Y. Cai, Z. Xiang, and X. Tang, "Secure transmissions in millimeter wave SWIPT UAV-based relay networks," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 785–788, 2019.

[6] X. Lu, W. Yang, X. Guan, and Y. Cai, "DCE-based secure transmission for massive MIMO relay system against active eavesdropper," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 11, pp. 13045–13059, 2020.

[7] Z. Xiang, W. Yang, Y. Cai, Z. Ding, and Y. Song, "Secure transmission design in HARQ assisted cognitive NOMA networks," *IEEE Transactions on Information Forensics and Security*, vol. 15, no. 20, pp. 2528–2541, 2020.

[8] J. H. Jinsu Kim and J. Kim, "A hybrid scheme of public-key encryption and somewhat homomorphic encryption," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 5, pp. 1052–1063, 2015.

[9] J. Talbot and D. Welsh, *Complexity and Cryptography: An Introduction*, Cambridge University Press, Cambridge, MA, USA, 2006.

[10] Y. Wang, S. Yan, W. Yang, and Y. Cai, "Covert communications with constrained age of information," *IEEE Wireless Communications Letters*, vol. 10, no. 2, pp. 368–372, 2021.

[11] L. Tao, W. Yang, S. Yan, D. Wu, X. Guan, and D. Chen, "Covert communication in downlink NOMA systems with random transmit power," *IEEE Wireless Communications Letters*, vol. 9, no. 11, pp. 2000–2004, 2020.

[12] Z. Xiang, W. Yang, G. Pan, Y. Cai, Y. Song, and Y. Zou, "Secure transmission in HARQ-assisted non-orthogonal multiple access networks," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 2171–2182, 2020.

[13] X. Lu, W. Yang, X. Guan, Q. Wu, and Y. Cai, "Robust and secure beamforming for intelligent reflecting surface aided mmWave MISO systems," *IEEE Wireless Communications Letters*, vol. 9, no. 12, pp. 2068–2072, 2020.

[14] B. A. Bash, D. Goeckel, and D. Towsley, "Limits of reliable communication with low probability of detection on AWGN channels," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1921–1930, 2013.

[15] A. Abdelaziz and C. E. Koksal, "Fundamental limits of covert communication over MIMO AWGN channel," in *Proceedings of IEEE Conference on Communications Network Security (CNS)*, pp. 1–9, Las Vegas, NV, USA, October 2017.

[16] L. Wang, G. W. Wornell, and L. Zheng, "Fundamental limits of communication with low probability of detection," *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 3493–3503, 2016.

[17] P. H. Che, M. Bakshi, and S. Jaggi, "Reliable deniable communication: hiding messages in noise," in *Proceedings of IEEE International Symposium on Information Theory*, pp. 2945–2949, Istanbul, Turkey, July 2013.

[18] B. He, S. Yan, X. Zhou, and V. K. N. Lau, "On covert communication with noise uncertainty," *IEEE Communications Letters*, vol. 21, no. 4, pp. 941–944, 2017.

[19] T. V. Sobers, B. A. Bash, S. Guha, D. Towsley, and D. Goeckel, "Covert communication in the presence of an uninformed jammer," *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 6193–6206, 2017.

[20] K. Shahzad, X. Zhou, and S. Yan, "Covert communication in fading channels under channel uncertainty," in *Proceedings of IEEE 85th Vehicular Technology Conference*, pp. 1–5, Sydney, Australia, June 2017.

[21] J. Hu, K. Shahzad, S. Yan, X. Zhou, F. Shu, and J. Li, "Covert communications with a full-duplex receiver over wireless fading channels," in *Proceedings of IEEE International Conference on Communications*, pp. 20–24, Beijing, China, May 2018.

[22] K. Shahzad, X. Zhou, S. Yan, J. Hu, F. Shu, and J. Li, "Achieving covert wireless communications using a full-duplex receiver," *IEEE Transactions on Wireless Communications*, vol. 17, no. 12, pp. 8517–8530, 2018.

[23] B. He, S. Yan, X. Zhou, and H. Jafarkhani, "Covert wireless communication with a Poisson field of interferers," *IEEE Transactions on Wireless Communications*, vol. 17, no. 9, pp. 6005–6017, 2018.

[24] B. A. Bash, D. Goeckel, and D. Towsley, "LPD communication when the warden does not know when," in *Proceedings of IEEE International Symposium on Information Theory*, pp. 606–610, Honolulu, HW, USA, July 2014.

[25] J. Hu, S. Yan, X. Zhou, F. Shu, and J. Wang, "Covert transmission with harvested energy by a wireless powered relay," in *Proceedings of ICC*, Kansas, MO, USA, May 2018.

[26] S. Yan, B. He, Y. Cong, and X. Zhou, "Covert communication with finite block length in AWGN channels," in *Proceedings of IEEE ICC*, pp. 1–6, Bikaner, India, May 2017.

[27] B. He, X. Zhou, Y. Cong, and A. L. Swindlehurst, "Delay-intolerant covert communications with either fixed or random transmit power," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 1, pp. 129–140, 2019.

[28] X. Zhou, S. Yan, J. Hu, J. Sun, J. Li, and F. Shu, "Joint optimization of a UAV trajectory and transmit power for covert communications," arXiv preprint arXiv:1812.00583v1, 2018.

[29] T.-X. Zheng, H.-M. Wang, D. W. K. Ng, and J. Yuan, "Multi-antenna covert communications in random wireless networks," *IEEE Transactions on Wireless Communications*, vol. 18, no. 3, pp. 1974–1987, 2019.

[30] L. Yang, W. Yang, S. Xu, L. Tang, and Z. He, "Achieving covert wireless communications using a full-duplex multi-antenna receiver," in *Proceedings of International Conference on Computer and Communications (ICCC)*, Changchun, China, December 2019.

[31] X. Lu, W. Yang, Y. Cai, and X. Guan, "Proactive eavesdropping via covert pilot spoofing attack in multi-antenna systems," *IEEE Access*, vol. 7, pp. 151295–151306, 2019.

[32] C. Wu, Y. Xiao, and P. Yang, "Covert information embedded spatial modulation," *IEEE Communications Letters*, vol. 24, no. 11, pp. 2426–2430, 2020.

[33] F. Shu, T. Xu, J. Hu, and S. Yan, "Delay-constrained covert communications with a full-duplex receiver," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 813–816, 2019.

[34] J. Wang, W. Tang, Q. Zhu, X. Li, H. Rao, and S. Li, "Covert communication with the help of relay and channel uncertainty," *IEEE Wireless Communications Letters*, vol. 8, no. 1, pp. 317–320, 2019.

[35] O. Shmuel, A. Cohen, O. Gurewitz, and A. Cohen, "Multi-antenna jamming in covert communication," *Computer Science*, 2020, https://arxiv.org/abs/2001.03146.

[36] C. Basilevsky, *Applied Matrix Algebra in the Statistical Sciences*, North-Holland, New York, NY, USA, 1983.

[37] E. Lehmann and J. Romano, *Testing Statistical Hypotheses*, Springer Press, New York, NY, USA, 3rd edition, 2005.

Hindawi
Complexity
Volume 2021, Article ID 6941701, 15 pages
https://doi.org/10.1155/2021/6941701

Research Article

Further Results on Exponentially Robust Stability of Uncertain Connection Weights of Neutral-Type Recurrent Neural Networks

Wenxiao Si, Tao Xie, and Biwen Li

Hubei Normal University, Huangshi 435002, Hubei, China

Correspondence should be addressed to Tao Xie; 1044806961@qq.com

Received 25 June 2021; Revised 11 August 2021; Accepted 30 August 2021; Published 23 September 2021

Academic Editor: Yue Song

Further results on the robustness of the global exponential stability of recurrent neural network with piecewise constant arguments and neutral terms (NPRNN) subject to uncertain connection weights are presented in this paper. Estimating the upper bounds of the two categories of interference factors and establishing a measuring mechanism for uncertain dual connection weights are the core tasks and challenges. Hence, on the one hand, the new sufficient criteria for the upper bounds of neutral terms and piecewise arguments to guarantee the global exponential stability of NPRNN are provided. On the other hand, the allowed enclosed region of dual connection weights is characterized by a four-variable transcendental equation based on the preceding stable NPRNN. In this way, two interference factors and dual uncertain connection weights are mutually restricted in the model of parameter-uncertainty NPRNN, which leads to a dynamic evolution relationship. Finally, the numerical simulation comparisons with stable and unstable cases are provided to verify the effectiveness of the deduced results.

1. Introduction

Since recurrent neural networks (RNNs) have the ability of parallel processing, distributed information storage, and associative deep learning, a series of neural networks such as Hopfield neural network, Cohen–Grossberg neural network, cellular neural network, BAM neural network, high-order cellular neural network, and shunt inhibition neural network, which are the typical representatives of RNNs, have attracted extensive attention over the years. Accordingly, with the in-depth study of RNNs, it can be seen that the stability is a forerunner condition for the multifarious practical applications. Hence, the research on the stability of the system is becoming more and more abundant [1–14], such as asymptotic stability [1], exponential stability [2–4], multistability [5], synchronization [6], dissipativity [3, 7, 8], region stability [9], memristor-based dynamic behavior stability [10], and exponential Lagrange stability [11, 12]. Additionally, in terms of the widespread application fields such as the visual optimization, image processing, language recognition, associative memory, and other fields, the stability of RNNs has become an indispensable dynamical behavior characteristic which must be further considered.

Exponential stability and robustness, as for the classical dynamic behaviors of neural networks, have been studied extensively in the past few years [2, 3, 13–22]. On the one hand, global exponential stability reveals the superiority that it guarantees the system can be fleetly stable at the equilibrium point with an exponential decay rate, which deservedly leads to a rapid stabilization and dramatically saves the response time. Furthermore, the decay rate value can be intuitively captured [14]. On the other hand, due to the tunable flexibility and broad applicability, the robustness is always endowed with different meanings in disparate practical scenarios. Taking [13] as an example, a natural question is raised and deeply explored: how much the interference intensity can a disturbed system bear to realize the stability again on the basis of its original stability, which implies the exact connotation of the robustness studied in this paper. In addition, the emerging literatures [1–23] have also indirectly confirmed that various generalized stability behaviors of the neural networks are inevitably and immensely subject to the category and quantity of disturbances,

for instance, time delays [1–8], stochastic disturbances [8, 13, 15], parameter disturbances [9], piecewise constant arguments [14], neutral terms [22], and Markov switching [12, 23]. Hence, the subsequent perturbations will be attached to RNNs to further examine and guarantee the robustness of global exponential stability (RoGES) of RNNs.

Neutral-type RNNs refer to RNNs with the neutral terms appeared in the derivative part, which makes the nonlinear perturbations affect not only the current states but also the states of derivative part. At present, neural networks with neutral terms have been applied in practice, for instance, the electrical interconnect and the electromagnetic interference design in digital computers are used as the specific physical application background for delayed neutral-type differential equations [24]. In combination with the statements mentioned above and some existing literatures, two species of exogenous interferences that diffusely appear in the modern engineering should be taken into consideration for neutral-type RNNs. On the one hand, due to the pervasive restrictions of the switching speed of each node of the neural networks, time delay is such a kind of inevitable interference element that it has been widely applied to various neural networks. And, piecewise constant argument studied here is another form of time delay, which unifies the hysteresis and advance. The investigation of RNNs subject to piecewise arguments is a breakthrough in literature [14], and the sufficient conditions for the exponential stability of a class of RNNs with piecewise arguments are given by constructing Lyapunov functions. At present, there are also some literatures further exploring the properties of the systems equipped with piecewise arguments [15, 25, 26]. On the other hand, it can be found that the parameter-intensity of the connection weight matrix is a crucial index affecting the stability of the systems [16–21, 27, 28]. Firstly, since parametric uncertainty was introduced in [16, 17, 27, 28], diverse forms of uncertain connection weights were applied to neural networks. Later, Zhu and Shen [18] visually depicted the boundary of the uncertain connection weights by the enclosed curve graph method and provided the sufficient conditions for RNNs disturbed by uncertain parameters to achieve the global exponential stability. Furthermore, time delays and random perturbations are additionally attached on RNNs in [19–21]. At present, uncertain connection weights have been applied to various fields, such as medicine, information transmission, and operations' research and planning [29–32]. It follows that the intensity of the connection weights is a highly mutable but immensely critical indicator.

By virtue of the existing literatures in the past decades, the stability of neutral-type RNNs with piecewise arguments or other time delays is explored by [22, 33–37] and some relatively mature methods have been widely used, such as the Euler–Maclaurin method, one-leg method, block boundary value method, and multidomain Legendre spectral collocation method. In addition, the investigations of RNNs instead of neutral-type RNNs with uncertain connection weights also have been carried out in [16–21, 27, 28]. However, there is hardly any studies aiming for the RoGES of the recurrent neural network with neutral terms and piecewise constant arguments (NPRNN) with uncertain connection weights. Hence, it is a notable problem that how much the parameter intensity of dual connection weights can a disturbed NPRNN endure to remain stable again on the basis of the original stable NPRNN.

Motivated by the above statements, here we investigate the RoGES of the NPRNN with uncertain connection weights. Naturally, the ultimate goal herein is to quantify the boundary values of the connection weight matrices on the basis of stable NPRNN. Hence, the main contributions are as follows. (1) The upper bounds of piecewise arguments and neutral terms that NPRNN can maintain stability are established. We solve the tricky neutral terms newly appeared in the derivative part of RNN by constructing a Lipschitz condition, and then, we obtain the upper bound of the neutral terms by solving a univariate transcendental equation. Afterwards, by fixing an appropriate value of neutral terms, the upper bound of piecewise arguments is settled by solving several different binary transcendental equations. (2) An enclosed curve about dual connection weights (σ and λ) is acquired by solving the newly established four-element (two disturbances and two uncertain connection weights) transcendental equation on the basis of the preceding stable NPRNN, which is the core significant results of this paper (more details can be seen in Remarks 2 and 3). (3) By virtue of the above two upper bounds of the interference factors and the characterized enclosed curve about the dual-parameter intensity, the simulation results indicate that, as long as one of the interference values exceeds the derived bounds, the preceding stable NPRNN will be unstable, which also intuitively confirmed the validity of deduced results. Based on what has been established above, these four interference elements are mutually restricted and the relationship among these factors established by a four-element transcendental equation is dynamic.

The rest of the paper is as follows. The preliminaries and the model descriptions are included in Section 2. Then, the feasible threshold values of the piecewise arguments, the neutral terms, and the uncertain connection weights to achieve the RoGES of the parameter-uncertainty NPRNN are discussed in Section 3. Accordingly, the simulation comparisons to verify the validity of the deduced results are shown in Section 4. Finally, a brief conclusion and some feasible prospects for future work are given in Section 5.

2. Problem Formulation

Based on this paper, denote $\mathbf{N}$ as the natural number set and $\mathbf{R}$ as the real number set. And, for any constant n, denote $\underline{n} = \{1, 2, \ldots, n\}$. $\mathbf{Z}^+$ is the positive integer set. $\mathbb{R}^n$, $\mathbb{R}^{n\times m}$, and $\mathbb{R}^+$ stand for n-dimensional Euclidean space, $n \times m$ real matrix space, and positive real space, respectively. Let $\|\cdot\|$ be the Euclidean norm, and the operator norm of matrix A is defined as $\|A\| = \sup\{\|Ax\|: \|x\| = 1\}$. Denote two piecewise constant argument real-value sequences $\{\theta_i\}$ and $\{\eta_i\}$, such that $\theta_i < \eta_i < \theta_{i+1}$, $\theta_i \longrightarrow \infty$ when $i \longrightarrow \infty$ and $i \in \mathbb{N}$.

Consider a NPRNN as

$$\begin{cases} d[v(t) - G(v(t))] = [-Av(t) + Bf(v(t)) + Cg(v(\gamma(t)))]dt, \\ v(t_0) = v_0 \in \mathbb{R}^n, \end{cases} \tag{1}$$

where $v(t) = (v_1(t), \ldots, v_n(t)) \in \mathbb{R}^n$ represents the state vector of the neurons, $G(v(t))$: $\mathbb{R}^n \longrightarrow \mathbb{R}^n$ is the neutral-type function of $v(t) \in \mathbb{R}^n$, $A = \text{diag}\{a_1, \ldots, a_n\}$ is a self-feedback connection weight matrix, $B = (b_{ij})_{n\times n}$ and $C = (c_{ij})_{n\times n}$ are the connection weight matrices of $f(v(t))$ and $g(v(\gamma(t)))$, respectively, where $A, B, C \in \mathbb{R}^{n\times n}$, $f(v(t))$ and $g(v(\gamma(t)))$ are continuous vector-value activation functions, and $v(t)$ and $v(\gamma(t))$ are the current state and piecewise argument state, respectively.

For the case that $G(v(t)) = 0$ and $\gamma(t) = t$, NPRNN (1) becomes the following RNN model:

$$\begin{cases} \dot{r}(t) = -Ar(t) + Bf(r(t)) + Cg(r(t)), \\ r(t_0) = r_0 \in \mathbb{R}^n. \end{cases} \tag{2}$$

Definition 1. For any $t_0 \in \mathbb{R}^+$ and $r_0 \in \mathbb{R}^n$, if there are $\alpha, \beta > 0$, such that $\|r(t, t_0, r_0)\| \le \alpha\|r_0\|\exp\{-\beta(t - t_0)\}$ holds, and the state $r(t, t_0, r_0)$ of (2) can achieve globally exponential stability.

Definition 2. For any $t_0 \in \mathbb{R}^+$ and $v_0 \in \mathbb{R}^n$, if there are $\widetilde{\alpha}, \widetilde{\beta} > 0$, such that $\|v(t, t_0, v_0)\| \le \widetilde{\alpha}\|v_0\|\exp\{-\widetilde{\beta}(t - t_0)\}$ holds, and the state $v(t, t_0, v_0)$ of (1) can achieve globally exponential stability.

To deduce the main results, some needed assumptions throughout the paper are given:

(H1) The activation functions $f(\cdot)$, $g(\cdot) \in \mathbb{R}^n$, and there are Lipschitz constants l_1 and $l_2 > 0$ such that

$$\begin{aligned} \|f(\varsigma) - f(\varrho)\| &\le l_1\|\varsigma - \varrho\|, \\ \|g(\varsigma) - g(\varrho)\| &\le l_2\|\varsigma - \varrho\|, \end{aligned} \tag{3}$$

hold for any $\varsigma, \varrho \in \mathbb{R}^n$, where $f(\cdot)$ and $g(\cdot)$ are endowed with the initial values $f(0) = 0$ and $g(0) = 0$.

(H2) Assume that neutral-type function $G(\cdot)$ satisfies a Lipschitz condition:

$$\|G(\varsigma) - G(\varrho)\| \le L\|\varsigma - \varrho\|, \tag{4}$$

where $\varsigma, \varrho \in \mathbb{R}^n$.

(H3) For a real-value sequence θ_k, there is a $\theta > 0$, which makes $\theta_{k+1} - \theta_k \le \theta$ true for any $k \in \mathbf{N}$.

(H4) Assume that

$$\begin{aligned} &(L + l_2\theta\|C\|) + \theta[\|A\| + l_1\|B\|]\left\{\frac{1 + L + l_2\theta\|C\|}{1 - L}\right. \\ &\left.\times \exp\left\{\frac{\theta(\|A\| + l_1\|B\|)}{1 - L}\right\}\right\} < 1. \end{aligned} \tag{5}$$

(H5) Assume that

$$\begin{aligned} &\alpha\exp\{-\beta(T - \theta)\} + \left[\frac{L\alpha\exp(-\beta\theta) + L}{1 - L} + \frac{2\alpha l_2\|C\|}{\beta(1 - L)^2}\right] \\ &\times \exp\left\{\frac{2T(\|A\| + l_1\|B\|)}{(1 - L)} + \frac{2T[(3 - L)l_2\|C\|]}{(1 - L)^2}\right\} < 1. \end{aligned} \tag{6}$$

Remark 1. For the sake of convenience, some symbol descriptions are listed as follows:

$$\begin{aligned} \varpi_1 &= \frac{1 + L + l_2\theta\|C\|}{1 - L}, \\ \varpi_2 &= \frac{[(1 + |\sigma|)\|A\| + (1 + |\lambda|)l_1\|B\|]}{1 - L}, \\ \varpi_3 &= \frac{[l_2\|C\|(1 + \Gamma) + |\sigma|\|A\| + |\lambda|l_1\|B\|]\alpha/\beta}{1 - L} + \frac{L\alpha\exp(-\beta\theta) + L}{1 - L}, \\ \varpi_4 &= \frac{[(1 + |\sigma|)\|A\| + l_1(1 + |\lambda|)\|B\| + l_2\|C\|(2 + \Gamma)]}{1 - L}, \\ \xi &= (L + l_2\theta\|C\|) + [(1 + |\sigma|)\|A\| + (1 + |\lambda|)l_1\|B\|] \times \theta\{\varpi_1\exp(\varpi_2\theta)\}, \\ \Gamma &= \frac{1 + L}{1 - \xi}, \quad \xi < 1, \\ T &> \ln\frac{\alpha}{\beta} > 0. \end{aligned} \tag{7}$$

3. Main Results

On the basis of NPRNN (1), in this paper, we mainly explore NPRNN with uncertain connection weights (called parameter-uncertainty NPRNN throughout the text):

$$\begin{cases} d[u(t) - G(u(t))] = [-(1+\sigma)Au(t) + (1+\lambda)Bf(u(t)) + Cg(u(\gamma(t)))]\mathrm{d}t, \\ u(t_0) = u_0 \in \mathbb{R}^n, \end{cases} \tag{8}$$

where σ and $\lambda \in \mathbf{R}$ are the extra interference intensity of connection weight matrices A and B, corresponding to the state $u(t)$ and activation function $f(u(t))$.

Then, the following auxiliary Lemma 1 aims to clarify the relationship between piecewise argument state $u(\gamma(t))$ and current state $u(t)$.

Lemma 1. *Denote that $u(t) = (u_1(t), \ldots, u_n(t))^T$ is a solution of (2), and (H1)-(H4) hold; then, the following inequality,*

$$\|u(\gamma(t))\| \le \Gamma \|u(t)\|, \tag{9}$$

holds, where

$$\begin{aligned} \varpi_1 &= \frac{1 + L + l_2\theta\|C\|}{1 - L}, \\ \varpi_2 &= \frac{[(1+|\sigma|)\|A\| + (1+|\lambda|)l_1\|B\|]}{1 - L}, \\ \Gamma &= \frac{1+L}{1-\xi}, \quad \xi < 1, \\ \xi &= (L + l_2\theta\|C\|) + [(1+|\sigma|)\|A\| + (1+|\lambda|)l_1\|B\|] \times \theta\{\varpi_1 \exp(\varpi_2\theta)\}. \end{aligned} \tag{10}$$

Proof. Fix $k \in \mathbf{N}$, and for $\gamma(t) = \eta_k$ and $t \in [\theta_k, \theta_{k+1})$, $t \in \mathbb{R}^+$, we have

$$\begin{aligned} & u(t) - G(u(t)) - [u(\eta_k) - G(u(\eta_k))] \\ &= \int_{\eta_k}^{t} [-(1+\sigma)Au(s) + (1+\lambda)Bf(u(s)) + Cg(u(\eta_k))]\mathrm{d}s. \end{aligned} \tag{11}$$

Applying the norm inequality on both sides of (11) and in accordance with (H1), we obtain

$$\begin{aligned} & \|u(t) - u(\eta_k)\| - \|G(u(\eta_k)) - G(u(t))\| \\ &\le (1+|\sigma|)\|A\| \int_{\eta_k}^{t} \|u(s)\|\mathrm{d}s + (1+|\lambda|)\|B\| \\ &\quad \times \int_{\eta_k}^{t} \|f(u(s))\|\mathrm{d}s + \|C\| \int_{\eta_k}^{t} \|g(u(\eta_k))\|\mathrm{d}s \\ &\le (1+|\sigma|)\|A\| \int_{\eta_k}^{t} \|u(s)\|\mathrm{d}s + (1+|\lambda|)\|B\| \\ &\quad \times \int_{\eta_k}^{t} l_1\|u(s)\|\mathrm{d}s + \|C\| \int_{\eta_k}^{t} l_2\|u(\eta_k)\|\mathrm{d}s \\ &\le (1+|\sigma|)\|A\| \int_{\eta_k}^{t} \|u(s)\|\mathrm{d}s + [(1+|\lambda|)l_1\|B\|] \\ &\quad \times \int_{\eta_k}^{t} \|u(s)\|\mathrm{d}s + l_2\|C\| \int_{\eta_k}^{t} \|u(\eta_k)\|\mathrm{d}s. \end{aligned} \tag{12}$$

In terms of (H2) and the norm inequality, we derive

$$\begin{aligned} \|u(t)\| &\le \|u(\eta_k)\| + L\|u(\eta_k)\| + L\|u(t)\| \\ &\quad + [(1+|\sigma|)\|A\| + (1+|\lambda|)l_1\|B\|] \int_{\eta_k}^{t} \|u(s)\|\mathrm{d}s \\ &\quad + l_2\|C\| \int_{\eta_k}^{t} \|u(\eta_k)\|\mathrm{d}s \\ &\le L\|u(t)\| + (1 + L + l_2\theta\|C\|)\|u(\eta_k)\| \\ &\quad + [(1+|\sigma|)\|A\| + (1+|\lambda|)l_1\|B\|] \int_{\eta_k}^{t} \|u(s)\|\mathrm{d}s. \end{aligned} \tag{13}$$

Directly, for $L \in (0,1)$, we have

$$\begin{aligned} \|u(t)\| &\le \frac{1 + L + l_2\theta\|C\|}{1 - L}\|u(\eta_k)\| \\ &\quad + \frac{[(1+|\sigma|)\|A\| + (1+|\lambda|)l_1\|B\|]}{1 - L} \int_{\eta_k}^{t} \|u(s)\|\mathrm{d}s \\ &\le \varpi_1\|u(\eta_k)\| + \varpi_2 \int_{\eta_k}^{t} \|u(s)\|\mathrm{d}s, \end{aligned} \tag{14}$$

where

$$\varpi_1 = \frac{1 + L + l_2\theta\|C\|}{1 - L},$$
$$\varpi_2 = \frac{[(1 + |\sigma|)\|A\| + (1 + |\lambda|)l_1\|B\|]}{1 - L}. \tag{15}$$

According to Gronwall–Bellman lemma, it follows that

$$\|u(t)\| \le \varpi_1 \exp\{\varpi_2\theta\}\|u(\eta_k)\|. \tag{16}$$

Moreover, for $t \in [\theta_k, \theta_{k+1})$, we obtain

$$\begin{aligned} &G(u(\eta_k)) - G(u(t)) \\ &= u(\eta_k) - u(t) + \int_{\eta_k}^{t} [-(1 + \sigma)Au(s) + (1 + \lambda)Bf(u(s)) + Cg(u(\eta_k))]\mathrm{d}s. \end{aligned} \tag{17}$$

In combination with (H1)–(H3) and (16), similarly, we obtain

$$\begin{aligned} \|u(\eta_k)\| &\le \|u(t)\| + L\|u(\eta_k)\| + L\|u(t)\| + [(1 + |\sigma|)\|A\| \\ &\quad + (1 + |\lambda|)l_1\|B\|] \int_{\eta_k}^{t} \|u(s)\|\mathrm{d}s \\ &\quad + l_2\|C\| \int_{\eta_k}^{t} \|u(\eta_k)\|\mathrm{d}s \\ &\le (1 + L)\|u(t)\| + L\|u(\eta_k)\| + [(1 + |\sigma|)\|A\| \\ &\quad + (1 + |\lambda|)l_1\|B\|]\theta\{\varpi_1 \exp\{\varpi_2\theta\}\}\|u(\eta_k)\| \\ &\quad + l_2\|C\| \int_{\eta_k}^{t} \|u(\eta_k)\|\mathrm{d}s \\ &\le (1 + L)\|u(t)\| + \{(L + l_2\theta\|C\|) + [(1 + |\sigma|)\|A\| \\ &\quad + (1 + |\lambda|)l_1\|B\|]\theta\{\varpi_1 \exp(\varpi_2\theta)\}\}\|u(\eta_k)\| \\ &\le (1 + L)\|u(t)\| + \xi\|u(\eta_k)\|, \end{aligned} \tag{18}$$

where

$$\begin{aligned} \xi = {} & (L + l_2\theta\|C\|) + [(1 + |\sigma|)\|A\| + (1 + |\lambda|)l_1\|B\|] \\ & \times \theta\{\varpi_1 \exp(\varpi_2\theta)\}. \end{aligned} \tag{19}$$

Therefore, by (18) and (H4), for $\eta_k = \gamma(t)$ and $k \in \mathbf{N}$, we can obtain

$$\begin{aligned} \|u(\gamma(t))\| &\le \frac{1 + L}{1 - \xi}\|u(t)\| \\ &=: \Gamma\|u(t)\|, \end{aligned} \tag{20}$$

where

$$\Gamma = \frac{1 + L}{1 - \xi}, \quad \xi < 1. \tag{21}$$

□

Theorem 1. *If (H1)–(H5) hold, (2) achieves globally exponential stability. Then, the conditions for achieving the RoGES of NPRNN can be given by the following (a) and (b):*

(a) Let $\overline{L}$ be the upper bound of neutral term compressibility coefficient; then, $L < \overline{L}$ and $\overline{L} \in (0, 1)$, and $\overline{L}$ is given by the following transcendental equation:

$$\begin{aligned} &\alpha \exp(-\beta T) + \frac{(\overline{L}\alpha + \overline{L})(1 - \overline{L})\beta + 2\alpha l_2\|C\|}{\beta(1 - \overline{L})^2} \\ &\times \exp\left\{\frac{2T[(1 - \overline{L})(\|A\| + l_1\|B\|) + (3 - \overline{L})l_2\|C\|]}{(1 - \overline{L})^2}\right\} < 1. \end{aligned} \tag{22}$$

(b) Let θ_3 be the upper bound of piecewise constant arguments, where θ_3 is given by

$$\theta < \theta_3 = \min\left\{\frac{T}{2}, \theta_1, \theta_2\right\}, \tag{23}$$

where θ_1 and θ_2 are the upper bounds satisfying (H4) and (H5), respectively, $T > (\ln \alpha)/\beta > 0$.

Furthermore, if the real selected values $L < \overline{L}$ and $\theta < \theta_3$ hold by (22) and (23), then the allowed intensity of (σ, λ) to achieve the RoGES of parameter-uncertainty NPRNN (8) should be in the inner of the following enclosed curve:

$$\alpha \exp(-\beta(T - \theta)) + \varpi_3 \exp\{2T\varpi_4\} = 1, \tag{24}$$

where

$$\begin{aligned} \varpi_3 &= \frac{L\alpha \exp(-\beta\theta) + L}{1 - L} \\ &\quad + \frac{[l_2\|C\|(1 + \Gamma) + |\sigma|\|A\| + |\lambda|l_1\|B\|]\alpha/\beta}{1 - L}, \\ \varpi_4 &= \frac{[(1 + |\sigma|)\|A\| + l_1(1 + |\lambda|)\|B\| + l_2\|C\|(2 + \Gamma)]}{1 - L}, \end{aligned} \tag{25}$$

and the other parameters are the same as Lemma 1. Besides, the boundary value of $|\sigma|$ is expressed as $|\sigma|_{\sup}$, and $|\lambda|$ is expressed as $|\lambda|_{\sup}$ hereinbelow.

Proof. From (2) and (8), for $t \ge t_0$, we obtain

$$
\begin{aligned}
u(t) - r(t) &= G(u(t)) - G\left(u\left(t_0\right)\right) \\
&+ \int_{t_0}^{t} \{-A[u(s) - r(s)] + B[f(u(s)) - f(r(s))] \\
&+ C[g(u(\gamma(s))) - g(u(s)) + g(u(s)) - g(r(s))] \\
&- \sigma A[u(s) - r(s) + r(s)] \\
&+ \lambda B[f(u(s)) - f(r(s)) + f(r(s))]\}\mathrm{d}s.
\end{aligned} \tag{26}
$$

Taking the absolute value of (26) for $t \le t_0 + 2T$, we obtain

$$
\begin{aligned}
\|u(t) - r(t)\| &\le \left\|G(u(t)) - G\left(u\left(t_0\right)\right)\right\| + \left[\|A\| + l_1\|B\| + l_2\|C\|\right. \\
&\quad \left. + |\sigma|\|A\| + |\lambda| l_1\|B\|\right] \int_{t_0}^{t} \|u(s) - r(s)\|\mathrm{d}s \\
&\quad + \|C\| \int_{t_0}^{t} \|g(u(\gamma(s))) - g(u(s))\|\mathrm{d}s \\
&\quad + \left(|\sigma|\|A\| + |\lambda| l_1\|B\|\right) \int_{t_0}^{t} \|r(s)\|\mathrm{d}s \\
&\le L\|u(t)\| + L\left\|u\left(t_0\right)\right\| + \left[(1 + |\sigma|)\|A\| + l_1(1 + |\lambda|)\|B\|\right. \\
&\quad \left. + l_2\|C\|\right] \int_{t_0}^{t} \|u(s) - r(s)\|\mathrm{d}s \\
&\quad + \|C\| l_2 \int_{t_0}^{t} (\|u(\gamma(s))\| + \|u(s)\|)\mathrm{d}s \\
&\quad + \left(|\sigma|\|A\| + |\lambda| l_1\|B\|\right) \int_{t_0}^{t} \|r(s)\|\mathrm{d}s \\
&\le L\|u(t) - r(t)\| + L\|r(t)\| + L\left\|u\left(t_0\right)\right\| + \left[(1 + |\sigma|)\|A\| + l_1(1 + |\lambda|)\|B\|\right. \\
&\quad \left. + l_2\|C\|\right] \int_{t_0}^{t} \|u(s) - r(s)\|\mathrm{d}s \\
&\quad + \|C\| l_2(1 + \Gamma) \int_{t_0}^{t} [\|u(s) - r(s)\| + \|r(s)\|]\mathrm{d}s \\
&\quad + \left(|\sigma|\|A\| + |\lambda| l_1\|B\|\right) \int_{t_0}^{t} \|r(s)\|\mathrm{d}s.
\end{aligned} \tag{27}
$$

In accordance with Definition 1 and the exponential stability of (2), $L \in (0, 1)$, when $t \in [t_0 - \theta, t_0 + \theta]$, further we derive

$$\begin{aligned}\|u(t)-r(t)\| \le{}& \frac{L}{1-L}\alpha\|u_0\|\exp(-\beta\theta)+\frac{L}{1-L}\|u_0\|\\ &+\frac{[(1+|\sigma|)\|A\|+l_1(1+|\lambda|)\|B\|+l_2\|C\|(2+\Gamma)]}{1-L}\\ &\times\int_{t_0}^{t}\|u(s)-r(s)\|\mathrm{d}s\\ &+\frac{l_2\|C\|(1+\Gamma)+|\sigma|\|A\|+|\lambda|l_1\|B\|}{1-L}\\ &\times\int_{t_0}^{t}\alpha\|r_0\|\exp(-\beta(s-t_0))\mathrm{d}s\\ \le{}&\left\{\frac{L\alpha\exp(-\beta\theta)+L}{1-L}\right.\\ &\left.+\frac{[l_2\|C\|(1+\Gamma)+|\sigma|\|A\|+|\lambda|l_1\|B\|]\alpha/\beta}{1-L}\right\}\|u_0\|\\ &+\frac{[(1+|\sigma|)\|A\|+l_1(1+|\lambda|)\|B\|+l_2\|C\|(2+\Gamma)]}{1-L}\\ &\times\int_{t_0}^{t}\|u(s)-r(s)\|\mathrm{d}s\\ \le{}&\varpi_3\|u_0\|+\varpi_4\int_{t_0}^{t}\|u(s)-r(s)\|\mathrm{d}s,\end{aligned}\tag{28}$$

where

$$\begin{aligned}\varpi_3 &= \frac{L\alpha\exp(-\beta\theta)+L}{1-L}+\frac{[l_2\|C\|(1+\Gamma)+|\sigma|\|A\|+|\lambda|l_1\|B\|]\alpha/\beta}{1-L},\\ \varpi_4 &= \frac{[(1+|\sigma|)\|A\|+l_1(1+|\lambda|)\|B\|+l_2\|C\|(2+\Gamma)]}{1-L}.\end{aligned}\tag{29}$$

By virtue of Gronwall inequality, for $t_0+\theta\le t\le t_0+2T$, we have

$$\begin{aligned}\|u(t)-r(t)\| &\le \varpi_3\|u_0\|\exp\{\varpi_4(t-t_0)\}\\ &\le \varpi_3\exp\{2T\varpi_4\}\sup_{t\in[t_0-\theta,t_0+\theta]}\|u(t)\|.\end{aligned}\tag{30}$$

Note that $\theta<(T/2)$, when $t_0-\theta+T\le t_0-\theta+2T$, by the exponential stability of (2), we further obtain

$$\begin{aligned}\|u(t)\| &\le \|r(t)\|+\|u(t)-r(t)\|\\ &\le [\alpha\exp(-\beta(T-\theta))+\varpi_3\exp\{2T\varpi_4\}]\\ &\quad\times\sup_{t_0-\theta\le t_0-\theta+T}\|u(t)\|\\ &=:\widetilde{C}\sup_{t_0-\theta\le t_0-\theta+T}\|u(t)\|,\end{aligned}\tag{31}$$

where $\widetilde{C}=\alpha\exp(-\beta(T-\theta))+\varpi_3\exp\{2T\varpi_4\}$. Then, if we define

$$F(\Gamma(\theta),L)=\alpha\exp\{-\beta(T-\theta)\}+\varpi_3\exp\{2T\varpi_4\},\tag{32}$$

obviously, there is

$$F(\Gamma(\theta),L)\le 1.\tag{33}$$

Firstly, by (H4), denote θ_1 as the unique maximal positive solution satisfying $\xi(\theta)|_{\sigma,\lambda=0}<1$, that is,

$$\begin{aligned}&(L+l_2\theta\|C\|)+\theta[\|A\|+l_1\|B\|]\left\{\frac{1+L+l_2\theta\|C\|}{1-L}\right.\\ &\left.\exp\left\{\frac{\theta(\|A\|+l_1\|B\|)}{1-L}\right\}\right\}<1,\end{aligned}\tag{34}$$

so

$$\Gamma(\theta)=\frac{1+L}{1-\xi(\theta)}\in\left(\frac{1+L}{1-L},\infty\right),\quad L\in(0,1),\tag{35}$$

holds for any $\theta\in(0,\theta_1)$. According to (32) and (H5), it derives

$$\begin{aligned}F\left(\frac{1+L}{1-L},L\right)|_{\lambda,\sigma=0} &= \alpha\exp\{-\beta(T-\theta)\}+\left[\frac{L\alpha\exp(-\beta\theta)+L}{1-L}+\frac{2\alpha l_2\|C\|}{\beta(1-L)^2}\right]\\ &\quad\times\exp\left\{\frac{2T(\|A\|+l_1\|B\|)}{(1-L)}+\frac{2T[(3-L)l_2\|C\|]}{(1-L)^2}\right\}<1,\end{aligned}\tag{36}$$

$$F(\infty,L)>1.\tag{37}$$

Thus, by the monotonicity of $F(\Gamma(\theta),L)$, there exists some $\widehat{\Gamma}(\theta)\in(((1+l)/(1-L)),\infty)$, which makes $F(\Gamma(\theta),L)|_{\lambda,\sigma=0}=1$ true. Then, if we denote θ_2 as the upper bound satisfying (36) and select

$$\theta_3=\min\left\{\frac{T}{2},\theta_1,\theta_2\right\},\tag{38}$$

hence, the upper bound of the piecewise arguments unaffected by the intensity of uncertain parameters is given.

Besides, since $\partial(\Gamma(\theta))/\partial\theta > 0$ and $\partial(F(\Gamma(\theta)), L)/\partial\theta > 0$, the upper bound of neutral terms can be given by

$$F\left(\frac{1+L}{1-L}, L\right)\Big|_{\theta,\lambda,\sigma=0} = \alpha \exp(-\beta T) + \frac{(L\alpha + L)(1-L)\beta + 2\alpha l_2\|C\|}{\beta(1-L)^2} \times \exp\left\{\frac{2T\left[(1-L)\left(\|A\| + l_1\|B\|\right) + (3-L)l_2\|C\|\right]}{(1-L)^2}\right\} < 1. \tag{39}$$

Finally, if we select the exact L and θ which stabilize NPRNN (1), the intensity boundary of dual parameters σ and λ to stabilize the new parameter-uncertainty NPRNN (8) will be obtained by the following transcendental equation in accordance with the above L and θ:

$$\alpha \exp(-\beta(T-\theta)) + \varpi_3 \exp\{2T\varpi_4\} = 1, \tag{40}$$

and some corresponding symbolic descriptions of (40) are shown in Remark 1.

Accordingly, selecting $\Lambda = -\ln\widetilde{C}/T$, where $\Lambda \geq 0$, and from (31), we can obtain

$$\sup_{t_0-\theta+T\leq t\leq t_0-\theta+2T}\left\|u(t, t_0, u_0)\right\| \leq \exp(-\Lambda T)\sup_{t_0-\theta\leq t\leq t_0-\theta+T}\left\|u(t, t_0, u_0)\right\|. \tag{41}$$

Subsequently, for any positive integer $q \in \underline{n}$, when $t \geq t_0 - \theta + (q-1)T$, by the continuity of the solutions of the system [13], invoking (41), further we have

$$\begin{aligned}
&\sup_{t_0-\theta+qT\leq t\leq t_0-\theta+(q+1)T}\left\|u(t; t_0, u_0)\right\| \\
&= \sup_{t_0-\theta+(q-1)T+T\leq t\leq t_0-\theta+qT+T}\Big\|u(t; t_0-\theta+(q-1)T, \\
&u(t_0-\theta+(q-1)T; t_0, u_0))\Big\| \\
&\leq \exp(-\Lambda T)\sup_{t_0-\theta+(q-1)T\leq t\leq t_0-\theta+qT}\left\|u(t; t_0, u_0)\right\| \\
&\leq \exp(-\Lambda qT)\sup_{t_0-\theta\leq t\leq t_0-\theta+T}\left\|u(t; t_0, u_0)\right\| \\
&= C\exp(-\Lambda qT),
\end{aligned} \tag{42}$$

where $C = \sup_{t_0-\theta\leq t\leq t_0-\theta+T}\|u(t, t_0, u_0)\|$. Hence, when $t_0 - \theta + qT \leq t \leq t_0 - \theta + (q+1)T$, for any integer $q \in \underline{n}$, we have

$$\begin{aligned}
\|u(t)\| &\leq C\exp(-\Lambda qT) \\
&\leq C\exp\{\Lambda[-(t-t_0) + (T-\theta)]\} \\
&\leq C\exp(\Lambda T)\exp(-\Lambda(t-t_0)).
\end{aligned} \tag{43}$$

In this way, (8) achieves globally exponential stability. □

Corollary 1. *Obviously, the allowed intensity of (σ, λ) to achieve the RoGES of parameter-uncertainty NPRNN (8) is derived in Theorem 1. In fact, Theorem 1 is also applicable for the following models (44) and (45):*

$$\begin{cases} d[u(t) - G(u(t))] = [-(1+\sigma)Au(t) + Bf(u(t)) + Cg(u(\gamma(t)))]\mathrm{d}t, \\ u(t_0) = u_0 \in \mathbb{R}^n, \end{cases} \tag{44}$$

$$\begin{cases} d[u(t) - G(u(t))] = [-Au(t) + (1+\lambda)Bf(u(t)) + Cg(y(\gamma(t)))]\mathrm{d}t, \\ u(t_0) = u_0 \in \mathbb{R}^n, \end{cases} \tag{45}$$

corresponding to the cases $\lambda = 0$ and $\sigma = 0$ for parameter-uncertainty NPRNN (8), respectively.

Proof. Similarly, if we set $\lambda = 0$ and $\sigma = 0$ for Theorem 1, respectively, the sufficient conditions to ensure the globally

exponential stability of (44) and (45) can be given promptly. □

Remark 2. It is not easy to directly handle (24) (i.e., (40)), a transcendental equation with four variables in Theorem 1. Hence, we adopt a special method: fix θ and L which can maintain the globally exponential stability of NPRNN (1), and then, in combination with some known parameters α, β, l_1, l_2, and T, (24) becomes an implicit transcendental equation with only two variables $|\sigma|$ and $|\lambda|$. Transcendental equation (24) is a key step to obtain the enclosed curve such as the one in Figure 1, which is also the core work of this paper.

Remark 3. The order of the calculation in this paper is as follows. At first, some parameters α, β, l_1, l_2, and T are given in advance, and the upper bound of the neutral terms ($\overline{L}$) can be derived by (22). Next, we fix an exact $L < \overline{L}$ which can ensure the exponential stability of NPRNN (1), and then, the supremum of the piecewise arguments (θ_3) is obtained by (23). Furthermore, in accordance with the appropriately selected values of L and θ (i.e., $L < \overline{L}$ and $\theta < \theta_3$) and equation (24), we can obtain the bounds of σ and λ which can guarantee the RoGES of parameter-uncertainty NPRNN (8) by MATLAB. The robustness of the system means that if the original NPRNN (1) is stable, the perturbed parameter-uncertainty NPRNN (8) can still remain stable as long as the selected values of σ and λ are included inside the closed curve given by equation (24).

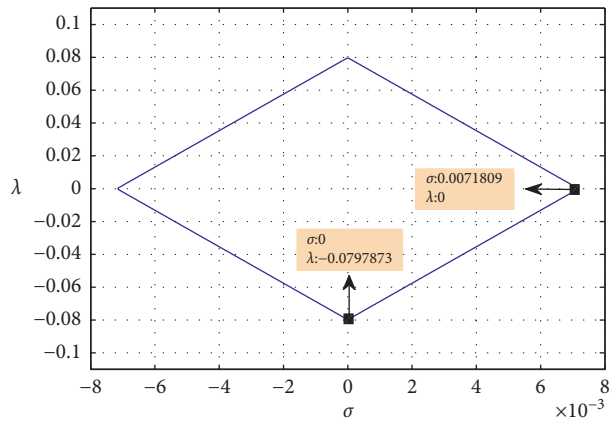


Figure 1: Stable enclosed curve of σ and λ in system (48).

4. Illustrative Example

Example 1. For a two-state RNN,

$$\begin{pmatrix} \dot{r}_1(t) \\ \dot{r}_2(t) \end{pmatrix} = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} r_1(t) \\ r_2(t) \end{pmatrix} + \begin{pmatrix} 0.09 & 0 \\ 0 & 0.09 \end{pmatrix}\begin{pmatrix} \tanh(r_1(t)) \\ \tanh(r_2(t)) \end{pmatrix} + \begin{pmatrix} 0.1 & 0.1 \\ 0.1 & 0.1 \end{pmatrix}\begin{pmatrix} \sin\left(\frac{r_1(t)}{2}\right) \\ \sin\left(\frac{r_2(t)}{2}\right) \end{pmatrix}. \tag{46}$$

Equation (46) can achieve globally exponential stability with $\alpha = 1$ and $\beta = 0.6$. According to Theorem 1, let $T = 0.5 > \ln 1/0.6 = 0$. From (H1), let $l_1 = 1$ and $l_2 = 0.5$. Besides, other parameters $\|A\| = 1$, $\|B\| = 0.09$, and $\|C\| = 0.02$ can also be listed; then, the stable state trajectory can be seen in Figure 2.

When the neutral terms and piecewise arguments are attached to RNN (46), (46) can be written as NPRNN (47):

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} v_1(t) + L\sin(v_1(t)) \\ v_2(t) + L\sin(v_2(t)) \end{pmatrix} = -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} v_1(t) \\ v_2(t) \end{pmatrix} + \begin{pmatrix} 0.09 & 0 \\ 0 & 0.09 \end{pmatrix}\begin{pmatrix} \tanh(v_1(t)) \\ \tanh(v_2(t)) \end{pmatrix} + \begin{pmatrix} 0.1 & 0.1 \\ 0.1 & 0.1 \end{pmatrix}\begin{pmatrix} \sin\left(\frac{v_1(\gamma(t))}{2}\right) \\ \sin\left(\frac{v_2(\gamma(t))}{2}\right) \end{pmatrix}, \tag{47}$$

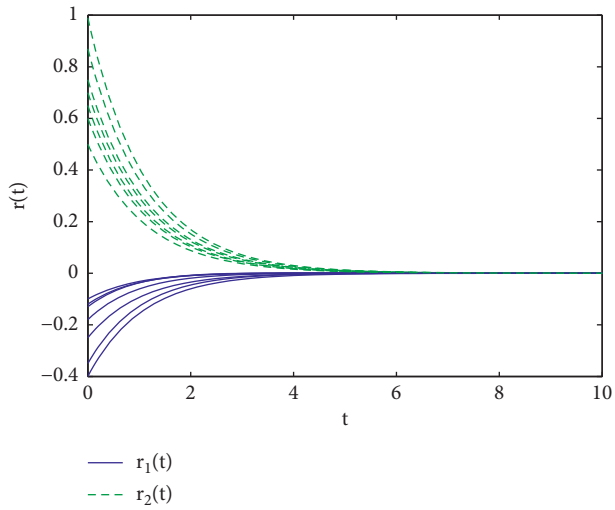


Figure 2: Stable state trajectories of $r_1(t)$ and $r_2(t)$ in (46).

where $\{\theta_k\} = \{k/10\}$ and $\{\eta_k\} = \{2k + 1/20\}$, $k \in \mathbf{N}$. The piecewise constant argument function $\gamma(t) = \eta_k$ when $t \in [\theta_k, \theta_{k+1})$, $k \in \mathbf{N}$.

Lastly, if we consider additional dual-parameter perturbations for (47), then parameter-uncertainty NPRNN (48) is formulated as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} u_1(t) + L\sin(u_1(t)) \\ u_2(t) + L\sin(u_2(t)) \end{pmatrix} = -(1+\sigma)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} + (1+\lambda)\begin{pmatrix} 0.09 & 0 \\ 0 & 0.09 \end{pmatrix} \times \begin{pmatrix} \tanh(u_1(t)) \\ \tanh(u_2(t)) \end{pmatrix} + \begin{pmatrix} 0.1 & 0.1 \\ 0.1 & 0.1 \end{pmatrix}\begin{pmatrix} \sin\left(\dfrac{u_1(\gamma(t))}{2}\right) \\ \sin\left(\dfrac{u_2(\gamma(t))}{2}\right) \end{pmatrix}. \tag{48}$$

In order to make the process clearer, the following explanations will be divided into two parts to illustrate the RoGES of system (48). On the one hand, we will explain how much the interference intensity of neutral terms and piecewise arguments the system (47) can tolerate to be stable again based on the stable RNN (46) depicted in Figure 2.

Firstly, calculating the following equation by MATLAB,

$$\exp(-0.3) + \frac{2L(1-L)\times 0.6 + 0.02}{0.6\times(1-L)^2} \times \exp\left\{\frac{1.09\times(1-L) + 0.01\times(3-L)}{(1-L)^2}\right\} = 1, \tag{49}$$

then the upper bound of L is obtained: $\overline{L} = 0.0231$.

Next, if we select $L = 0.01 < \overline{L} = 0.0231$ and substitute L, α, β, l_1, l_2, and T into (H4) and (H5), we get $\theta_1 = 0.5053$ and $\theta_2 = 0.2004$, respectively. So, the upper bound of θ is given by the following θ_3:

$$\begin{aligned} \theta < \theta_3 &= \min\left\{\frac{T}{2}, \theta_1, \theta_2\right\} \\ &= \min\{0.25, 0.5053, 0.2004\} \\ &= 0.2004. \end{aligned} \tag{50}$$

According to Theorem 1, the states $v_1(t)$ and $v_2(t)$ of (47) will be stable with $L = 0.01$ and $\theta = 0.1$, which is shown in Figure 3.

On the other hand, we will explain the intensity of connection weights σ and λ that parameter-uncertainty NPRNN (48) can tolerate based on the stable NPRNN (47) shown in Figure 3.

Therefore, we fix parameters $L = 0.01$ and $\theta = 0.1$ so as to satisfy the stable conditions in Theorem 1 and be consistent with the parameter setting in Figure 3. Subsequently, the stable region with (σ, λ) can be solved by a transcendental equation by MATLAB:

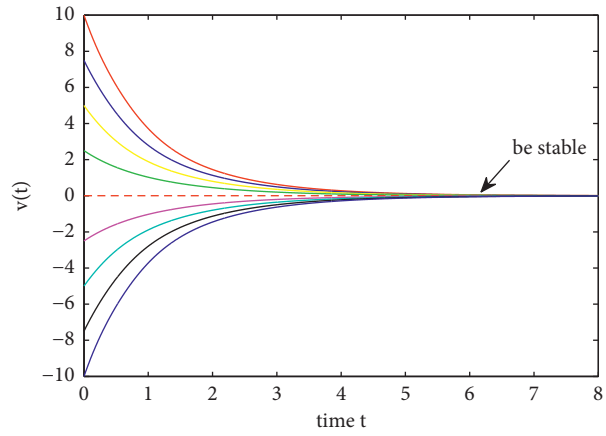


Figure 3: Stability of $v_1(t)$ and $v_2(t)$ in (47) with $L = 0.01$ and $\theta = 0.1$.

$$\exp(-0.24) + \frac{0.01 \times \exp(-0.06) + 0.01}{0.99} + \frac{[0.01 \times (1 + (1.01/(1-\xi))) + |\sigma| + 0.09|\lambda|] \times (1/0.6)}{0.99} \times \exp\left\{\frac{(1+|\sigma|) + 0.09(1+|\lambda|) + 0.01 \times (2 + (1.01/(1-\xi)))}{0.99}\right\} = 1, \tag{51}$$

where

$$\xi = 0.011 + [(1+|\sigma|) + 0.09(1+|\lambda|)] \times 0.1 \times \left\{\frac{1.011}{0.99} \exp\left\{\frac{0.1 \times [(1+|\sigma|) + 0.09(1+|\lambda|)]}{0.99}\right\}\right\}. \tag{52}$$

According to the above implicit transcendental equation, the stable region of σ and λ is depicted in Figure 1, that is, if the intensity range of (σ, λ) is in the inner of the enclosed curve in Figure 1, then parameter-uncertainty NPRNN (48) will be stable again based on the stable NPRNN (47). Therefore, some stable and unstable cases are given to verify the RoGES of system (48), such as the stable case (i) and unstable cases (ii)-(v).

(i) If $(\sigma, \lambda) = (0.001, 0.01)$, where $|\sigma| = 0.001 < |\sigma|_{\sup} = 0.0071809$ and $|\lambda| = 0.01 < |\lambda|_{\sup} = 0.0797873$, according to Figure 1, the stable state trajectories of $u_1(t)$ and $u_2(t)$ are shown in Figure 4.

Otherwise, some unstable cases should be provided to illustrate. If the connection weight σ exceeds the boundary value $|\sigma|_{\sup}$ or λ exceeds the boundary value $|\lambda|_{\sup}$ in Figure 1, parameter-uncertainty NPRNN (48) will be unstable even based on stable NPRNN (47), which can be illustrated in the following:

(ii) If $(\sigma, \lambda) = (0, 18)$, where $|\lambda| = 18 > |\lambda|_{\sup} = 0.0797873$, the unstable state trajectories are shown in Figure 5.

(iii) If $(\sigma, \lambda) = (-1.6, 0)$, where $|\sigma| = 1.6 > |\sigma|_{\sup} = 0.0071809$, the unstable state trajectories are shown in Figure 6.

(iv) If $(\sigma, \lambda) = (0.5, 20)$, where $|\sigma| = 0.5 > |\sigma|_{\sup} = 0.0071809$ and $|\lambda| = 20 > |\lambda|_{\sup} = 0.0797873$, the unstable state trajectories are shown in Figure 7.

(v) If $(\sigma, \lambda) = (-0.5, 20)$, similarly, the unstable state trajectories are shown in Figure 8.

Remark 4. Figure 1 is an enclosed curve about σ and λ with the boundary value $|\sigma|_{\sup}$ and $|\lambda|_{\sup}$. Figure 1 is depicted by (24); more details of the calculation process of (24) can be seen in Remark 2. According to Remark 2, we transform (24) from a four-variable equation to a two-variable equation only related to $|\sigma|$ and $|\lambda|$. By virtue of computer software MATLAB, an enclosed region of σ and λ, and the boundary values $|\sigma|_{\sup} = 0.0071809$ and $|\lambda|_{\sup} = 0.0797873$ are all directly exhibited in Figure 1 to demonstrate the RoGES of parameter-uncertainty NPRNN (48). If we select $|\sigma| < |\sigma|_{\sup}$ and $|\lambda| < |\lambda|_{\sup}$, system (48) will be stable again based on stable Figure 3. Or else, if $|\sigma| > |\sigma|_{\sup}$ or $|\lambda| > |\lambda|_{\sup}$ holds, system (48) will lose its stability even on the basis of stable Figure 3.

Remark 5. Figures 1–8 systematically and intuitively prove the robustness of the system (48). From Theorem 1, the supremum of neutral terms $\overline{L}$ is derived by (22), and the supremum of piecewise arguments θ_3 is derived by (23), and the boundary of uncertain-parameter intensity is characterized by (24). Figures 1–8 are obtained as follows. (1) Fix some necessary parameters α, β, l_1, l_2, and T; an initial globally exponentially stable RNN (46) is given by Figure 2. (2) According to $\overline{L}$ and θ_3 settled by (22) and (23), here we select $L = 0.01 < \overline{L} = 0.0231$ and

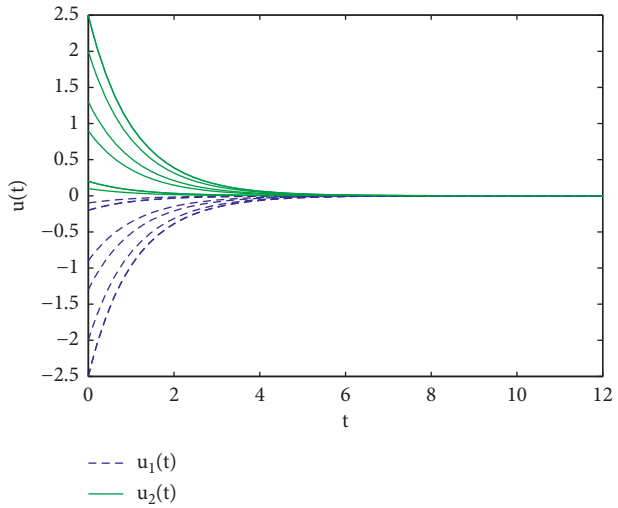


Figure 4: Stability of $u_1(t)$ and $u_2(t)$ in (48) with $(\sigma, \lambda) = (0.001, 0.01)$.

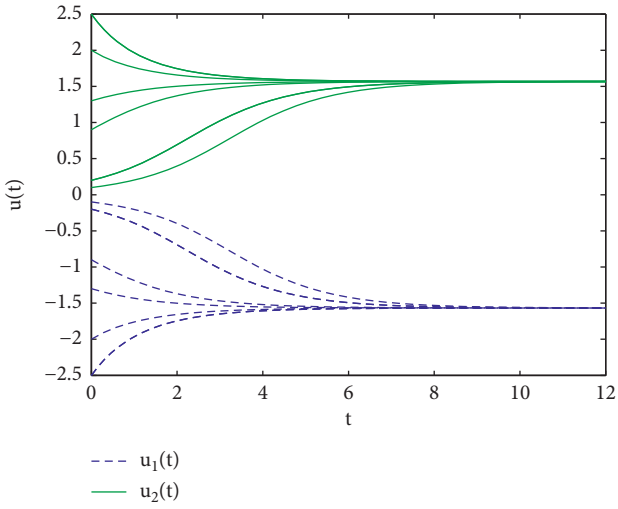


Figure 5: Instability of $u_1(t)$ and $u_2(t)$ in (48) with $(\sigma, \lambda) = (0, 18)$.

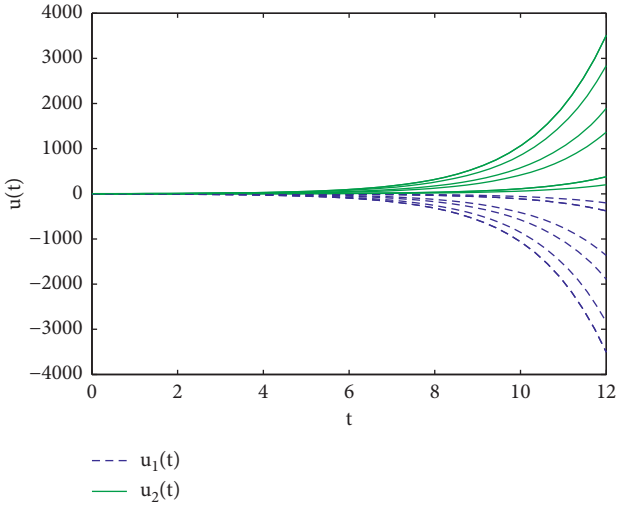


Figure 6: Instability of $u_1(t)$ and $u_2(t)$ in (48) with $(\sigma, \lambda) = (-1.6, 0)$.

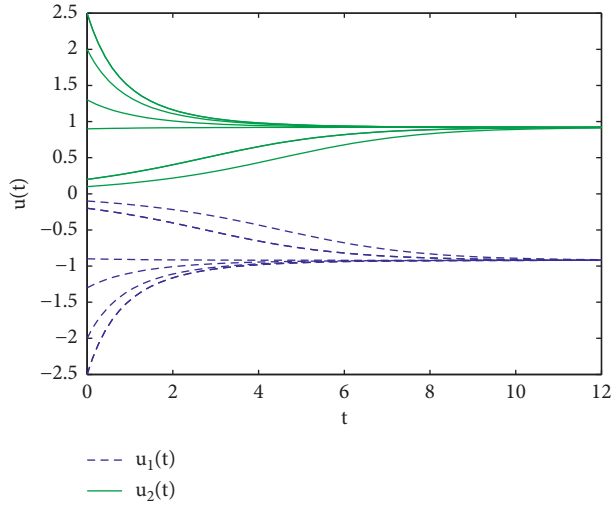


FIGURE 7: Instability of $u_1(t)$ and $u_2(t)$ in (48) with $(\sigma, \lambda) = (0.5, 20)$.

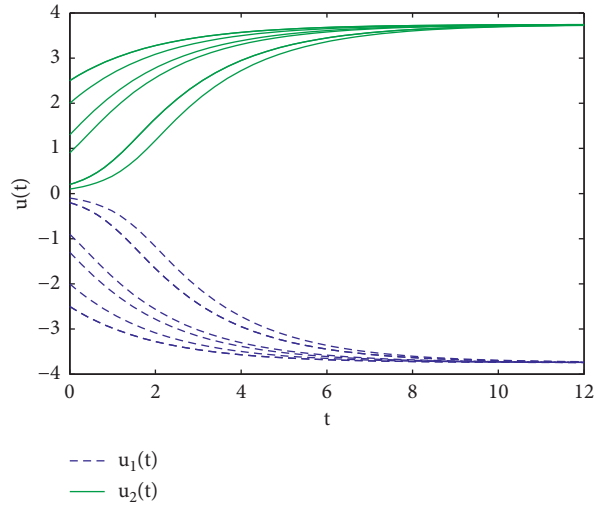


FIGURE 8: Instability of $u_1(t)$ and $u_2(t)$ in (48) with $(\sigma, \lambda) = (-0.5, 20)$.

$\theta = 0.1 < \theta_3 = 0.2004$; then, the NPRNN (47) can achieve exponential stability and the stable trajectories are shown in Figure 3. (3) Figure 1 is an stable enclosed curve about σ and λ depicted by (24) (more details for Figure 1 is shown in Remark 4). (4) If connection weights $|\sigma| = 0.001 < |\sigma|_{\text{sup}} = 0.0071809$ and $|\lambda| = 0.01 < |\lambda|_{\text{sup}} = 0.0797873$, the parameter-uncertainty NPRNN (48) will be stable again, and the stable trajectories are shown in Figure 4. (5) However, as long as either $|\sigma| > |\sigma|_{\text{sup}}$ or $|\lambda| > |\lambda|_{\text{sup}}$ is true, system (48) will be unstable, that is, the final-time state value does not converge to the same equilibrium point, which can be seen in Figures 5–8. More importantly, Figures 5–8 are essential and favourable reference frames for demonstrating the robustness of the parameter-uncertainty NPRNN. Furthermore, a brief calculation flowchart of Figures 1–8 is provided in Figure 9.

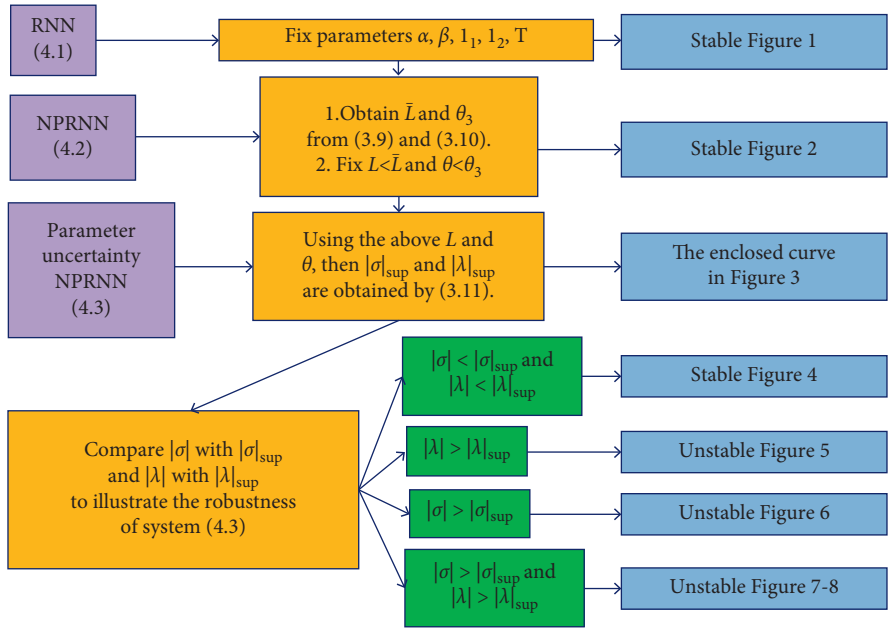


Figure 9: The calculation process of Figures 1–8 in Example 1.

5. Conclusion

This paper further explored the robustness of global exponential stability of NPRNN with uncertain dual connection weight matrices. Firstly, a foundational lemma is given. The relationship between the piecewise argument state and the current state is obtained by virtue of Gronwall inequality and the Lipchitz conditions. In addition, a main theorem is given. The upper bounds of the two interference factors and the coupling restrictions of uncertain dual connection weights are deduced by solving a four-element transcendental equation based on stable NPRNN. Finally, the stable and unstable cases of parameter-uncertainty NPRNN are analyzed by a systematic comparison of several numerical simulations, and it is also visually explained that if the two categories of interferences (L and θ) and dual uncertain connection weights (σ and λ) are lower than the deduced values ($\overline{L}$, θ_3, $|\sigma|_{sup}$, and $|\lambda|_{sup}$) in Theorem 1, respectively, the disturbed parameter-uncertainty NPRNN based on stable NPRNN can be stabilized again.

Additionally, there is a prospect for future work. On the one hand, future work may focus on the other diverse dynamical stability analysis of NPRNN with connection weight matrices, such as passivity and dissipativity [3, 8], multistability [5], and asynchronous analysis [6, 23]. On the other hand, we may also pay attention to attaching more exogenous interferences on the system established in this paper, such as mixed time-varying delays [3], stochastic disturbances [8], and Markov switching [23]. Furthermore, the uncertain-parameter NPRNN can be extended to more multidimensional spaces, such as fractional-order systems [1] and high-order systems [4].

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors thank everyone who helped for this paper. Additionally, the authors express their thanks to their school and academy for the fund support. This work was supported by the Natural Science Foundation of China, under Grants 62072164 and 12074111, the Directive Project of Science and Technology Research Plan of Hubei Education Department in 2020, under Grant B2020127, and the Project of Hubei Provincial Education Department in 2021, under Grant BXLBX0536.

References

[1] Y. Xu, J. Yu, W. Li, and J. Feng, "Global asymptotic stability of fractional-order competitive neural networks with multiple time-varying-delay links," *Applied Mathematics and Computation*, vol. 389, Article ID 125498, 2021.

[2] Z. Zhang, X. Yang, X. Deng, and L. Li, "A varying-gain recurrent neural-network with super exponential convergence rate for solving nonlinear time-varying systems," *Neurocomputing*, vol. 351, pp. 10–18, 2019.

[3] R. Saravanakumar, G. Rajchakit, C. K. Ahn, and H. R. Karimi, "Exponential stability, passivity, and dissipativity analysis of generalized neural networks with mixed time-varying delays," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 2, pp. 395–405, 2019.

[4] Z. Dong, X. Zhang, and X. Wang, "Global exponential stability of discrete-time higher-order Cohen-Grossberg neural networks with time-varying delays, connection weights and impulses," *Journal of the Franklin Institute*, vol. 358, no. 11, pp. 5931–5950, 2021.

[5] P. Liu, Z. Zeng, and J. Wang, "Multistability of recurrent neural networks with nonmonotonic activation functions and unbounded time-varying delays," *IEEE Transactions on*

Neural Networks and Learning Systems, vol. 29, no. 7, pp. 3000–3010, 2018.

[6] J.-J. Xiong, G.-B. Zhang, J.-X. Wang, and T.-H. Yan, "Improved sliding mode control for finite-time synchronization of nonidentical delayed recurrent neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 6, pp. 2209–2216, 2020.

[7] Y. Zhou, J. Xia, H. Shen, J. Zhou, and Z. Wang, "Extended dissipative learning of time-delay recurrent neural networks," *Journal of the Franklin Institute*, vol. 356, no. 15, pp. 8745–8769, 2019.

[8] R. Saravanakumar, H. Mukaidani, and P. Muthukumar, "Extended dissipative state estimation of delayed stochastic neural networks," *Neurocomputing*, vol. 406, pp. 244–252, 2020.

[9] G. Bao, Y. Peng, X. Zhou, and S. Gong, "Region stability and stabilization of recurrent neural network with parameter disturbances," *Neural Processing Letters*, vol. 52, no. 3, pp. 2175–2188, 2020.

[10] J. Li, M. Hu, and L. Guo, "Exponential stability of stochastic memristor-based recurrent neural networks with time-varying delays," *Neurocomputing*, vol. 138, pp. 92–98, 2014.

[11] W. Jiang, L. Li, Z. Tu, and Y. Feng, "Exponential Lagrange stability for impulses in discrete-time delayed recurrent neural networks," *International Journal of Systems Science*, vol. 50, no. 1, pp. 50–59, 2019.

[12] Q. Chen, L. Liu, and A. Wu, "Mean-square global exponential stability in Lagrange sense for delayed recurrent neural networks with Markovian switching," *Neurocomputing*, vol. 226, pp. 58–65, 2017.

[13] Y. Jun Wang and J. Wang, "Robustness analysis of global exponential stability of recurrent neural networks in the presence of time delays and random disturbances," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 1, pp. 87–96, 2012.

[14] M. U. Akhmet, D. Aruğaslan, and E. Yılmaz, "Stability analysis of recurrent neural networks with piecewise constant argument of generalized type," *Neural Networks*, vol. 23, no. 7, pp. 805–811, 2010.

[15] J.-E. Zhang, "Robustness analysis of global exponential stability of nonlinear systems with deviating argument and stochastic disturbance," *IEEE Access*, vol. 5, pp. 13446–13454, 2017.

[16] Y. He, Q.-G. Wang, and W.-X. Zheng, "Global robust stability for delayed neural networks with polytopic type uncertainties," *Chaos, Solitons & Fractals*, vol. 26, no. 5, pp. 1349–1354, 2005.

[17] V. Singh, "Global robust stability of delayed neural networks: estimating upper limit of norm of delayed connection weight matrix," *Chaos, Solitons & Fractals*, vol. 32, no. 1, pp. 259–263, 2007.

[18] S. Zhu and Y. Shen, "Robustness analysis for connection weight matrix of global exponential stability recurrent neural networks," *Neurocomputing*, vol. 101, pp. 370–374, 2013.

[19] S. Zhu and Y. Shen, "Robustness analysis for connection weight matrices of global exponential stability of stochastic recurrent neural networks," *Neural Networks*, vol. 38, pp. 17–22, 2013.

[20] S. Zhu and Y. Shen, "Robustness analysis for connection weight matrices of global exponential stable time varying delayed recurrent neural networks," *Neurocomputing*, vol. 113, pp. 220–226, 2013.

[21] S. Zhu, K. Zhong, and Y. Zhang, "Robustness analysis for parameter matrices of global exponential stable stochastic time varying delay systems," *Communications in Nonlinear Science and Numerical Simulation*, vol. 19, no. 1, pp. 128–138, 2014.

[22] W. X. Si, T. Xie, and B. W. Li, "Exploration on robustness of exponentially global stability of recurrent neural networks with neutral terms and generalized piecewise constant arguments," *Discrete Dynamics in Nature and Society*, vol. 2021, Article ID 9941881, 13 pages, 2021.

[23] J. Cheng, Y. Wu, L. Xiong, J. Cao, and J. H. Park, "Resilient asynchronous state estimation of Markov switching neural networks: a hierarchical structure approach," *Neural Networks*, vol. 135, pp. 29–37, 2021.

[24] A. Bellen, N. Guglielmi, and M. Zennaro, "On the contractivity and asymptotic stability of systems of delay differential equations of neutral type," *Bit Numerical Mathematics*, vol. 39, no. 1, pp. 1–24, 1999.

[25] M. Hui, N. Luo, Q. Wu, R. Yao, and L. Bai, "New results of finite-time synchronization via piecewise control for memristive Cohen-Grossberg neural networks with time-varying delays," *IEEE Access*, vol. 7, pp. 79173–79185, 2019.

[26] M. Song, Y. Geng, and M. Liu, "Stability equivalence among stochastic differential equations and stochastic differential equations with piecewise continuous arguments and corresponding Euler-Maruyama methods," *Applied Mathematics and Computation*, vol. 400, Article ID 125813, 2021.

[27] D. Peaucelle, D. Arzelier, O. Bachelier, and J. Bernussou, "A new robust D-stability condition for real convex polytopic uncertainty," *Systems & Control Letters*, vol. 40, no. 1, pp. 21–30, 2000.

[28] Y. Q. Xia and Y. M. Jia, "Robust control of state delayed systems with polytopic type uncertainties via parameter-dependent Lyapunov functionals," *Systems & Control Letters*, vol. 50, no. 3, pp. 183–193, 2003.

[29] J. Olejnik and A. Olejnik, "QML estimation with non-summable weight matrices," *Journal of Geographical Systems*, vol. 22, pp. 469–495, 2020.

[30] F. Corpas-Burgos and M. A. Martinez-Beneito, "On the use of adaptive spatial weight matrices from disease mapping multivariate analyses," *Stochastic Environmental Research and Risk Assessment*, vol. 34, pp. 531–544, 2020.

[31] J. S. Zhang, J. Y. Hu, and J. M. Liu, "Neural network with multiple connection weights," *Pattern Recognition*, vol. 107, Article ID 107481, 2020.

[32] N. Rathi, P. Panda, and K. Roy, "STDP-based pruning of connections and weight quantization in spiking neural networks for energy-efficient recognition," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 38, no. 4, pp. 668–677, 2019.

[33] W. J. Lv, Z. W. Yang, and M. Z. Liu, "Stability of the Euler-Maclaurin methods for neutral differential equations with piecewise continuous arguments," *Applied Mathematics and Computation*, vol. 186, no. 2, pp. 1480–1487, 2007.

[34] C. J. Zhang and Y. Y. He, "The extended one-leg methods for nonlinear neutral delay-integrodifferential equations," *Applied Numerical Mathematics*, vol. 59, no. 6, pp. 1409–1418, 2009.

[35] C. Li and C. G. Zhang, "Block boundary value methods applied to functional differential equations with piecewise continuous arguments," *Applied Numerical Mathematics*, vol. 115, pp. 214–224, 2017.

[36] C. J. Zhang, W. S. Wang, B. C. Liu, and T. T. Qin, "A multidomain Legendre spectral collocation method for nonlinear neutral equations with piecewise continuous argument," *International Journal of Computer Mathematics*, vol. 95, no. 12, pp. 2419–2432, 2018.

[37] C. J. Zhang, C. Li, and J. Y. Jiang, "Extended block boundary value methods for neutral equations with piecewise constant argument," *Applied Numerical Mathematics*, vol. 150, pp. 182–193, 2020.

Hindawi
Complexity
Volume 2021, Article ID 6981719, 13 pages
https://doi.org/10.1155/2021/6981719

WILEY | Hindawi

Research Article

Spectrum Handoff Aware AODV Routing Protocol for Cognitive Radio Vehicular Ad Hoc Networks

Abdur Rashid Sangi,[1] **Mohammed S Alkatheiri**,[2] **Satish Anamalamudi**,[3] **Mohammed A. Alqarni**,[4] **Muhammad Hammad Memon**,[1] **and Wanan Yang**[1]

[1]*School of Artificial Intelligence and Big Data, Yibin University, Yibin 644000, Sichuan, China*
[2]*Department of Cybersecurity, College of Computer Science & Engineering, University of Jeddah, Jeddah, Saudi Arabia*
[3]*Department of Computer Science Engineering, SRM University, Amaravati, Andhra Pradesh, India*
[4]*Department of Software Engineering, College of Computer Science and Engineering, University of Jeddah, Jeddah, Saudi Arabia*

Correspondence should be addressed to Abdur Rashid Sangi; sangi_bahrian@yahoo.com

Received 24 May 2021; Revised 11 July 2021; Accepted 13 August 2021; Published 15 September 2021

Academic Editor: Yue Song

End-to-end application performance and throughput of vehicular cognitive transport control protocol depend on how efficiently the segments (TCP header) are being transmitted from source to destination. One way to enhance the performance of vehicular cognitive TCP protocol is by reducing the packet drops in between the source and destination. In general, packet drops occur in between the source and destination of Cognitive Radio Vehicular Ad Hoc Network (CR-VANET) because of spectrum handoff or cognitive node handoff, or network congestion. In this paper, we focus on enhancing the performance of cognitive TCP protocol through "cognitive AODV routing protocol with the spectrum handoff mechanism." In the proposed work, channel-route control messages of cognitive AODV routing protocol are updated with the support of spectrum handoff which helps to provide the backup opportunistic channel during PU active and helps to reduce the end-to-end spectrum handoff packet drops. Simulation results reveal that the overall performance of the vehicular cognitive TCP protocol with the proposed spectrum handoff aware cognitive AODV routing protocol is enhanced as compared to the existing cognitive TCP protocol.

1. Introduction

With enhanced wireless technology, a wide range of applications are being deployed and operated in current static unlicensed ISM spectrum bands. This results in spectrum scarcity and increased spectral congestion. On the contrary, FCC (Federal Communication Commission) revealed that 15% to 85% in the current static allocated spectrum bands below 3 GHz are not efficiently utilized concerning temporal and spatial domains [1–4]. This clearly shows that the current static spectrum allocation is not utilized optimally rather than physical spectrum scarcity. One way to overcome this issue is by utilizing the temporal and spatial domain unused spectrum bands through dynamic spectrum access. With dynamic spectrum access (DSA), the existing spectrum scarcity problem will be alleviated as well as it increases the spectrum utilization. These result in developing an adaptive wireless device called cognitive radio together with the introduction of dynamic spectrum access through secondary spectrum licensing called "CRN (Cognitive Radio Networking)." With CRN, cognitive radio devices are capable of accessing the spatial spectrum dynamically through spectrum holes or white spaces [5–7]. Moreover, a CR node with "SDR (Software Defined Radio)" is capable of switching in a wide range of spectrum bands (spectrum hole) through available RSSI (signal strength), transmission quality, and dynamic PU activity. In general, Cognitive Radio Networks can be deployed in one of the two architectures, namely, "infrastructure-based CRN" and "ad-hoc CRN networks," which are also called "CRAHNs (Cognitive Radio Ad Hoc Networks)" [8–12]. Recent advancement of wireless applications in vehicular networks leads to spectrum scarcity [4, 5] which can be overcome through the CR-VANETs (Cognitive Radio Vehicular Ad Hoc Networks).

This paper mainly focuses on the performance enhancement of "vehicular CRAHNS" by reducing the packet drops due to spectrum handoff and mobile handoff of the vehicular nodes. Currently, the main research focus in cognitive radio networking is well explored in the areas of channel estimation, spectrum sensing, spectrum sharing, cognitive routing, and MAC (medium access control). In general, the list of PU-free channels in between one-hop CR nodes will be managed by the "connection management" [13–15]. During the route discovery at the network layer, the routing protocol has to concurrently select the PU-free channel (spectrum hole) and next-hop route to transmit the application data. Thus, there should be tight coordination between "connection management" and "routing tables" during the end-to-end channel-route discovery. In [16], the design of cognitive AODV routing protocol is being proposed to concurrently select the channel along with the route during the route discovery. Later, an end-to-end channel route is identified by exchanging the cognitive control message with a "common control channel" for application data transmission. In static spectrum-allocation-based mobile ad-hoc networks, the application data packets in between the source to destination CR node can be dropped due to bandwidth reduction or mobile handoff. When it comes to Cognitive Radio Ad Hoc Networks, there is a new type of packet drop that occurs because of spectrum handoff from the current spectrum hole to another spectrum hole at the time of PU active in the current CR communication channel [17–19]. State-of-the-art research proposals are proposed and designed to reduce the packet drops due to CR node handoff and bandwidth degradation. But, very few research activities are done to explore the packet buffering mechanisms during spectrum handoff. When it comes to dynamic spectrum access, it is crucial to reduce the application packet drops during spectrum handover to achieve enhanced end-to-end application throughput. To accomplish this, this paper proposes a "packet buffering and forwarding mechanism" to the "cognitive AODV routing protocol." With this, the end-to-end achievable TCP throughput will be enhanced by reducing on-the-fly packet drops and reducing the channel-route reconstruction time during spectrum handover at the intermediate CR node. Our paper is organized as follows. Section 2 briefly explains the advantages and disadvantages of the existing CRAHNs without handoff mechanisms. Section 3 describes the proposed "spectrum handoff" aware cognitive AODV protocol through packet buffering and forwarding mechanism to enhance the performance of end-to-end connection-oriented cognitive TCP protocol. Besides, a RED (Random Early Detection) buffer is proposed to use in our solution to minimize the packet drops during packet forwarding in between intermediate CR nodes. Section 4 explains the simulation results collected by simulating in the Cognitive Radio Network Simulator (CRNS.2.35) and compares the performance of existing cognitive TCP protocols with proposed simulation results with spectrum handoff packet buffering and forwarding mechanism. Section 5 discusses the conclusion with future work.

2. Motivation

In standalone wired or wireless networks, application packet drops are mainly due to increased traffic flow (buffer overflow) that results in network congested packet drops. In addition, packet drops may occur at the end nodes or intermediate nodes due to errors within the packet. Extensive research is carried out to minimize the packet drops due to buffer overflow and errors in the packet. Later on, once the mobility support is deployed within the wireless networks, then node mobility packet drops are increased during node handover from one base station to another base station. To overcome this, many researchers focus on enhancing the performance of the end-to-end packet transmission by introducing the node mobility handover mechanisms. But, the node mobility handover mechanisms cannot be directly deployed in cognitive radio vehicular networks due to opportunistic channel utilization and high-speed node mobility. Hence, a vehicular node should aware of the opportunistic PU-free channel along with the node mobility to integrate the spectrum handover and node mobility in real-time vehicular communication. To achieve this, our work proposes a spectrum handoff-aware routing protocol that considers robust dynamic PU channel selection along with node mobility support at the time of discovering the end-to-end channel route from the source to destination vehicular nodes.

3. Related Work

The end-to-end connection-oriented application throughput mainly depends on the reliability of the channel routes that are being discovered through the routing protocol. Hence, the cognitive routing protocol in the network layer plays a critical role in enhancing the end-to-end throughput of the cognitive TCP protocol.

In general, MAC protocols are broadly classified into contention-based (random access) and contention-free access protocols to schedule the one-hop data transmission. Traditional random access-based MAC protocols make use of CSMA/CA for accessing the one-hop channel whereas MAC contention-free protocols use either predetermined global or local time-slotted-based channel access. With CSMA/CA, mobile nodes at the MAC layer will contend for the shared channel (nonoverlapping channel) to transmit the application data, whereas virtual carrier sensing with NAV (Network Allocation Vector) helps to avoid the nodes' hidden terminal problem at the sensing range. Random access-based MAC protocols are further classified into RTS/CTS-based and tone-based [6] to alleviate the hidden and exposed terminal issues. But, when it comes to cognitive radio networks, there is no single global unlicensed channel (2.4 GHz/5 GHz ISM bands) available to assign as a control channel for cognitive control message exchange. To overcome this, in [20–22], the authors proposed a hybrid CCC-based MAC and routing protocols for channel-route control message exchange at the network layer and one-hop scheduling operations at the MAC layer of Cognitive Radio Ad Hoc Networks. An overall summarization of well-known routing protocols is shown in Table 1.

The comparison was made with different state-of-the art routing protocols in VANET and Cognitive Radio Ad Hoc Networks. To date, most of the existing solutions focus on

Table 1: Summarization of routing protocols.

Routing protocol	Routing metric	Channel selection	Energy conservation	CCC
OBPF [23]	Link quality	Single channel	No	No
SE-AOMDV [24]	Hop count	Single nonoverlapping channel	No	No
CR-AODV [16]	Optimized channel route	Single channel	Yes	Yes
ERMR [25]	Energy and stability	Channel availability time	Yes	Dedicated
MRPC [26]	Hop count	Stability	No	Dedicated
PMRC [27]	Route path	Single channel	No	Dedicated

channel-route discovery at the network layer for application data transmission. But, there is little research focused on how to handle the spectrum handover during application data transmission. Besides, end-to-end packet delays in cognitive routing protocol and CR node energy consumption mainly depend on the PU activity (spectrum handoff), per-hop channel saturation, and CR node handoff. Moreover, the number of simultaneous noninterfering transmissions among the neighbor CR nodes will have a direct impact on achievable end-to-end network throughput in multihop cognitive communication [28–31]. Furthermore, CCC (Common Control Channel) [20, 32] plays a vital role in discovering the end-to-end channel-route path for application data transmission. In common mobile IP-based networks, the end-to-end path failure can be mainly due to node mobility from one subnetwork to another subnetwork. When it comes to Cognitive Radio Ad Hoc Networks, a new path failure in between the source to destination CR node can happen due to spectrum handoff in PU transmission channels. As shown in Figure 1, whenever the PU (primary user) is active in the current CR node communication channel, then on-the-fly packets have to be buffered until a new one-hop channel route is being locally discovered. Cutting-edge routing protocols are being proposed to provide the shortest end-to-end channel route between the source and destination CR node. But, whenever there is a spectrum handoff, there is no existing solution to handle on-the-fly packet drops until a new local channel route is being reconstructed. Also, there should be a tight synchronization between the network layer and the spectrum management to check the common PU-free channels available between the intermediate CR nodes. Apart from this, when the intermediate CR node is unable to find the backup PU-free channel at the time of spectrum handoff, then it generates the channel RERR (route-error) message and also unicasts back to the source CR node.

Once the RERR message arrives back at the Source CR node, then it will re-initiate the channel-route discovery through Route-Request and Route-Response to handle the spectrum handover connection failure. An enormous amount of research is well focusing on solution designing for channel route failures due to the spectrum handover with efficient PU-free back up the channel to continue the ongoing transmission. In [21, 22], the authors mainly focus on the PU channel selection at the gateway devices to integrate multiple unlicensed radio access networks. But, the node mobility is not integrated with PU channel selection which results in packet drops during the active primary user in the ongoing CR vehicular communication. In [29, 30], the authors focus on ways to enhance PU-free channel utilization due to the proliferation of radio devices and the heterogeneous environment of wireless communications. In their contribution, the SU channel switching rate is considered when the primary user is active in the current CR communication. But, there is no mechanism explored about how the SU channel switching has to be integrated along with the node mobility and the traffic characteristics to minimize the application packet drops. To date, there is no research focus on how to reduce on-the-fly packet drops through packet buffering and packet forwarding mechanisms during cognitive channel switching and CR node mobility. In this paper, a RED-based packet buffering and forwarding mechanism is being proposed to minimize on-the-fly packet drops to enhance the performance of cognitive TCP protocol with respect to spectrum mobility and CR node mobility. The main contribution of this paper is threefold: (1) packet buffering and forwarding mechanism to buffer on-the-fly spectrum handover packet drops, (2) modified CR-AODV protocol to record the PCL (PU-free channel list) during the channel-route discovery, and (3) mechanism to trigger the spectrum management with the available PU-free channel list to buffer the packets during spectrum handover.

4. The Proposed Scheme

Spectrum handoff-aware cognitive AODV routing protocol works under the half-duplex radio transceiver, i.e., a secondary radio device can either receive or transmit at any given time through the opportunistic PU-free spectrum bands. In traditional IEEE 802.11 ad-hoc networks, the network designer has to mainly focus on how to efficiently utilize the nonoverlapping spectrum in either 2.4 GHz or 5 GHz ISM bands. Thus, a PLRR (Preemptive Local Route Repair) is being proposed to locally reconstruct the link during end-to-end application data transmission. With this, the achievable TCP throughput in traditional ad-hoc networks can be improved by locally re-constructing the routes. Initially, when a sender/source node has application data, then it has to check for the end-to-end channel route. If the channel route is unavailable within the routing table of the sender/source CR node, then it has to initiate the RREQ broadcast message.

Since there is no fixed static channel available in Cognitive Radio Ad Hoc Networks, it is a challenging task to broadcast the route discovery control messages from the source to destination CR node. In other words, with decentralized cognitive ad-hoc radio networks, it is clear that

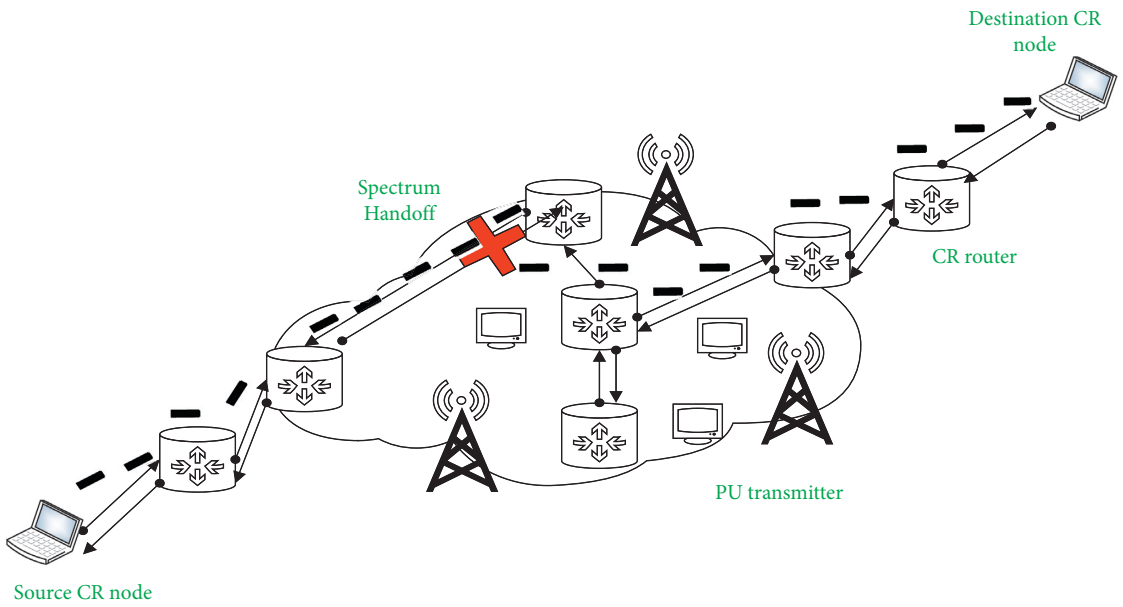


Figure 1: Spectrum handoff in between intermediate CR nodes due to PU active.

any CR node synchronization with other nearby CR nodes for cognitive control exchange is a huge challenge. Moreover, the AODV routing protocol with dynamic spectrum access has certain limitations to select the end-to-end channel-route in between the source and destination CR node. This is due to increased route-channel control overhead, longer channel rendezvous end-to-end delays, and increased collision probability because of the multichannel hidden terminal issue. Hence, hybrid CCC (Common Control Channel) is being proposed to transmit the cognitive and route control messages among source and destination CR nodes. As shown in Figure 2, hybrid CCC makes use of 902 MHz ISM band for Time "*T*-1" to exchange its PU-free channel list to one-hop neighbor CR nodes. Subsequently, at *T*-2, the source CR node will start listening to its PU-free channel list to learn about the PU-free channel list of its one-hop away neighbor nodes. With this, the sender/source CR node will get the PCL list of its one-hop away neighbor CR nodes. Later, the source CR node broadcasts the DCRequest in a common PU-free channel to its one-hop away neighbor nodes. CR nodes that have the same PU-free channel will reply with the DCResponse back to the source CR node. With this, the source CR node will have a synchronized common PU-free channel available with its one-hop neighbor CR nodes. Thus, the routing and MAC control messages (RREQ, RTS, CTS, and RREP) will be exchanged in between the source and destination to find out an end-to-end shortest path to transmit the application data. As shown in Figure 2, packet drops can occur at the intermediate CR node due to PU active in the ongoing CR node communication channel. At this point, the per-node packet buffering technique can be used to avoid the application packet drops. This will, in turn, enhance the achievable application throughput at the cognitive-TCP protocol. In this paper, omnidirectional antenna-based simulation is being used for both cognitive control exchange through CCF (Cognitive Coordination Function) and application data transmission through DCF (Distributed Coordination Function), as shown in Figure 2. If the end-to-end channel route does not exist at the source CR node, then it will activate the medium access layer to initiate the CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) for one-hop link-level channel contention to broadcast the channel RREQ (route-discovery messages). Figure 3 shows how the packet buffering mechanism works in our proposed work.

On-the-fly packets will be recorded at the SDR radio interface and buffered at the input RED buffer. When there is no issue with the outgoing link, then the buffered packet in the input buffer will be sent to the output RED buffer. Whenever there is a link failure at the outgoing RED buffer towards the next-hop CR node, then a separate RED packet buffer will be created for the next-hop neighbor CR node. In this paper, spectrum handover in-flight packets from the old CR intermediate node to a new CR intermediate node have to be forwarded through this priority-based RED buffer. At one point in time, RED buffer will have either Avg_quelength < Thresoldmin or Thresoldmin < Avg_quelength < Thresoldmax, or Avg_quelength > Thresoldmax. In mobile ad-hoc networks, a separate packet buffering for mobile handover will be created when the average length of the input RED buffer is in between Minimum_Thresold to Maximum_Thresold or when the average length is greater than the Maximum_Thresold. But for spectrum handover packet drops, a separate spectrum handover buffer will be created for in-flight packets, when the old CR intermediate node handovers its opportunistic channel due to PU active. Thus, the received in-flight spectrum handover packets at the old CR node will be buffered at the spectrum handover packet buffer. It is noteworthy that a separate spectrum

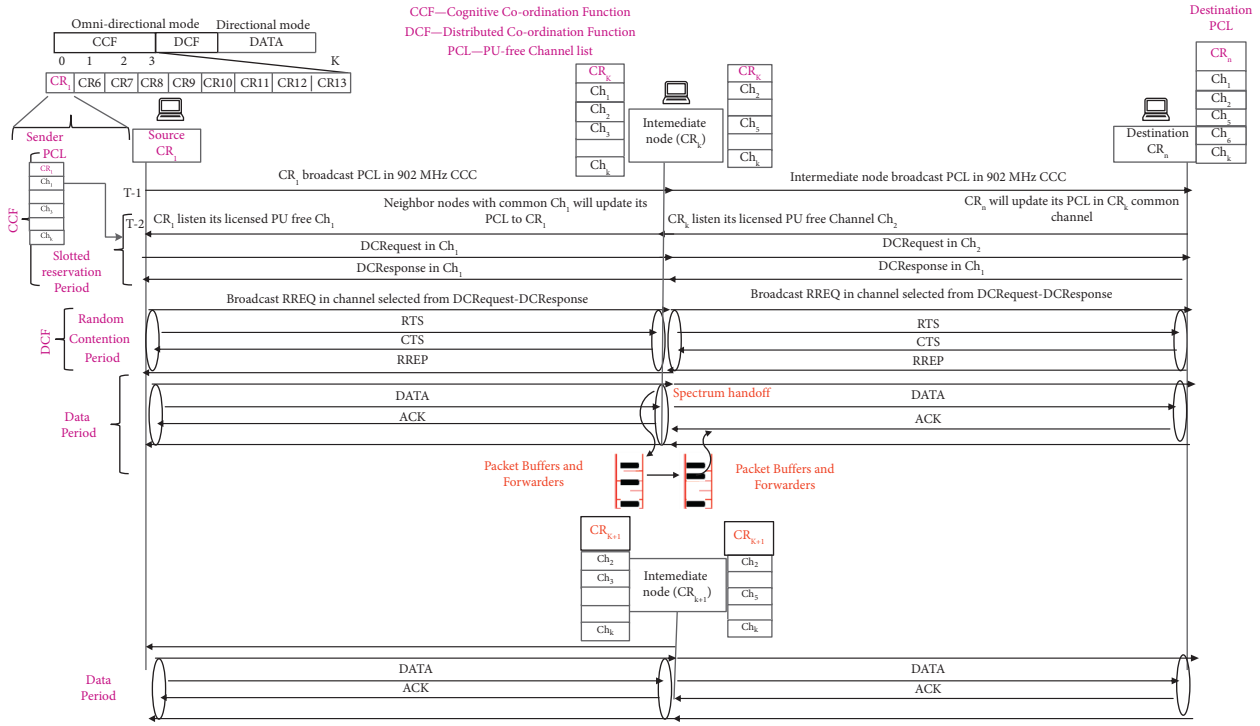


Figure 2: Hybrid CCC-based CR-AODV protocol with spectrum handoff through packet buffering and forwarding.

handover packet buffer will be created for every next-hop CR node that has spectrum handover link failure.

Algorithm 1 explains the packet buffering operation when the average length of the RED buffer is less than the predefined threshold value. Once there is a spectrum handover, then the old CR node will check for the local channel-link re-construction through PLRR (Preemptive Local Route Repair). Later, spectrum handover buffer packets will be retransmitted from the spectrum handover buffer to the next-hop CR node (new CR node) through the outgoing RED buffer. In this case, once the link is locally reconstructed, then the spectrum handover buffer packets are transmitted like normal best-effort IP traffic to the next-hop CR node (new CR node). But there are cases where the RED buffers can get overflow through network congestion. This usually happens due to the store-and-forward mechanism of multiple traffic flows. In this case, whenever there is network congestion, then the outgoing RED buffer will start dropping the packets randomly at the output RED buffer. This in turn results in spectrum handover packet drops even with the packet buffers. To overcome this, a priority-based spectrum packet buffering and forwarding mechanism is being proposed to avoid the spectrum handover packet drops during the network congestion at the output RED buffer. Figure 4 describes how to add a priority option to the spectrum handover buffer packets at the output RED buffer. As shown in Figure 4, a one-bit priority will be added to the spectrum handover packet whenever the average length of the output RED buffer is either Thresoldmin < Avg_que_length < Thresoldmax or Avg_quelength > Thresoldmax. At this point in time, priority packets (spectrum handover packets enabled with priority bit $P = 1$) cannot be discarded at the output RED buffer of the old CR node during the network congestion. Whenever the average length of the output RED buffer at the old CR node is Avg_que-length < Thresoldmin, then the priority bit is disabled (priority bit $P = 0$) and packets will be transmitted like normal IP packets. With this, the probability of spectrum handover buffer packet drops will be reduced especially when there is network congestion due to increased best-effort traffic flows at the output RED buffer of the old CR node. This helps to engage the opportunistic communication link to transmit many application packets than the route-control messages. Thus, the end-to-end application throughput and performance of cognitive TCP protocol with proposed priority packet buffering will be enhanced in comparison with the existing spectrum handover-based packet buffering and forwarding. In the worst case, there may be packet drops at the input RED buffer due to network congestion. Algorithm 2 explains the packet buffering operation when the average length of the RED buffer is in between a minimum threshold and a maximum threshold of a predefined threshold value. For such cases, especially for deterministic applications, a separate priority input buffer will be created along with the normal input buffer at the network layer.

Once the packets with higher priority arrive at the old CR node, then it is buffered within the priority buffer. Later, the CR node checks for the outgoing PU-free channel to forward the buffered priority packet. When there is a spectrum handover, then the buffered priority packet will be forwarded to the spectrum handover buffer and waited until the new one-hop link is established from the current CR node. It is noteworthy that the buffered spectrum handover

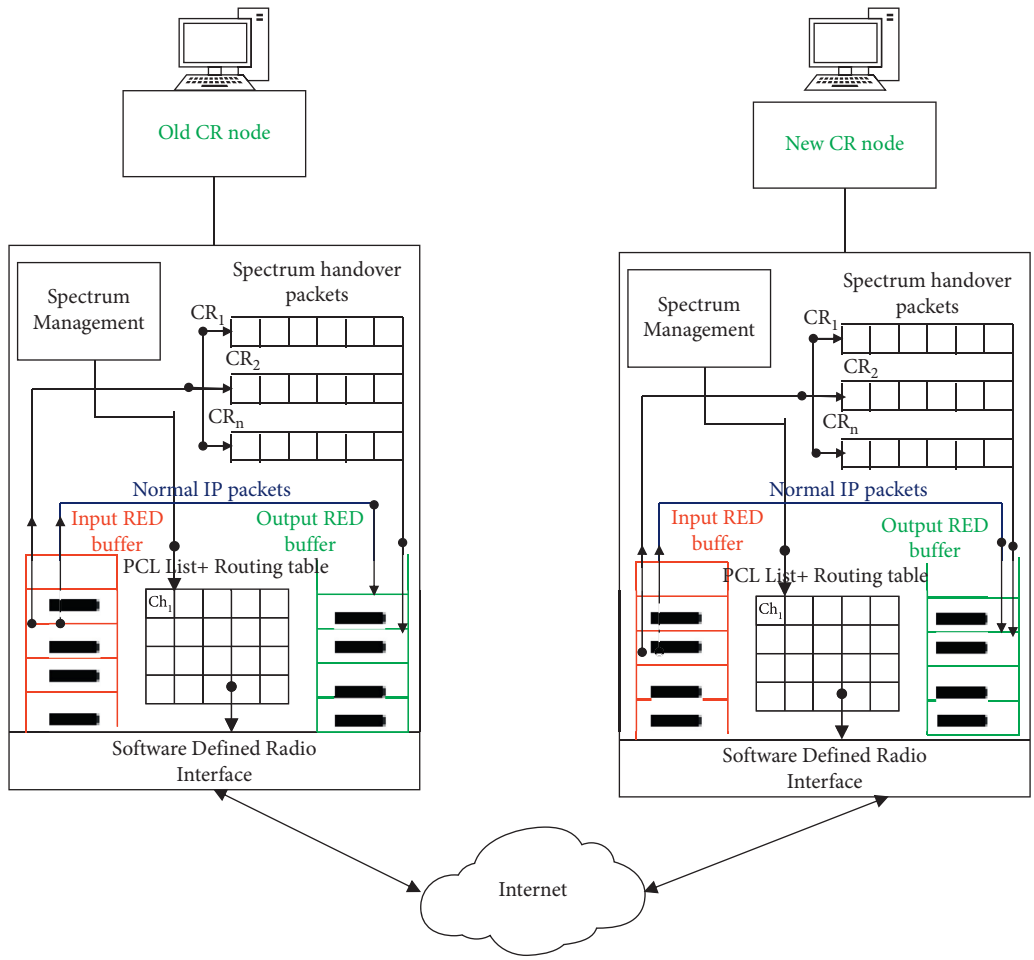


Figure 3: Spectrum handoff packet buffering at the intermediate CR node in CRAHNs.

packet will be forwarded with either priority bit as "1" or "0" based on the average length of the output RED buffer. Once the packet arrives at the next hop intermediate CR node (new CR node), then it will check the status of the input RED buffer and decide whether it has to be buffered in the input RED buffer or priority input buffer.

4.1. End-to-End Channel-Route Maintenance. Due to "dynamic opportunistic spectrum access," there will be frequent link failures due to active licensed primary users. In widely used MANETs (mobile ad-hoc networks), there is a control message called "RERR (Route Error)" that triggers when the outgoing link cannot be locally reconstructed. This usually happens when there is network congestion of mobile handoff in traditional mobile ad-hoc networks. But, in Cognitive Radio Ad Hoc Networks, there is an additional packet drop due to channel switching at the time of PU active. Hence, the RERR (Route Error) message should be updated with the support of the "spectrum handover packet drops" to the source CR node. In [21], the authors described a new control message format to support the spectrum handover packet drops at an intermediate CR node to the sender/source CR node. When the link cannot be locally reconstructed for more than an RTT (Round Trip Time), then this intermediate CR node has to re-initiate the "RERR (Route Error)" message back to the source CR node. Thus, the RERR (Route Error) message with error bit is set to "11" at the time of transmitting from the intermediate CR node to the source CR node [21]. Once, the source CR node receives the "Route Error" message with the "Spectrum handover" link failure, then it will start to look for an existing alternate route. When there is no existing alternate route, then it initiates the route discovery message to determine the fresh channel route from source towards destination CR nodes.

4.2. Spectrum Handover Queuing Model. Whenever the average queue length of output RED buffer at intermediate CR node is Avg_Que_length_CR < thresholdmin, then there is no need of having priority bit enabled at the spectrum handover packet buffers. Whenever the average queue length of the output RED buffer is in between Thresoldmin < Avg_que length < Thresoldmax and Avg_que-length > Thresoldmax, then the spectrum handover packet buffers will be enabled with priority bit at the spectrum handover buffers. Hence, it is crucial to know the average length of the output RED buffer through a mathematical model to enable the priority bit for congested output RED buffer. An M/M/1 queuing model was designed to figure out

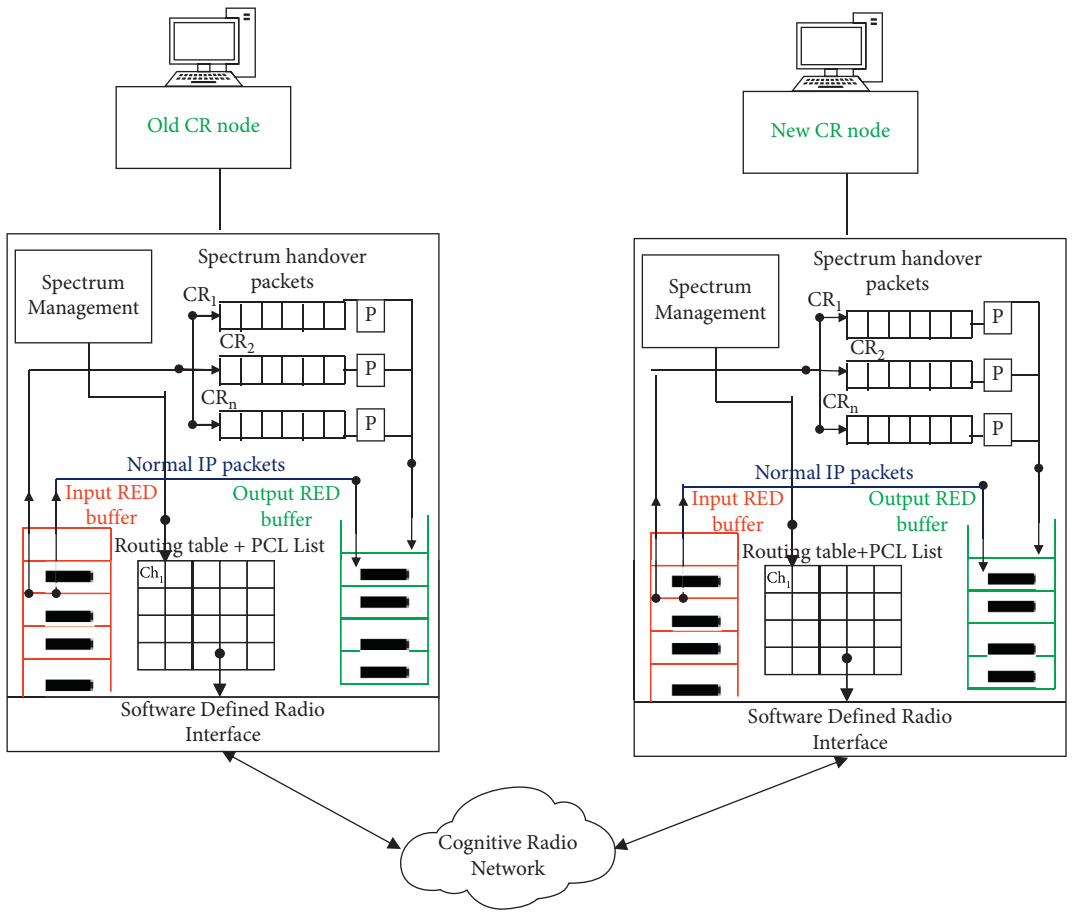


Figure 4: Spectrum handoff packet buffering with priority buffers at the intermediate CR node in CRAHNs.

```
Spectrum handoff based CR-AODV {
  while (Avg_Queue_length_old-CR-node < RED_minthresold) {
    Priority_bit = 0; //packet with no priority.
    RED_priority = Fa1se; //No need to create Priority buffer.
    Marker_priority = 0; //packets with no priority
    REDInput ()//InputRED buffer at new-CR-node. {
      REDInput = Packets_Non−priority + NormalIP;
      CongestionNotification = 0;
      Commonchannel_old node−new node = True;
    }
  }
}
```

Algorithm 1: //Spectrum handoff packet buffering at intermediate CR nodes//

the average length of the output RED buffer at a certain point in time. Based on the analysis of the mathematical model, a priority bit will be decided to enable/disable at the spectrum handover buffer. In queuing theory, an M/M/1 Queue represents the length of the queue in a system having a single server, and the packet arrival is determined by the Poisson process, and the job service times are based on the exponential distribution. The state space in M/M/1 queue is the set of {0,1,2,3, ...} whose value corresponds to the number of receiving packets (normal IP traffic flows and spectrum handover packets) from one intermediate CR node (old CR node) to another intermediate CR node (new CR node). In this paper, the following assumptions are as follows: packet arrival (both spectrum handover + normal IP packets) rate at the old CR node is "α" from state "P" to "$P + 1$." Service time is defined as 1/€ where "€" is the service rate. The server follows the FCFS (First Come First Serve) basis to receive the packets. When there is no spectrum handover, then the packet leaves the input RED buffer, thus reduced to "$N - 1$" if the total number of the packet is known to be "N." In

```
Spectrum handoff based CR-AODV {
  while (((Avg_queue > minthresold) && (Avg_queue < maxthresold) || (Avg_queue > maxthresold)) {
    Priority_bit = 1; //Priority packet.
    RED_priority = True; //create Priority buffer.
    Marker_priority = 1; //priority packets
    REDInput ()//InputRED buffer at new-CR-node. {
      RED_Input = Norma1_IP;
      RED_Priority = Packet_spriority;
      if(Avg_queue > maxthresold)
        CongestionNotification = 1;
      else
        CongestionNotification = 0;
        Commonchannel_oldnode-newnode = True;
    }
  }
}
```

Algorithm 2: //Spectrum handoff packet buffering with priority buffers at intermediate CR nodes//

general, it is assumed that the RED buffer is of finite size (N) which is having a limit on the storage of the packets within the queue. The average length (1) of the proposed queuing model is

$$\mathrm{Lq} = \sum_{n=0}^{\infty} (n-1)p_n = \frac{2p - p^2}{1-p} + \frac{p^{N-1}(Np+1)}{1-p^{n-1}}. \tag{1}$$

The detailed derivation of "Average Queue Length (Lq)" is explained in [22]. Here, when $P = 1$, then $\mathrm{Lq} = 0$, i.e., there will not be spectrum handover packet buffering at the input RED buffer. When $P > 1$, then Lq is equal to infinity. Hence, we only analyze when $P < 1$ to determine the average length of the queue.

4.3. One-Hop Theoretical Throughput Calculation. Theoretical upper bound of one-hop throughput from the old CR node to the new CR node helps to determine the collision rate concerning to the data rate in the PU-free opportunistic shared channel. Furthermore, this helps to predict the congestion window size variation at the cognitive transport layer to improve the end-to-end throughput and overall CRAHN network performance. In theoretical throughput estimation, we assume that the channel is completely available for the CR node to transmit the application data. This helps for an ideal case to figure out the maximum link achievable throughput without any collisions. In experimental results, the PU active in current CR communication is considered to show how the throughput and performance of the TCP vary with respect to PU active:

$$\begin{aligned} \mathrm{Throughput}_{\mathrm{per-hop}} &= \frac{T_{\mathrm{data}}}{\mathrm{Delay}}, \\ \mathrm{Throughput}_{\mathrm{per-hop}} &= \frac{T_{\mathrm{data}}}{T_{\mathrm{DCF}} + T_{\mathrm{PCL}}}, \end{aligned} \tag{2}$$

where

T_{PCL} = Time to select the PU channel,

$$\begin{aligned} T_{\mathrm{DCF}} &= T_{\mathrm{RTS}} + T_{\mathrm{SIFS}} + T_{\mathrm{CTS}} + T_{\mathrm{SIFS}} + T_{\mathrm{DATA}} + T_{\mathrm{SIFS}} + T_{\mathrm{ACK}} + T_{\mathrm{DIFS}} + T_{\mathrm{BACKOFF}}, \\ T_{\mathrm{DCF}} &= T_{\mathrm{RTS}} + T_{\mathrm{SIFS}} + T_{\mathrm{CTS}} + T_{\mathrm{SIFS}} + \left(T_{\mathrm{PR}} + T_{\mathrm{PHY}} + \frac{8(L_{\mathrm{MAC}} + \mathrm{MSDU})}{\mathrm{Data}_{\mathrm{Rate}}}\right) \\ &\quad + T_{\mathrm{SIFS}} + \left(T_{\mathrm{PR}} + T_{\mathrm{PHY}} + \frac{8(L_{\mathrm{MAC}})}{\mathrm{Data}_{\mathrm{Rate}}}\right) + 2 * T_{\mathrm{slottime}} + T_{\mathrm{SIFS}} + \frac{Cw_{\min} * T_{\mathrm{slottime}}}{2}, \end{aligned} \tag{3}$$

$$T_{\mathrm{PCL}} = T_{\mathrm{PR}} + T_{\mathrm{PHY}} + \frac{8(n * L_{\mathrm{PCL}} + L_{\mathrm{sender\,IP}} + L_{\mathrm{sender\,MAC}})}{\mathrm{Data\ rate}},$$

$$\mathrm{Throughput}_{\mathrm{end-to-end}} = N * \left(\frac{T_{\mathrm{data}}}{\mathrm{Delay}}\right), \tag{4}$$

where N = total number of hops in between the source to destination CR node.

Equation (4) clearly shows that the achievable end-to-end application throughput directly depends on the route control messages. Whenever there are spectrum handoff packet drops, then the time taken to reconstruct the channel route increases. Thus, the end-to-end delay on data transmission is higher with increased control message transmission.

4.4. Use Case of Proposed CR-AODV Protocol in CR Vehicular Ad Hoc Networks. Packet drops in high-speed vehicular networks, such as automated ambulance, are a significant issue that needs to be addressed in mission-critical applications. In addition, relying on the unlicensed spectrum bands to transmit the application data should also be addressed due to heavy traffic flows by heterogeneous networks in the ISM bands. To overcome both the issues, this paper proposes a cognitive-AODV routing protocol that utilizes an unlicensed spectrum band along with the spectrum mobility and CR vehicular node mobility with RED spectrum mobility priority packet buffers. The proposed cognitive AODV-routing protocol mainly addresses the packet drops with respect to node mobility, spectrum mobility, and aggregate packet drops at the intermediate CR vehicular nodes. The performance of the proposed CR routing protocol is tested with different performance metrics in Sections 5 and 6.

5. Implementation in NS-2 Simulator

CR-NS-2.31 (Cognitive Radio Network Simulator) [18, 33–37] with dynamic channel access support is implemented to analyze the congestion window variation, performance of cognitive TCP, and end-to-end throughput with respect to spectrum handover (PU active) at the intermediate CR nodes in Cognitive Radio Ad Hoc Networks. This work makes use of the 8 MHz "primary TV white space" for opportunistic data transmission among the source CR node and destination CR node within the CRAHNs. Table 2 below describes the parameters used for our simulation. The total simulation area is considered as $1000 * 1000$ meters. Cognitive Ad-hoc on-demand distance vector (Cognitive-AODV) routing protocol is used to select the unreliable end-to-end concurrent channel route from the source CR node to the destination CR node. The total number of CR nodes in between the sender CR node to destination CR node is considered as 100 whereas PU nodes are considered as 10. The cognitive TCP protocol is being used to provide reliable end-to-end connection-oriented traffic flow.

During the simulation, it is assumed that there is no PU transmitter active at the current CR communication. Once the connection setup (CR-TCP) and end-to-end channel-route discovery (CR-AODV) are done, then application data start to transfer from the sender CR node to the destination CR node. At this point in time, the PU transmitter will be active (PU-ON) for either 10 or 15, or 20 seconds. Subsequently, the spectrum handover will get triggered at an intermediate CR node, and on-the-fly spectrum handover packets get dropped at an intermediate CR node with the existing CR routing protocols.

6. Results and Discussions

With the implementation of spectrum handover-based CR-AODV protocol in cognitive network simulator, the packets at the intermediate CR-node get buffered at the spectrum handover buffers and re-transmitted back once the local channel-route is discovered for the spectrum handover link failure. Figure 5(a) describes the CR-TCP congestion window variation with varied simulation times. It is noteworthy that the end-to-end CR throughput and the performance of cognitive-TCP are directly dependent on the congestion window.

Whenever there is a packet drop, then the ACK will not be reached back to the sender TCP within RTT (Round Trip Time). Due to the longer RTT delays, the sender congestion window gets reduced which in turn reflects on the rate of packet transmission. When there is a PU active in current CR communication, then the intermediate CR node drops the spectrum handover packets due to channel-route failure. In Figure 5(a), CR-AODV without spectrum handover buffer is having more packet drops than the intermediate CR nodes, which is why the size of the congestion window is less in comparison with the other two techniques. CR-AODV with only spectrum handover buffer is having a better congestion window size in comparison with the CR-AODV without spectrum handover buffer. This is due to a separate spectrum handover buffer at an intermediate CR node for the next-hop CR node that has a channel-link failure. CR-AODV with spectrum handover buffer and priority bit is having the highest congestion window. This is due to reduced packet drops at the output RED buffer. In other words, whenever the best-effort traffic flow is close to maximun_thresold, then the spectrum handover packets also get dropped along with the normal IP packets. To overcome this, the proposed solution added a priority bit to the spectrum handover buffer packets. With this, the output RED buffer will try to not drop the packet that is having priority bit but will continue to drop the normal IP packets. This results in an increased TCP congestion window in comparison with the other two techniques. Figure 5(b) describes the end-to-end CR-TCP throughput (Mbps) with respect to simulation time. In reality, the end-to-end throughput is directly proportional to the size of the congestion window. With this, cognitive CR-AODV with spectrum handover buffer with priority bit is having better achievable throughput in comparison with CR-AODV with only spectrum handover buffer and CR-AODV without spectrum handover buffer. Figure 5(c) describes the local channel-route discovery with different no. of CR nodes. With a higher number of one-hop CR nodes, there will be increased traffic flow at the input RED buffer and output RED buffer. This results in network congestion and increased packet drops at the input RED buffer and output RED buffer. Since the proposed CR-AODV with spectrum handover buffer is having priority bit enabled, the output

Table 2: Simulation parameters.

Parameter name	Description
Topology	1000 ∗ 1000 flatgrid
No. of CR nodes	100
No. of PU channels	10 (8 MHz) channels
No. of PU transmitters	10
PU active probability	10, 15, and 20 seconds
Mobility model	Random waypoint model
Receiver threshold	−95 dBm
Carrier sense threshold	−115 dBm
Input CR transmit power	10 mW
CR Tx range	200 m, licensed channel
PU Tx range	500 m, licensed channel
Data rate	2 Mbps
Interface queue length	50
Simulation time (s)	200 seconds
Traffic type	TCP
Packet size	512 and 1024 bytes

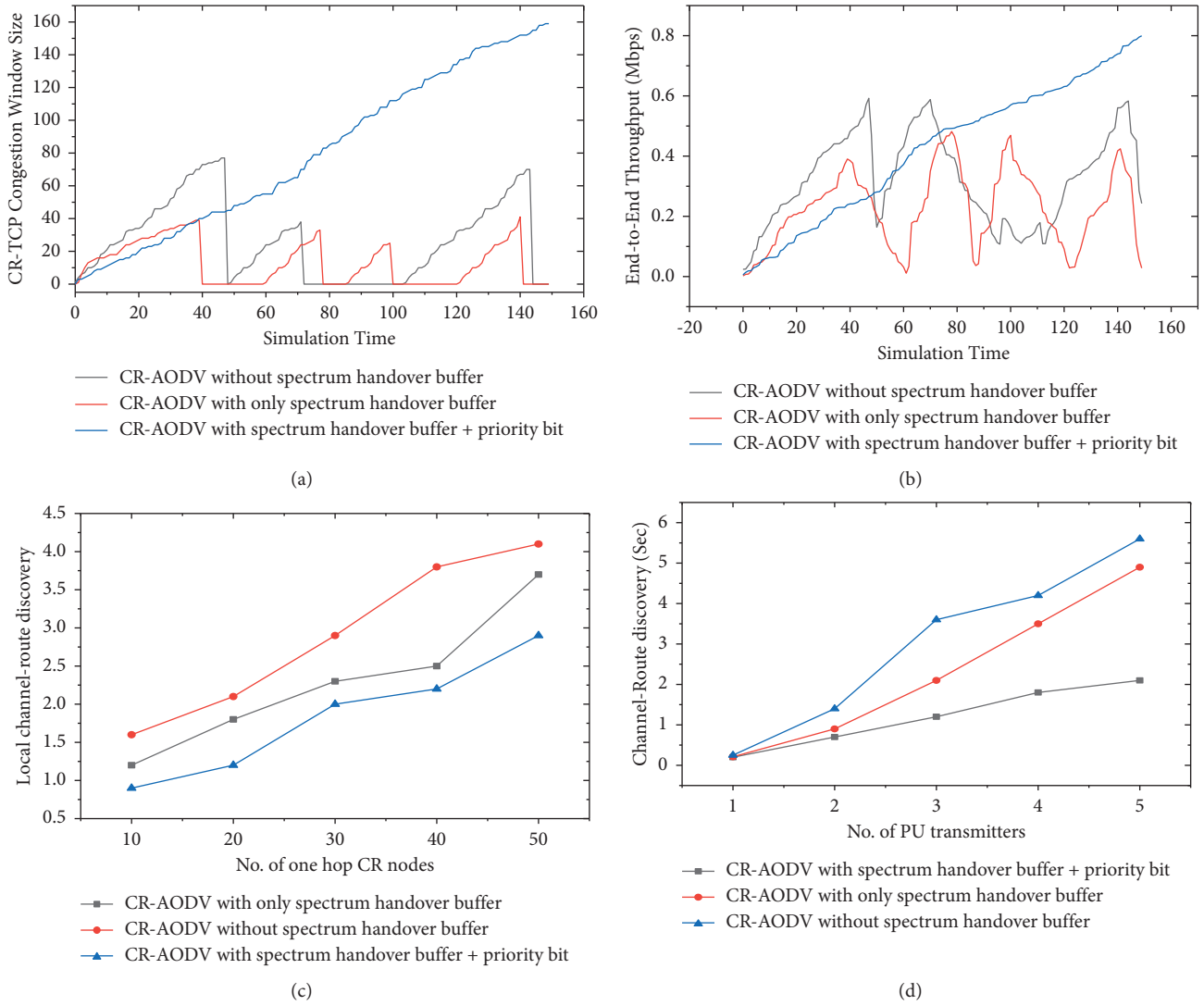


Figure 5: Continued.

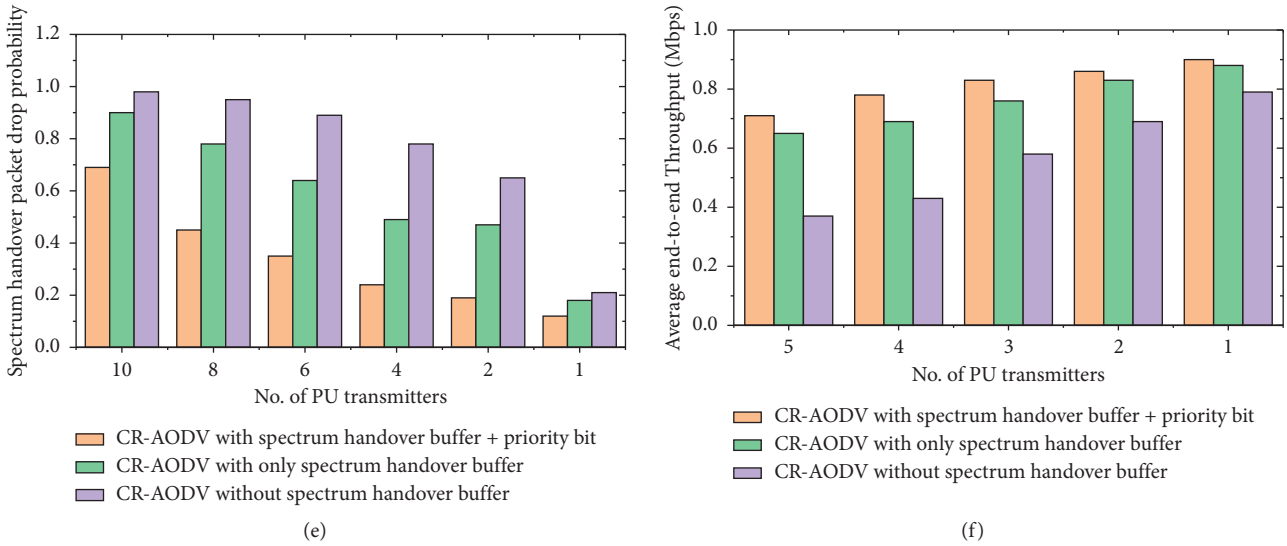


FIGURE 5: Cognitive TCP performance evaluation with respect to spectrum handoff. (a) CR-TCP congestion window variation with varied simulation time. (b) End-to-end TCP throughput with respect to simulation time. (c) Spectrum handoff local channel-route discovery with respect to no. of CR nodes. (d) Spectrum handoff local channel-route discovery with respect to the PU transmitters. (e) Spectrum handover packet drop probability with respect to no. of PU transmitters. (f) Average end-to-end CR-TCP throughput with respect to no. of PU transmitters.

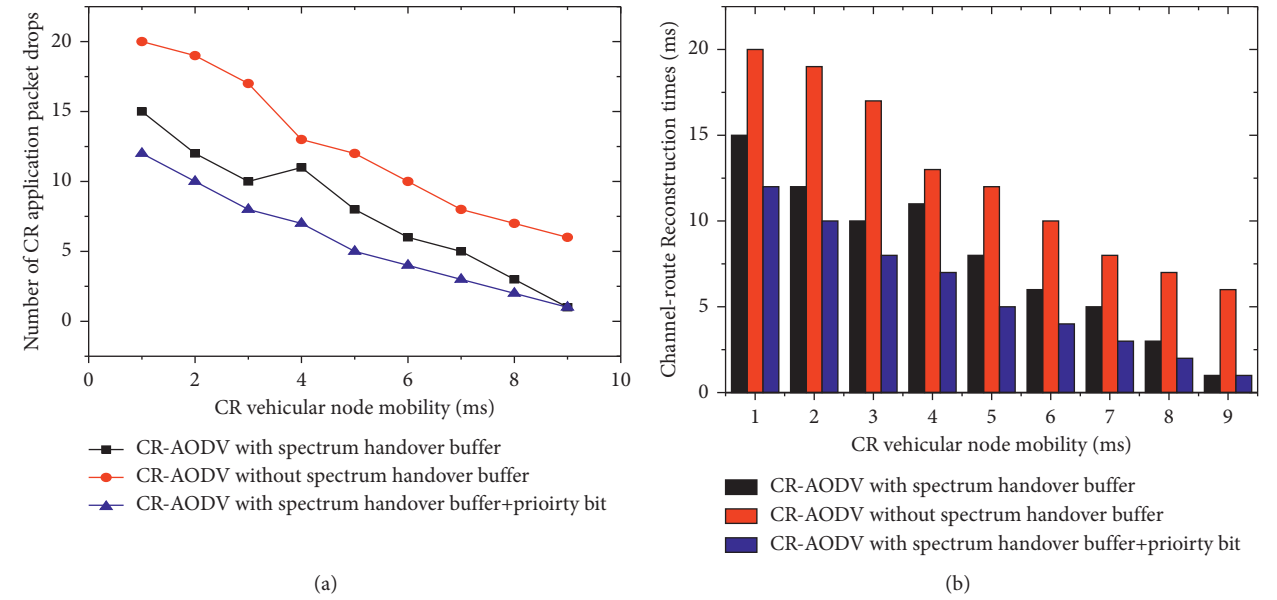


FIGURE 6: Performance of the CR vehicular network. (a) CR vehicular node mobility variation with respect to the number of application packet drops. (b) CR node mobility variation with respect to channel-route reconstruction time.

RED buffer will not drop the outgoing packets with priority enabled, whereas for the other two cases, there is more probability of packet drops at the output RED buffer. Figure 5(d) describes the local channel-route discovery with respect to the number of PU transmitters. In general, with increased PU transmitters, the probability of getting the opportunistic channel to transmit the CR application data gets reduced. In addition, the ongoing traffic flow gets packet drops because of spectrum handover at the intermediate CR nodes. Thus, the solutions with separate handover buffers at an intermediate CR node will attain better throughput in comparison with the solutions that does not support the packet buffers and priority bit. Figure 5(e) describes the probability of spectrum handover packet drops with respect to the number of PU transmitters. As shown in Figure 5(e), it is clear that the probability of the packet drop with respect to spectrum handover is directly proportional to the number of PU transmitters. Moreover, packet drop at the intermediate CR node gets dropped when there are no separate handover buffers. In other words, the input and output RED buffers

cannot hold on-the-fly CR packets until a new local channel link gets discovered. Thus, the probability of the CR packet drop is higher with the solutions that do not support spectrum handover buffers. Figure 5(f) describes the average end-to-end CR-TCP throughput with respect to the number of PU transmitters. CR-AODV with spectrum handover buffer along with priority bit is having higher achievable average end-to-end throughput in comparison with the other two techniques. This is due to reduced on-the-fly packet drops and reduced network congestion CR packet drops at the intermediate CR node.

Figure 6 explains about the CR vehicular ad-hoc network performance with respect to CR node handover from one base station to another base station. In general, when an intermediate CR vehicular node transits from one coverage area to other coverage areas, then the in-flight forwarded packets have to be either stored in a buffer of aggregate traffic flow buffer or CR node mobility priority buffer. In our experimental results of Figure 6(a), we tested the performance of the CR vehicular ad-hoc network with different node handover mobility scenarios. As shown in Figure 6(a), the packets drops are high when the probability of node handover from one base station to another base station is very high. In other words, the packet drops with node mobility variation of 1 msec are much higher than the packet drops with node mobility variation of 8 msec. This is due to heavy inflight traffic flow at the intermediate CR nodes at the time of CR node mobility. The CR node mobility variation with respect to channel reconstruction time is compared in Figure 6(b). It is noteworthy that when a CR node handover to another base station, then the opportunistic licensed channel needs to be tested before continuing with the data transmission. Otherwise, there may be a collision with the primary user data which needs to be protected with higher priority. In Figure 6(b), channel reconstruction times increase with frequent CR vehicular node handover in comparison with the nodes that hand over with longer delays. This is because when a CR node handovers from one base station to another base station more frequently, then inflight packets have to be buffered in more number of times at the base station that results in congestion with the aggregate traffic flows.

7. Conclusion

Performance of CR-TCP in CR-VANET directly depends on the number of CR packet drops in between the source and destination CR nodes. With opportunistic dynamic channel access, there is a new packet drop due to spectrum handover (PU active in the current CR communication channel) in between the source and destination CR nodes. To alleviate the spectrum handover packet drops, this paper proposes a "spectrum handover-based AODV routing protocol in CRAHNs." With the proposed packet buffering and forwarding mechanism, the achievable end-to-end average throughput is enhanced as compared to the existing CR-AODV routing protocols.

Data Availability

The data used to support the findings of the study are included within this article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Key Project of Sichuan Province Science and Technology (2021YFG0186).

References

[1] S. H. Chun and R. J. La, "Secondary spectrum trading-auction-based framework for spectrum allocation and profit sharing," *IEEE/ACM Transactions on Networking*, vol. 21, no. 1, 2013.

[2] C. Lăcătuşş, D. Akopian, P. Yaddanapudi, and M. Shadaram, "Flexible spectrum and power allocation for OFDM unlicensed wireless systems," *IEEE System Journal*, vol. 3, no. 2, 2009.

[3] T. Baykas, M. Kasslin, M. Cummings et al., "Developing a standard for TV white space coexistence: technical challenges and solution approaches," *IEEE Wireless Communications*, vol. 19, no. 1, 2012.

[4] F. B. S. D. Carvalho, W. T. A. Lopes, M. S. Alencar, and J. V. S. Filho, "Cognitive vehicular networks: an overview," *Procedia Computer Science*, vol. 65, 2015.

[5] R. S. Aqeel and A. Y. Kok-Lim, "Spectrum leasing in cognitive radio networks: A survey," *International Journal of Distributed Sensor Networks*, vol. 10, p. 2, 2014.

[6] R. Tandra, S. M. Mishra, and A. Sahai, "What is a spectrum hole and what does it take to recognize one?" 2009.

[7] B. F. Lo, "A survey of common control channel design in cognitive radio networks," *Physical Communication*, vol. 4, no. 1, 2011.

[8] M. Cesana, F. Cuomo, and E. Ekici, "Routing in cognitive radio networks: challenges and solutions," *Ad Hoc Networks*, vol. 9, no. 3, pp. 228–248, 2011.

[9] uting Wang, G. Zheng, H. Ma, Li Yang, and J. Li, "A joint channel selection and routing protocol for cognitive radio network," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 6848641, 7 pages, 2018.

[10] H. Khalifé, N. Malouch, and S. Fdida, "Multihop cognitive radio networks: to route or not to route," *IEEE Network*, vol. 23, no. 4, 2009.

[11] N. Devroye, M. Vu, and V. Tarokh, "Cognitive radio networks: highlights of information theoretic limits, models, and design," *IEEE Signal Processing Magazine*, vol. 26, no. 7, 2008.

[12] A. Guirguis, F. Digham, K. G. Seddik, M. Ibrahim, K. A. Harras, and M. Youssef, "Primary user-aware optimal discovery routing for cognitive radio networks," *IEEE Transactions on Mobile Computing*, vol. 18, no. 1, pp. 193–206, 2019.

[13] S. W. Kim, M. Pan, G. P. Joshi, O. Gazi, J. He, and M. Coupechoux, "Applications of cognitive radio networks: recent advances and future directions," *International Journal of Distributed Sensor Networks*, vol. 12, p. 1, 2016.

[14] I. F. Akyildiz, W. Y. Lee, and K. R. Chowdhury, "Spectrum management in cognitive radio ad hoc networks," *IEEE Network*, vol. 23, no. 4, 2009.

[15] K. Kotobi, P. Mainwaring, C. Tucker, and S. Bilén, "Data-throughput enhancement using data mining-informed cognitive radio," *Electronics*, vol. 4, no. 2, 2015.

[16] S. Anamalamudi and M. Jin, "Low rate common control channel based AODV routing protocol for cognitive radio ad-hoc networks," in *Proceedings of the International Conference on Ubiquitous and Future Networks, ICUFN*, Da Nang, Vietnam, July 2013.

[17] M. Zareei, A. K. M. Muzahidul Islam, S. Baharun, C. Vargas-Rosales, L. Azpilicueta, and N. Mansoor, "Medium access control protocols for cognitive radio ad hoc networks: A survey," *Sensors (Switzerland)*, vol. 17, no. 9, 2017.

[18] K. Singh and S. Moh, "An energy-efficient and robust multipath routing protocol for cognitive radio ad hoc networks," *Sensors (Switzerland)*, vol. 17, no. 9, 2017.

[19] S. Anamalamudi and M. Jin, "Performance enhancement of TCP in cognitive mobile IP based networks," in *Proceedings of the International Conference on Ubiquitous and Future Networks, ICUFN*, Shanghai, China, July 2014.

[20] S. Anamalamudi and M. Jin, "Energy-efficient hybrid CCC-based MAC protocol for cognitive radio ad hoc networks," *IEEE Systems Journal*, vol. 10, no. 1, pp. 358–369, 2016.

[21] T. Safdar Malik and M. Hilmi Hasan, "Reinforcement learning-based routing protocol to minimize channel switching and interference for cognitive radio networks," *Complexity*, vol. 2020, Article ID 8257168, 24 pages, 2020.

[22] A. Iqbal, R. Hussain, A. Shakeel et al., "Enhanced spectrum access for QoS provisioning in multi-class cognitive D2D communication system," *IEEE Access*, vol. 9, pp. 33608–33624, 2021.

[23] A. R. Sangi, M. S. Alkatheiri, S. Anamalamudi, and J. Liu, "Cognitive AODV routing protocol with novel channel-route failure detection," *Multimedia Tools and Applications*, vol. 79, 2019.

[24] S. Anamalamudi, M. Jin, J. M. Kim, and C. Liu, "On demand priority packet forwarding for TCP performance enhancement in cognitive mobile IP networks," *Wireless Personal Communications*, vol. 86, no. 4, 2016.

[25] K. N. Qureshi, A. H. Abdullah, M. Bukhari, and R. W. Anwar, "SSNM-Smart sensor network model for vehicular ad hoc networks," in *Proceedings of the 2015 International Conference on Smart Sensors and Application (ICSSA)*, pp. 82–87, Kuala Lumpur, Malaysia, May 2015.

[26] K. N. Qureshi, S. Din, G. Jeon, and F. Piccialli, "Link quality and energy utilization based preferable next hop selection routing for wireless body area networks," *Computer Communications*, vol. 149, pp. 382–392, 2020.

[27] K. N. Qureshi, A. H. Abdullah, O. Kaiwartya, U. Fasee, I. Saleem, and A. Ayman, "Weighted link quality and forward progress coupled with modified RTS/CTS for beaconless packet forwarding protocol (B-PFP) in VANETs," *Telecommunication Systems*, vol. 75, 2016.

[28] K. N. Qureshi, A. H. Abdullah, J. Lloret, and A. Altameem, "OBPF: opportunistic beaconless packet forwarding strategy for vehicular ad hoc networks," *KSII Transactions on Internet and Information Systems*, vol. 10, no. 5, pp. 2144–2165, 2016.

[29] A. Meddeb Makhlouf and M. Guizani, "SE-AOMDV: Secure and efficient AOMDV routing protocol for vehicular communications," *International Journal of Information Security*, vol. 18, no. 5, pp. 665–676, 2019.

[30] A. Ali, L. Abbas, M. Shafiq et al., "Hybrid fuzzy logic scheme for efficient channel utilization in cognitive radio networks," *IEEE Access*, vol. 7, pp. 24463–24476, 2019.

[31] Z. Ye, Y. Wang, and P. Wan, "Joint channel allocation and power control based on long short-term memory deep Q network in cognitive radio networks," *Complexity*, vol. 2020, Article ID 1628023, 11 pages, 2020.

[32] C. Cormio and K. R. Chowdhury, "Common control channel design for cognitive radio wireless ad hoc networks using adaptive frequency hopping," *Ad Hoc Networks*, vol. 8, no. 4, 2010.

[33] A. Misra and S. Banerjee, "MRPC: Maximizing network lifetime for reliable routing in wireless environments," in *Proceedings of the 2002 IEEE Wireless Communications and Networking Conference Record. WCNC 2002 (Cat. No.02TH8609)*, Orlando, FL, USA, March 2002.

[34] Q. Li, M. Yang, H. Wang, Y. Jiang, and J. Zeng, "A finite queue model analysis of PMRC-based wireless sensor networks," in *Proceedings of the 2008 International Conference on Wireless Networks, ICWN 2008*, pp. 30–35, Las Vegas, NV, USA, July 2008.

[35] J Zhong, "Development of NS-2 based cognitive radio cognitive network simulator," M. S. thesis, Michigan Technological University, Houghton, MI, USA, 2009.

[36] C. Cordeiro, K. Challapali, D. Birru, and N. Sai Shankar, "IEEE 802.22: The First worldwide wireless standard based on cognitive radios," in *Proceedings of the 2005 1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN 2005*, Baltimore, MD, USA, November 2005.

[37] M. Nekovee, "A survey of cognitive radio access to TV white spaces," in *Proceedings of the International Journal of Data Mining and Bioinformatics*, St. Petersburg, Russia, October 2010.

Hindawi
Complexity
Volume 2021, Article ID 7257248, 12 pages
https://doi.org/10.1155/2021/7257248

Research Article

Distributed Observer Design for Linear Systems under Time-Varying Communication Delay

Kairui Chen,[1,2] **Junwei Wang**,[3] **Xiaojing Zhong**,[1] **and Guanyu Lai**[4]

[1]*School of Mechanical and Electrical Engineering, Guangzhou University, Guangzhou 510006, China*
[2]*Key Laboratory of System Control and Information Processing, Ministry of Education, Shanghai 200240, China*
[3]*School of Mathematics and Statistics, Guangdong University of Foreign Studies, Guangzhou 510006, China*
[4]*School of Automation, Guangdong University of Technology, Guangzhou 510006, China*

Correspondence should be addressed to Junwei Wang; wangjwlj@gmail.com

Received 23 June 2021; Revised 24 July 2021; Accepted 20 August 2021; Published 7 September 2021

Academic Editor: Zhiwei Gao

The paper investigates the state estimation problem of general continuous-time linear systems with the consideration of time-varying communication delay. A solution is proposed in terms of the networked distributed observer, which consists of multiple local observers. Each local observer relies on only part of the system output and exchanges information with neighbors through undirected links modeled by a prespecified communication graph. A simple approach for computing observer parameters is presented by solving a parametric algebraic Riccati equation. Furthermore, by the Lyapunov–Krasovskii stability theorem, an upper bound of the delay could be calculated explicitly and together with the conditions of joint observability and connectivity of the communication graph; the resulting distributed observers work coordinately to achieve an asymptotic estimate of the full plant state. An illustrative example is provided to confirm the analytical results.

1. Introduction

State estimation or reconstruction is one of the most essential problems in control science and technology. In many control strategies (e.g., state feedback control, process monitoring, and fault detection), we need to have the state variables to construct the controller. In practice, it would be hard to directly obtain all states of an actual system because of technical limitations or perhaps immense cost by using complex sensors [1, 2]. Therefore, there is an urgent necessity to achieve an accurate estimation of the unmeasurable state variables. For deterministic linear time-invariant multivariable systems, the well-known Luenberger observer [3, 4] offers a comprehensive solution to accurately reconstruct the state variables by using the observability of the systems.

During the past decades, based on the Luenberger observer, many observer design schemes were proposed for various linear and nonlinear systems [5, 6], switched/hybrid systems [7–9], and systems modeled by partial differential equations or differential equations [10–14]. So far, most works in the literature of the observer design consider the scenario that the observer has access to the full output information of the system to be observed. However, when estimating the states of a spatially deployed system, it is difficult to obtain full information of its output signals individually [15]. This requires us to find a strategy to guarantee an asymptotic estimate of the full plant states while each observer receives only a portion of the measured outputs [16–18]. One feasible strategy to consider this issue is the so-called distributed observer consisting of networked local observers. A key feature of networked distributed observers is that each local observer rely on only its local measured outputs and cooperates with its neighboring observers to estimate the full plant states. This feature can be helpful in formational control of vehicles, autonomous underwater vehicles (AUVs), and spacecrafts [19–21].

Inspired by the extension of Kalman filtering by the idea of distributed control [22, 23], the design of the distributed observer for time-invariant linear systems has recently attracted lots of attention. For example, the authors in [24, 25] proposed a network of augmented observers and achieved the right state estimation of discrete-time linear time-invariant systems.

However, the introduction of an augmented state in each local observer results in the expanded dimensions of the augmented observer, which could induce a heavy computational burden in implementing their proposed schemes. For continuous-time linear time-invariant systems, the networked Luenberger observers proposed in [26–29] provided a feasible solution. The particularity of their formulations is that the effectiveness of the distributed observers can be verified by examining the observability of the system and connectivity of the communication network separately. Along with the idea in [26], different algorithmic procedures to design gain parameters of each local observer using local information of the plant dynamics and its own measurement output were presented in [30, 31]. However, the distributed observer schemes proposed by the above-mentioned works considered ideal interconnection networks without communication time delay among local observers.

In many practical systems, time delays are inevitable, which could degrade the performance of the system or even result instability. Some research studies concerning the time-delay systems have been reported, such as the time-delay state-space system with disturbances [32] and systems with state and input time delays [33]. In network communication systems, the information exchange on the network is often affected by time delay due to physical limitations in communication channels, finite-time information processing, or time-response of actuators [34–37]. As these communication delays may induce instability or poor performance of the systems, they should be taken into consideration to ensure performance of the proposed distributed observer. In [38], an augmented observer, which integrated robust estimation and compensation technique, was proposed for nonlinear systems subjected to unknown measurement delays. Up to now, there are few works on networked distributed observer design for continuous-time linear systems with the consideration communication delays. In [39, 40], two kinds of distributed Luenberger observer, tolerating arbitrarily large communication delays were proposed, while their sufficient conditions in both papers required all eigenvalues of the system matrix on the imaginary axis. In addition, the communication delays needed to be constant and their values to be known exactly in [39, 40], which is quite restrictive as communication delay is often time varying.

The lack of considering time-varying communication delay and the restriction on the system matrix motivate us to develop a novel distributed observer algorithm for general continuous-time linear systems by assuming the following. (1) The linear plants have general poles, i.e., poles of the system matrix can have zero, negative, or positive real parts. (2) The communication delay in the communication network is time varying. Therefore, the local observers in this paper are deployed to estimate linear plants with general poles under time-varying communication delay. The ability of observing general linear systems is critical for observers as they are generally used in the feedback loops. By solving the parametric algebraic Riccati equation, the gain matrix for local measurement output is determined. Furthermore, by applying the Lyapunov–Krasovskii theorem, we derive the process of computing coupling gains for the information exchange among local observers and the upper bound of the time-varying delay, under which distributed observers achieve omniscience eventually.

The remainder sections of this paper are organized in the following way. Section 2 reviews some necessary preliminaries and introduces the problem statement. In Section 3, we present the detailed design of the distributed observers for linear systems with time-varying communication delay. Simulation results are reported in Section 4. Finally, concluding remarks and further research directions are presented in Section 5.

Let $\mathbb{R}^+$, $\mathbb{R}^n$, and $\mathbb{R}^{n\times m}$ stand for the nonnegative real number, n-dimensional Euclidean space, and the set of all $n \times m$ real matrices, respectively, O_n and I_n signify the zero matrix and the identity matrix, 1_n denotes the n-dimensional column vector with all elements as 1, $X > 0$ ($X < 0$) means that $X \in \mathbb{R}^{n\times n}$ is a positive definite matrix (negative definite matrix), $\otimes$ is the Kronecker product, and $\mathrm{Diag}(\alpha_1, \alpha_2, \ldots, \alpha_n)$ stands for a diagonal matrix with diagonal elements α_i, $i = 1, 2, \ldots, n$, being real numbers (or a block diagonal matrix with diagonal blocks α_i, $i = 1, 2, \ldots, n$, being matrices). All matrices are assumed to be compatible in this paper.

2. Preliminaries and Problem Formulation

2.1. Graph Theory. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an undirected graph with the node set $\mathcal{V} = \{1, \ldots, N\}$ and the edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. An edge in $\mathcal{E}$ is denoted by an unordered pair of distinct nodes (i, j), and $(i, j) \in \mathcal{E}$ if and only if $(j, i) \in \mathcal{E}$. The neighbor set of node i is denoted by $N_i = \{j \in \mathcal{V} : (i, j) \in \mathcal{E}\}$. $G = [g_{ij}] \in \mathbb{R}^{N\times N}$ represents the adjacency matrix associated with $\mathcal{G}$, where $g_{ij} = g_{ji} > 0$ if $(i, j) \in \mathcal{E}$ and $g_{ij} = g_{ji} = 0$ otherwise. Assuming that there is no self-cycle for each node, i.e., $g_{ii} = 0$, correspondingly, the Laplacian matrix $L = [l_{ij}] \in \mathbb{R}^{N\times N}$ is defined as $l_{ii} = \sum_{k\in N_i} g_{ik}$ and $l_{ij} = -g_{ij}$, for $j \neq i$. For an undirected graph $\mathcal{G}$, both its adjacency matrix and Laplacian matrix are symmetric. A sequence of edges $(i_1, i_2), (i_2, i_3), \ldots, (i_{k-1}, i_k)$ is called a path from node i_1 to node i_k. An undirected graph is called connected if, for any node pair $i, j \in \mathcal{V}$, there exists a path from i to j.

The following lemma is a well-known result in algebraic graph theory [41].

Lemma 1. *For an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with N nodes, its Laplacian matrix L has N real eigenvalues which can be arranged in an ascending order as $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$. Moreover, if $\mathcal{G}$ is connected, $\lambda_1 = 0$ is a simple eigenvalue of L with corresponding eigenvector 1_N.*

2.2. Problem Formulation. In this paper, the linear system to be observed is

$$\begin{aligned} \dot{x} &= Ax, \\ y &= Cx = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_N \end{bmatrix} x = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}, \end{aligned} \tag{1}$$

where $x = [x_1, \ldots, x_n]^T \in \mathbb{R}^n$ is the state, $y \in \mathbb{R}^q$ is the measurement output, and the matrices $A \in \mathbb{R}^{n\times n}$ and

$C \in \mathbb{R}^{q\times n}$ are system matrices. We have partitioned the output y as $y = [y_1^T, \ldots, y_N^T]^T$ according to the partition of the matrix $C = [C_1^T, \ldots, C_N^T]^T$, where $y_i \in \mathbb{R}^{q_i}$ and $C_i \in \mathbb{R}^{q_i\times n}$.

For plant (1), with the given undirected communication graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, our objective is to construct a group of N networked observers, where each observer computes a state estimate that asymptotically converges to the entire plant state. The following constraints are challenging in the distributed estimation problem: (i) each local observer can only access the portion $y_i = C_i x$ of the plant output $y = Cx$ and (ii) the information exchange among local observers is restricted by a sparsity constraint described by the communication topology, where local observers are represented by nodes and the edges in $\mathcal{E}$ which determine the information links among local observers. In light of the structure of the distributed observers in [1, 39], we propose a kind of distributed observer taking the following form:

$$\begin{aligned}\dot{\hat{x}}_i &= A\hat{x}_i + H_i\left(y_i - C_i\hat{x}_i\right)\\ &\quad + k\sum_{j\in N_i} g_{ij}\left(\hat{x}_j(t-d(t))\right) - \left(\hat{x}_i(t-d(t))\right).\end{aligned} \tag{2}$$

where $\hat{x}_i = [\hat{x}_{i1}, \ldots, \hat{x}_{in}]^T \in \mathbb{R}^n$ is an estimate of the state generated by the ith local observer, g_{ij} shows the relation between observer i and observer j, N_i is the set of neighboring observers of observer i, $H_i \in \mathbb{R}^{n\times q_i}$ is the gain matrix for local output measurements to be determined, and $k > 0$ is the coupling strength for neighboring local observers to be designed. $d(t)$: $\mathbb{R}^+ \longrightarrow \mathbb{R}^+$ is a time-varying communication delay. The delay is assumed to be bounded by a constant $D > 0$ (i.e., $0 \le d(t) \le D$), which could be estimated by empirical observations or testings. To make the distributed observer (2) work under the delayed communication, we additionally set the initial condition $x(\theta) = \varphi(\theta)$ and $\hat{x}_i(\theta) = \varphi_i(\theta)$, for any $\theta \in [-D, 0]$, where $\varphi(\theta)$ and $\varphi_i(\theta)$ are known continuous initial state-vector functions.

Remark 1. In the designed distributed observer (2), the term $\hat{x}_j(t-d(t)) - \hat{x}_i(t-d(t))$ represents the relative state between observer i and observer j, which is subject to communication time delay $d(t)$. Therefore, the exactly data of $d(t)$ is not required in the distributed observer. In other words, we intend to measure $(\hat{x}_j(t) - \hat{x}_i(t))$, but the information that we obtain is delayed by the communication network. The time delay $d(t)$ exists inevitably in the communication channel, and we use $d(t)$ to imitate the time delay. Thus, we want to design a kind of distributed observer that considers the delay in order to depress the degradation caused by the delay.

Assumption 1. The undirected graph $\mathcal{G}$ is connected.

Assumption 2. The pair (A, C) is observable.

Remark 2. According to dynamics (2), each local observer recuperates the full plant state x from its local measurements y_i and relative delayed information $\hat{x}_{ji}(t-d(t))\,(j \in N_i)$ of its neighbors. For the special case that there is one C_i such that (A, C_i) is detectable, then other observers are unnecessary and the distributed structure becomes trivial. Therefore, we assume the pair (A, C) is detectable but each separated pair (A, C_i) is not; then, exchanging the estimates among neighboring observers is obligatory.

Let $\eta_i = \hat{x}_i - x$ be the estimation error of the ith observer. Then, the dynamics of observer error η_i can be obtained from (1) and (2) as follows:

$$\dot{\eta}_i = \left(A - H_iC_i\right)\eta_i - k\sum_{j\in N_i} l_{ij}\eta_j(t-d(t)), \tag{3}$$

where l_{ij} is the element of the Laplacian matrix L of $\mathcal{G}$. Stack the local observer errors $\eta_i(t)$ to get the joint vector of the observer error $\eta = (\eta_1^T, \eta_2^T, \ldots, \eta_N^T)^T$ and denote $\phi(t) = t - d(t)$. Then, the observer error dynamics governed by (3) can be rewritten as

$$\begin{aligned}&\dot{\eta}(t)\\ &\quad= \left(I_N\otimes A - \overline{HC}\right)\eta(t) - k\left(L\otimes I_n\right)\eta(\phi(t))\\ &\quad= \left(I_N\otimes A - \overline{HC} - k\left(L\otimes I_n\right)\right)\eta(t)\\ &\qquad + k\left(L\otimes I_n\right)(\eta(t) - \eta(\phi(t)))\\ &\quad= \left(I_N\otimes A - \overline{HC} - k\left(L\otimes I_n\right)\right)\eta(t) + k\left(L\otimes I_n\right)\rho(t),\end{aligned} \tag{4}$$

where

$$\begin{aligned}\rho(t) &= \eta(t) - \eta(\phi(t)),\\ \overline{H} &= \mathrm{Diag}\{H_1, H_2, \ldots, H_N\},\\ \overline{C} &= \mathrm{Diag}\{C_1, C_2, \ldots, C_N\}.\end{aligned} \tag{5}$$

Before proceeding further, we first introduce the concept of asymptotic omniscience as stated in [25].

Definition 1. For plant (1), the distributed observer (2) is said to achieve asymptotic omniscience if $\lim_{t\longrightarrow+\infty}\|\eta_i(t)\| = \lim_{t\longrightarrow+\infty}\|\hat{x}_i(t) - x(t)\| = 0$, i.e., each observer state $\hat{x}_i(t)$ converges to $x(t)$ asymptotically.

Remark 3. From the multiagent point of view, each local observer in (2) is in the situation of a follower and system (1) to be observed is analogous to the leader. Thus, the dynamical structure of (1) and (2) is in fact a special case of the leader-following multiagent system. A difference lies in that each follower often receives full state vector of the leader, while each local observer in (2) has access to only partial information of the output [39]. To the best of the authors' knowledge, rare result on the consensus problem of such a type of multiagent systems was reported by researchers. Therefore, the problem of designing an observer is still an open problem for the general form (2), especially with the consideration of time-varying communication delay.

In what follows, we will focus on computing the coupling strength k and the observer gain H_i such that the designed distributed observer (2) with time-varying communication delay $d(t)$ achieves omniscience asymptotically.

3. Distributed Observer for General Linear Systems

In this section, we are going to design distributed observers for general linear system (1), where the matrix A is general, i.e., may have eigenvalues with zero, negative, or positive real parts. Two kinds of distributed observer algorithms, with communication delay and without communication delay, are designed with rigorous proofs.

Before providing the distributed observer design method, some key technical lemmas for the design and proof of the distributed observer are introduced.

Lemma 2 (see [42]). *For (A, C) satisfying Assumption 2, let*

$$\epsilon > -2\min\{\mathrm{Re}(\lambda(A))\}, \tag{6}$$

where $\mathrm{Re}(\lambda(A))$ is the real parts of $\lambda(A)$ which are eigenvalues of A. Then, a unique positive definite solution of the following ARE,

$$PA^T + AP - PC^TCP = -\epsilon P, \tag{7}$$

is given by $P(\epsilon) = W^{-1}(\epsilon)$, where $W(\epsilon)$ can be obtained by solving the following Lyapunov equation:

$$W\left(A + \frac{\epsilon}{2}I_n\right) + \left(A + \frac{\epsilon}{2}I_n\right)^T W = C^TC. \tag{8}$$

Moreover, $(\mathrm{d}P(\epsilon)/d(\epsilon)) > 0$ and $P(\epsilon)$ has the following properties:

(i) $\mathrm{tr}(CP(\epsilon)C^T) = 2\mathrm{tr}(A) + n\epsilon$

(ii) $P(\epsilon)C^TCP(\epsilon) \le (2\mathrm{tr}(A) + n\epsilon)P(\epsilon)$

Lemma 3. *Let (A, C) satisfy Assumption 2 and $P(\epsilon) > 0$ be the solution to (7). Then, we have*

$$A^TP^{-1}A \le \zeta(\epsilon)P^{-1}, \tag{9}$$

where

$$\begin{aligned}\zeta(\epsilon) = {} & \frac{1}{2}(2\mathrm{tr}(A) + n\epsilon)^2 - \frac{1}{2}\epsilon(2\mathrm{tr}(A) + n\epsilon)) \\ & - \epsilon\mathrm{tr}(A) - \mathrm{tr}\left(A^2\right).\end{aligned} \tag{10}$$

Proof. From the fact $\mathrm{tr}(XY) = \mathrm{tr}(YX)$, we have

$$\begin{aligned}A^TP^{-1}A &= P^{-1}PA^TP^{-1}APP^{-1} \\ &= P^{-(1/2)}\left(P^{-(1/2)}PA^TP^{-1}APP^{-(1/2)}\right)P^{-(1/2)} \\ &\le \mathrm{tr}\left(P^{-(1/2)}PA^TP^{-1}APP^{-(1/2)}\right)P^{-1} \\ &= \mathrm{tr}\left(PA^TP^{-1}A\right)P^{-1}.\end{aligned} \tag{11}$$

Postmultiplying the two sides of ARE (7) with $P^{-1}A$ gives

$$PA^TP^{-1}A + A^2 - PC^TCA = -\epsilon A, \tag{12}$$

which leads to

$$\mathrm{tr}\left(PA^TP^{-1}A\right) = \mathrm{tr}\left(PC^TCA\right) - \epsilon\mathrm{tr}(A) - \mathrm{tr}\left(A^2\right). \tag{13}$$

Using ARE (7) again, we obtain

$$\begin{aligned}&\mathrm{tr}\left(PC^TCA\right) \\ &\quad = \mathrm{tr}\left(C^TCAP\right) = \mathrm{tr}\left(C^TC\left(PC^TCP - PA^T - \epsilon P\right)\right) \\ &\quad = \mathrm{tr}\left(\left(CPC^T\right)^2\right) - \mathrm{tr}\left(PC^TCA\right) - \epsilon\mathrm{tr}\left(CPC^T\right).\end{aligned} \tag{14}$$

Notice that $\mathrm{tr}(XY) \le \mathrm{tr}(X)\mathrm{tr}(Y)$ for $X \ge 0$ and $Y \ge 0$. Then, it follows from (14) and Lemma 2 that

$$\begin{aligned}\mathrm{tr}\left(PC^TCA\right) &= \frac{1}{2}\mathrm{tr}\left(\left(CPC^T\right)^2\right) - \frac{1}{2}\epsilon\mathrm{tr}\left(CPC^T\right) \\ &\le \frac{1}{2}\left(\mathrm{tr}\left(CPC^T\right)\right)^2 - \frac{1}{2}\epsilon\mathrm{tr}\left(CPC^T\right) \\ &= \frac{1}{2}(2\mathrm{tr}(A) + n\epsilon)^2 - \frac{1}{2}\epsilon(2\mathrm{tr}(A) + n\epsilon)).\end{aligned} \tag{15}$$

Substituting (15) into (13) gives

$$\begin{aligned}\mathrm{tr}\left(PA^TP^{-1}A\right) &= \mathrm{tr}\left(PC^TCA\right) - \epsilon\mathrm{tr}(A) - \mathrm{tr}\left(A^2\right) \\ &\le \frac{1}{2}(2\mathrm{tr}(A) + n\epsilon)^2 - \frac{1}{2}\epsilon(2\mathrm{tr}(A) + n\epsilon)) \\ &\quad - \epsilon\mathrm{tr}(A) - \mathrm{tr}\left(A^2\right).\end{aligned} \tag{16}$$

Consequently, inequality (9) follows from (11) and (16). This completes the proof. □

Lemma 4 (Schur complement [43]). *For a given symmetric matrix $S = \begin{bmatrix} S_{11} & S_{12} \\ S_{12}^T & S_{22} \end{bmatrix}$ with $S_{11} = S_{11}^T$ and $S_{22} = S_{22}^T$, $S < 0$ is equivalent to $S_{11} < 0, S_{22} - S_{12}^TS_{11}^{-1}S_{12} < 0$.*

Lemma 5 (Jensen's inequality [44]). *Let t_1 and t_2 be scalars with $t_2 \ge t_1$, and a vector-valued function $\eta(t)$: $[t_1, t_2] \longrightarrow \mathbb{R}^n$ so that the related integrals are well defined. Then, for any positive definite matrix $P \in \mathbb{R}^{n\times n}$, we have the following inequality:*

$$\left(\int_{t_1}^{t_2}\eta(t)\mathrm{d}t\right)^T P\left(\int_{t_1}^{t_2}\eta(t)\mathrm{d}t\right) \le (t_2 - t_1)\int_{t_1}^{t_2}\eta^T(t)P\eta(t)\mathrm{d}t. \tag{17}$$

Based on the above preparations, we are ready to propose the following theorem.

Theorem 1. *Consider the linear system (1) under Assumptions 1 and 2. Let $0 = \lambda_1 < \lambda_2 \le \cdots \le \lambda_N$ denote N real eigenvalues of L and design feedback gain H_i and coupling strength k in the distributed observer (2) as*

$$H_i = \frac{N}{2}P(\epsilon)C_i^T, k > \frac{\rho_1(\epsilon)}{\lambda_2}, \tag{18}$$

where $P(\epsilon) > 0$ is the solution to (7) and

$$\rho_1(\epsilon) = (4N-3)(2\mathrm{tr}(A)+n\epsilon) + 1.5\epsilon + \frac{2}{\epsilon}(1+N)^2(2\mathrm{tr}(A)+n\epsilon)^2. \tag{19}$$

Then, the distributed observer (2) achieves asymptotic omniscience if the upper bound D of the time-varying delay $d(t)$ satisfies

$$D < \sqrt{\frac{\epsilon}{3\epsilon k^2\lambda_N^2 + 2k\lambda_N\rho_2(\epsilon)}}, \tag{20}$$

where

$$\rho_2(\epsilon) = \left(3+1.5N^2\right)(2\mathrm{tr}(A)+n\epsilon)^2 + 3k^2\lambda_N^2 - 3\epsilon(2\mathrm{tr}(A)+n\epsilon) - 6\epsilon\mathrm{tr}(A) - 6\mathrm{tr}\left(A^2\right). \tag{21}$$

Proof. It can be seen that the distributed observer (2) achieves asymptotic omniscience if and only if the observer error system (4) is asymptotically stable. Thus, we choose the following Lyapunov functional candidate to prove the convergence of the global observer error system (4):

$$V(t) = V_1(t) + V_2(t), \tag{22}$$

where

$$V_1(t) = \eta^T(t)\overline{P}\eta(t), \qquad V_2(t) = \sigma\int_{-D}^{0}\int_{t+\theta}^{t}\dot{\eta}^T(s)\overline{P}\dot{\eta}(s)\mathrm{d}s\mathrm{d}\theta, \tag{23}$$

and $\overline{P} = I_N \otimes P^{-1}$, and $\sigma > 0$ is a scalar to be designed.

Along the solutions of (4) with H_i in (19), computing the derivative of $V_1(t)$ yields

$$\begin{aligned}\dot{V}_1(t) &= \eta^T(t)\left(I_N\otimes\left(A^TP^{-1}+P^{-1}A\right)\right)\eta(t) - N\eta^T(t)\overline{C}^T\overline{C}\eta(t) - 2k\eta^T(t)\left(L\otimes P^{-1}\right)\eta(t) + 2k\eta^T(t)\left(L\otimes P^{-1}\right)\rho(t)\\ &\le \eta^T(t)\left(I_N\otimes\left(A^TP^{-1}+P^{-1}A\right)\right)\eta(t) - N\eta^T(t)\overline{C}^T\overline{C}\eta(t) - k\eta^T(t)\left(L\otimes P^{-1}\right)\eta(t) + k\rho^T(t)\left(L\otimes P^{-1}\right)\rho(t)\\ &= \eta^T(t)\Pi(k)\eta(t) + k\rho^T(t)\left(L\otimes P^{-1}\right)\rho(t),\end{aligned} \tag{24}$$

where

$$\begin{aligned}\Pi(k) &= I_N\otimes\left(A^TP^{-1}+P^{-1}A\right) - N\overline{C}^T\overline{C} - kL\otimes P^{-1}\\ &= \mathrm{Diag}\{\Upsilon_1,\Upsilon_2,\ldots,\Upsilon_N\} - kL\otimes P^{-1},\end{aligned} \tag{25}$$

with $\Upsilon_i = A^TP^{-1} + P^{-1}A - NC_i^TC_i$. Then, according to (7), we have $\sum_{i=1}^{N}\Upsilon_i = N(A^TP^{-1}+P^{-1}A - \sum_{i=1}^{N}C_i^TC_i) = -\epsilon NP^{-1}$.

To analyze $\Pi(k)$, we first show that, for any scalar $\epsilon > 0$, there is a $k > 0$ such that

$$\Pi(k) < -\frac{1}{2}\epsilon\overline{P}. \tag{26}$$

By Assumption 1 and Lemma 1, we know that $(1/\sqrt{N})1_N$ is the right eigenvector of L associated with the zero eigenvalue. Let $U = [(1/\sqrt{N})1_N, U_1]$ be the orthogonal matrix such that

$$U^TLU = \mathrm{Diag}\left\{0,\widehat{L}\right\}, U^{-1} = U^T, \tag{27}$$

where $\widehat{L} = \mathrm{Diag}\{\lambda_2,\ldots,\lambda_N\}$. By premultiplying both sides of (26) with $U^T\otimes I_n$ and postmultiplying with $U\otimes I_n$, we have

$$\left(U^T\otimes I_n\right)\Pi(k)\left(U\otimes I_n\right) < -\frac{1}{2}\epsilon\overline{P}, \tag{28}$$

which can be rewritten as in (29)

$$\begin{pmatrix} -\frac{1}{2}\epsilon P^{-1} & \frac{1}{\sqrt{N}}\left(1_N^T\otimes I_n\right)\overline{\Upsilon}\left(U_1\otimes I_n\right)\\ * & \left(U_1^T\otimes I_n\right)\overline{\Upsilon}\left(U_1\otimes I_n\right) + \left(\frac{1}{2}\epsilon I_{N-1} - k\widehat{L}\right)\otimes P^{-1}\end{pmatrix} < 0. \tag{29}$$

By the Schur complement in Lemma 4, (29) is equivalent to

$$\left(U_1^T\otimes I_n\right)\overline{\Upsilon}\left(U_1\otimes I_n\right) + \left(\frac{1}{2}\epsilon I_{N-1} - k\widehat{L}\right)\otimes P^{-1} + \frac{2}{\epsilon N}\left(U_1^T\otimes I_n\right)\overline{\Upsilon}\left(1_N1_N^T\otimes P\right)\overline{\Upsilon}\left(U_1\otimes I_n\right) < 0, \tag{30}$$

where $\overline{\Upsilon} = \mathrm{Diag}\{\Upsilon_1,\Upsilon_2,\ldots,\Upsilon_N\}$.

From Lemma 2, we can get $\mathrm{tr}(C_iPC_i^T) \le \mathrm{tr}\ (CPC^T) = 2\mathrm{tr}(A) + n\epsilon$ and $C_i^TC_i \le C^TC \le (2\mathrm{tr}(A)+n\epsilon)P^{-1}$, and then, $C_i^TC_iPC_i^TC_i \le C_i^T\ \ \mathrm{tr}(C_iPC_i^T)C_i \le (2\mathrm{tr}(A)+n\epsilon)C_i^TC_i \le (2\mathrm{tr}(A)+n\epsilon)^2P^{-1}$. Hence, it can be verified that

$$\begin{aligned}\Upsilon_i &= -\epsilon P^{-1} + C^TC - NC_i^TC_i\\ &\le -\epsilon P^{-1} + C^TC \le (2\mathrm{tr}(A)+n\epsilon-\epsilon)P^{-1}\\ \Upsilon_iP\Upsilon_i &= \epsilon^2P^{-1} - 2\epsilon C^TC + 2\epsilon NC_i^TC_i + C^TCPC^TC - NC^TCPC_i^TC_i - NC_i^TC_iPC^TC + N^2C_i^TC_iPC_i^TC_i\\ &\le \epsilon^2P^{-1} + 2\epsilon(N-1)C^TC + (N+1)C^TCPC^TC + \left(N^2+N\right)C_i^TC_iPC_i^TC_i\\ &\le \left[\epsilon^2 + 2\epsilon(N-1)(2\mathrm{tr}(A)+n\epsilon) + (1+N)^2(2\mathrm{tr}(A)+n\epsilon)^2\right]P^{-1},\end{aligned} \tag{31}$$

which implies that

$$\left(U_1^T\otimes I_n\right)\overline{\Upsilon}\left(U_1\otimes I_n\right) \le (2\mathrm{tr}(A)+n\epsilon-\epsilon)\left(I_{N-1}\otimes P^{-1}\right), \tag{32}$$

and

$$\begin{aligned}&\left(U_1^T\otimes I_n\right)\overline{\Upsilon}\left(1_N1_N^T\otimes P\right)\overline{\Upsilon}\left(U_1\otimes I_n\right)\\&\quad\le N\left(U_1^T\otimes I_n\right)\overline{\Upsilon}\left(I_N\otimes P\right)\overline{\Upsilon}\left(U_1\otimes I_n\right)\\&\quad\le N\left[\epsilon^2+2\epsilon\left(N-1\right)\left(2\mathrm{tr}\left(A\right)+n\epsilon\right)\right.\\&\qquad\left.+\left(1+N\right)^2\left(2\mathrm{tr}\left(A\right)+n\epsilon\right)^2\right]\left(I_{N-1}\otimes P^{-1}\right),\end{aligned}\tag{33}$$

where $U_1^TU_1=I_{N-1}$ has been used.

Combining (32) and (33), the left-hand side of inequality (27) satisfies

$$\begin{aligned}&\left(U_1^T\otimes I_n\right)\overline{\Upsilon}\left(U_1\otimes I_n\right)+\left(\frac{1}{2}\epsilon I_{N-1}-k\widehat{L}\right)\otimes P^{-1}\\&\qquad+\frac{2}{\epsilon N}\left(U_1^T\otimes I_n\right)\overline{\Upsilon}\left(1_N1_N^T\otimes P\right)\overline{\Upsilon}\left(U_1\otimes I_n\right)\\&\qquad\le-\left(k\lambda_2-\rho_1\left(\epsilon\right)\right)\left(I_{N-1}\otimes P^{-1}\right),\end{aligned}\tag{34}$$

with $\rho_1(\epsilon)$ as defined in (19). By choosing the coupling strength k as in (18), the inequality (26) is thus proved.

We now simplify the term $k\rho^T(t)(L\otimes P^{-1})\rho(t)$ in (24). By Jensen's inequality in Lemma 5, one obtains

$$\begin{aligned}&k\rho^T\left(t\right)\left(L\otimes P^{-1}\right)\rho\left(t\right)\\&\quad=k\left(\int_{\phi\left(t\right)}^t\dot\eta\left(s\right)\mathrm{d}s\right)^T\left(L\otimes P^{-1}\right)\left(\int_{\phi\left(t\right)}^t\dot\eta\left(s\right)\mathrm{d}s\right)\\&\quad\le k\left(t-\phi\left(t\right)\right)\int_{\phi\left(t\right)}^t\dot\eta^T\left(s\right)\left(L\otimes P^{-1}\right)\dot\eta\left(s\right)\mathrm{d}s\\&\quad\le k\lambda_ND\int_{t-D}^t\dot\eta^T\left(s\right)\overline{P}\dot\eta\left(s\right)\mathrm{d}s,\end{aligned}\tag{35}$$

where we use the knowledge that $d(t)$ is bounded.

Substituting inequalities (26) and (35) into (24) gives

$$\dot V_1\left(t\right)\le-\frac{1}{2}\epsilon\eta^T\left(t\right)\overline{P}\eta\left(t\right)+k\lambda_ND\int_{t-D}^t\dot\eta^T\left(s\right)\overline{P}\dot\eta\left(s\right)\mathrm{d}s.\tag{36}$$

Next, the derivative of $V_2(t)$ along (4) with H_i in (18) is

$$\begin{aligned}\dot V_2\left(t\right)&=\sigma\int_{-D}^0\left(\dot\eta^T\left(t\right)\overline{P}\dot\eta\left(t\right)-\dot\eta^T\left(t+\theta\right)\overline{P}\dot\eta\left(t+\theta\right)\right)\mathrm{d}\theta\\&=\sigma D\dot\eta^T\left(t\right)\overline{P}\dot\eta\left(t\right)-\sigma\int_{t-D}^t\dot\eta^T\left(s\right)\overline{P}\dot\eta\left(s\right)\mathrm{d}s.\end{aligned}\tag{37}$$

By Lemma 3 and Young's inequality, the first term in (37) becomes

$$\begin{aligned}&\sigma D\dot\eta^T\left(t\right)\overline{P}\dot\eta\left(t\right)\\&\quad=\sigma D\eta^T\left(I_N\otimes A-\overline{HC}\right)^T\overline{P}\left(I_N\otimes A-\overline{HC}\right)\eta-2\sigma Dk\eta^T\\&\left(L^T\otimes P^{-1}\right)\left(I_N\otimes A-\overline{HC}\right)\eta+\sigma Dk^2\eta^T\left(L^TL\otimes P^{-1}\right)\eta\\&\qquad+2\sigma Dk\eta^T\left(I_N\otimes A-\overline{HC}\right)^T\left(L\otimes P^{-1}\right)\rho\\&\qquad-2\sigma Dk^2\eta^T\left(L^TL\otimes P^{-1}\right)\rho+\sigma Dk^2\rho^T\left(L^TL\otimes P^{-1}\right)\rho\\&\qquad\le3\sigma D\eta^T\left(t\right)\left(I_N\otimes A-\overline{HC}\right)^T\overline{P}\\&\left(I_N\otimes A-\overline{HC}\right)\eta\left(t\right)+3\sigma Dk^2\lambda_N^2\eta^T\left(t\right)\overline{P}\eta\left(t\right)\\&\qquad+3\sigma k^2\lambda_N^2D^2\int_{t-D}^t\dot\eta^T\left(s\right)\overline{P}\dot\eta\left(s\right)\mathrm{d}s\\&\quad=3\sigma D\eta^T\mathrm{Diag}\{\Psi_1,\Psi_2,\ldots,\Psi_N\}\eta\left(t\right)+3\sigma Dk^2\lambda_N^2\eta^T\left(t\right)\\&\overline{P}\eta\left(t\right)+3\sigma k^2\lambda_N^2D^2\int_{t-D}^t\dot\eta^T\left(s\right)\overline{P}\dot\eta\left(s\right)\mathrm{d}s,\end{aligned}\tag{38}$$

where

$$\begin{aligned}\Psi_i&=A^TP^{-1}A-\frac{N}{2}A^TC_i^TC_i-\frac{N}{2}C_i^TC_iA\\&\quad+\frac{N^2}{4}C_i^TC_iPC_i^TC_i\\&\le2A^TP^{-1}A+\frac{N^2}{2}C_i^TC_iPC_i^TC_i\\&\le2\zeta\left(\epsilon\right)P^{-1}+\frac{N^2}{2}\left(2\mathrm{tr}\left(A\right)+n\epsilon\right)^2P^{-1},\end{aligned}\tag{39}$$

with $\zeta(\epsilon)$ as defined in (10). With $\rho_2(\epsilon)$ in (21), we obtain

$$\begin{aligned}&\sigma D\dot\eta^T\left(t\right)\overline{P}\dot\eta\left(t\right)\\&\quad\le\sigma D\left(6\zeta\left(\epsilon\right)+\frac{3}{2}N^2\left(2\mathrm{tr}\left(A\right)+n\epsilon\right)^2+3k^2\lambda_N^2\right)\eta^T\left(t\right)\\&\overline{P}\eta\left(t\right)+3\sigma k^2\lambda_N^2D^2\int_{t-D}^t\dot\eta^T\left(s\right)\overline{P}\dot\eta\left(s\right)\mathrm{d}s\\&\quad\le\sigma D\rho_2\left(\epsilon\right)\eta^T\left(t\right)\overline{P}\eta\left(t\right)+3\sigma k^2\lambda_N^2D^2\int_{t-D}^t\dot\eta^T\left(s\right)\overline{P}\dot\eta\left(s\right)\mathrm{d}s.\end{aligned}\tag{40}$$

Then, substitution of (36) and (37) with (40) into the derivative of $V(t)$ in (22) results in

$$\begin{aligned}\dot V\left(t\right)&=\dot V_1\left(t\right)+\dot V_2\left(t\right)\\&\le-\frac{1}{2}\epsilon\eta^T\left(t\right)\overline{P}\eta\left(t\right)+k\lambda_ND\int_{t-D}^t\dot\eta^T\left(s\right)\overline{P}\dot\eta\left(s\right)\mathrm{d}s\\&\quad+\sigma D\dot\eta^T\left(t\right)\overline{P}\dot\eta\left(t\right)-\sigma\int_{t-D}^t\dot\eta^T\left(s\right)\overline{P}\dot\eta\left(s\right)\mathrm{d}s\\&\le-\left(\frac{1}{2}\epsilon-\sigma D\rho_2\left(\epsilon\right)\right)\eta^T\left(t\right)\overline{P}\eta\left(t\right)\\&\quad+\left(3\sigma k^2\lambda_N^2D^2+k\lambda_ND-\sigma\right)\int_{t-D}^t\dot\eta^T\left(s\right)\overline{P}\dot\eta\left(s\right)\mathrm{d}s.\end{aligned}\tag{41}$$

Therefore, by taking

$$\sigma=\frac{k\lambda_ND}{1-3k^2\lambda_N^2D^2},\tag{42}$$

which is positive based on the delay bound condition (20), we have

$$\begin{aligned}\dot{V}(t) &\leq -\left(\frac{1}{2}\epsilon - \sigma\ D\rho_2(\epsilon)\right)\eta^T(t)\overline{P}\eta(t)\\ &\leq -\left(\frac{1}{2}\epsilon - \frac{k\lambda_N D^2\rho_2(\epsilon)}{1-3k^2\lambda_N^2 D^2}\right)\eta^T(t)\overline{P}\eta(t).\end{aligned} \tag{43}$$

Note that the delay bound condition (20) implies $1/2\epsilon > k\lambda_N D^2\rho_2(\epsilon)/1-3k^2\lambda_N^2 D^2$, and thus, $\dot{V}(t) < 0$. By the Lyapunov–Krasovskii stability theorem [45], asymptotic omniscience of the distributed observer (2) has been achieved under the gain parameter choice satisfying (18)–(21). □

Remark 4. According to the sufficient condition (18)–(21) for the distributed observer in Theorem 1, the parameter ϵ needs to be chosen before we can construct the observers. Firstly, ϵ is a parameter in ARE (7), and $(d/\mathrm{d}\epsilon)P(\epsilon) > 0$. Therefore, ϵ is usually a small parameter to keep H_i as a small feedback gain. Secondly, there is a low bound for ϵ according to (6) in Lemma 2. Moreover, note that ϵ affects the upper bound of the time-varying delay $d(t)$ according to (20) and affects the convergence rate of the distributed observers according to (43). These tradeoffs need to be considered when selecting a proper ϵ.

Remark 5. A similar setting of state estimation problem was discussed in [39], where the time delay was assumed to be constant and known. However, the above proof demonstrates that the stability analysis of the observer error system (4) involves only the information on the bound D of the time-varying delay $d(t)$, without requiring its exact knowledge. Therefore, our proposed distributed observer scheme is applicable to more general cases (e.g., fast-varying communication delay) and is also robust to the uncertainty in the communication time delay.

For the special case with no communication time delays, the distributed observer (2) becomes

$$\dot{\widehat{x}}_i = A\widehat{x}_i + H_i\left(y_i - C_i\widehat{x}_i\right) + k\sum_{j\in N_i} g_{ij}\left(\widehat{x}_j(t) - \widehat{x}_i(t)\right). \tag{44}$$

Regarding this circumstance, a sufficient condition for the distributed observer (44) to achieve asymptotic omniscience is described by the following corollary.

Corollary 1. *Consider the linear system (1) under Assumptions 1 and 2. Then, the distributed observer (44) achieves asymptotic omniscience with the feedback gain $H_i = (N/2)P(\epsilon)C_i^T$ if*

$$k > \frac{\rho_1(\epsilon)}{2\lambda_2}, \tag{45}$$

where $P(\epsilon)$ is the solution to (7), $\lambda_2 > 0$ is the second smallest eigenvalue of L, and $\rho_1(\epsilon)$ is as in (19).

Proof. Without the communication delay, the error system between (1) and (44) reduces to

$$\dot{\eta}(t) = \left(I_N \otimes A - \overline{HC} - k\left(L\otimes I_n\right)\right)\eta(t), \tag{46}$$

where $\overline{H}$ and $\overline{C}$ are as defined in (4). To prove the omniscience of this distributed observer, a Lyapunov function for (46) can be defined as follows:

$$V(t) = \eta^T(t)\overline{P}\eta(t), \tag{47}$$

where $\overline{P} = I_N \otimes P^{-1}$.

For the feedback gain, $H_i = (N/2)P(\epsilon)C_i^T$, and the coupling strength k in (45), the fact that the error system (46) is asymptotically stable is shown similar with the corresponding proof in Theorem 1. □

Remark 6. Both observers (2) and (44) can be implemented in a distributed manner, since they compute the state estimates of the local observer by the information exchanged with neighbors. However, the choice of the coupling gain k in (19) and (45) needs some global information (e.g., λ_2) of the communication graph $\mathcal{G}$, whose design is not a decentralized one. Without the global knowledge of the communication graph, one can take k sufficiently large to make inequality (19) or (45) hold. Another feasible method to select the coupling strength k in (2) and (44) is the adaptive regulation law [46, 47].

4. Numerical Example

Consider the following system matrices for system (1):

$$\begin{aligned}A &= \begin{bmatrix} -0.04 & 0 & 2 & 0\\ 0 & 0.02 & 0 & -1\\ 0 & 0 & 0 & -1\\ 0 & 0 & 1 & 0\end{bmatrix},\\ C &= \begin{bmatrix} C_1\\ C_2\\ C_3\end{bmatrix},\end{aligned} \tag{48}$$

where $C_1 = [2,0,0,0]$, $C_2 = [0,2,0,0]$, and $C_3 = [0,0,0,2]$. The open-loop dynamics of this system is exponentially unstable, since system matrix A has eigenvalues $\sigma(A) = \{-0.04, 0.02, \pm i\}$.

By direct calculation, we have (A,C) is detectable, $\mathrm{tr}(A) = -0.02$, and $\mathrm{tr}(A^2) = -2.0$. It can be easily verified that none of (A,C_i) is detectable. Hence, we construct a network of $N = 3$ local observers, and observer i can obtain $y_i = C_i x$ instead of $y = Cx$. The relationship among these observers is given in Figure 1. For simplicity, here, we suppose that all the weights are set as 2. The corresponding Laplacian matrix is given by

$$L = 2\begin{bmatrix} 1 & -1 & 0\\ -1 & 2 & -1\\ 0 & -1 & 1\end{bmatrix}, \tag{49}$$

whose eigenvalues are $\lambda_1(L) = 0, \lambda_2(L) = 2$, and $\lambda_3(L) = 6$.

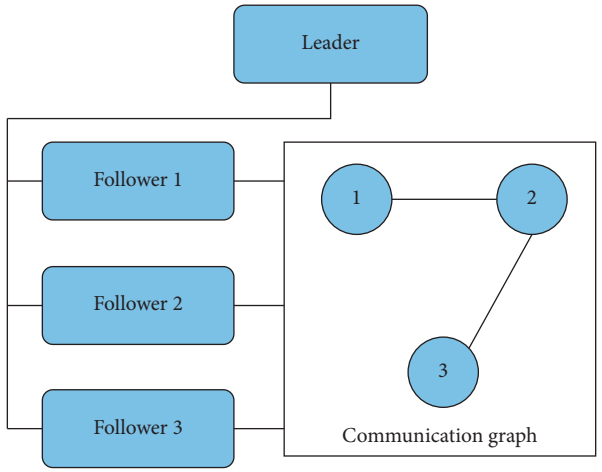


Figure 1: Framework of the networked distributed observer consisting of $N = 3$ local observers.

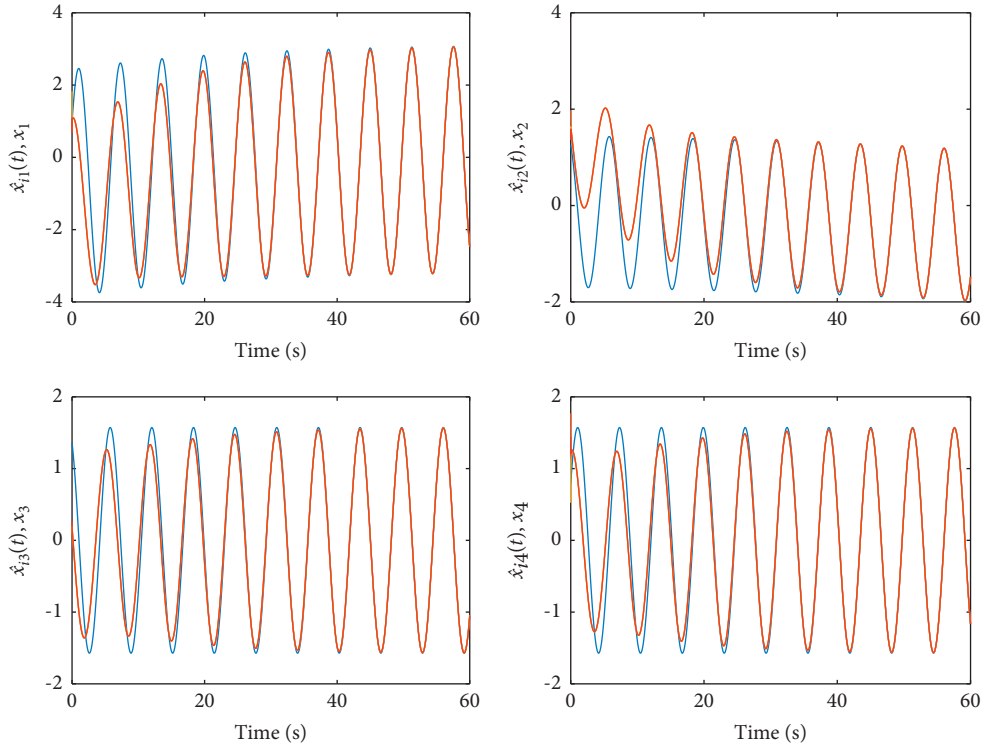


Figure 2: Observer states $\hat{x}_i(t)(i = 1, 2, 3)$ and the plant state $x(t)$ for time-varying communication delay.

By selecting the parameter $\epsilon = 0.18$, the solution of ARE (7) is calculated as

$$P = \begin{bmatrix} 0.0852 & 0.0001 & 0.0060 & 0.0299 \\ 0.0001 & 0.0696 & 0.0148 & -0.0020 \\ 0.0060 & 0.0148 & 0.0156 & 0.0009 \\ 0.0299 & -0.0020 & 0.0009 & 0.0151 \end{bmatrix}. \quad (50)$$

In the delayed case, by calculating (18) and (19), we have the coupling strength $k > 44.40$. Let $k = 44.5$, the corresponding delay bound can be calculated by (20). Figures 2 and 3 show the trajectories and the observer errors of the distributed observers with a time-varying communication delay $d(t) = 0.1 \sin (t)^2$. According to the results, we can find that all observers follow the states of the linear plant with very small errors. For the delay-free case, by (41), we have $k > 22.2$, and other parameters are the same with the delayed case. With $k = 22.3$, Figures 4 and 5 show the trajectories and the observer errors of the distributed observers without communication delay. The delay-free case allows a smaller coupling strength with satisfied estimation results. The results of the simulations verify the effectiveness of the distributed observers with/without communication delay.

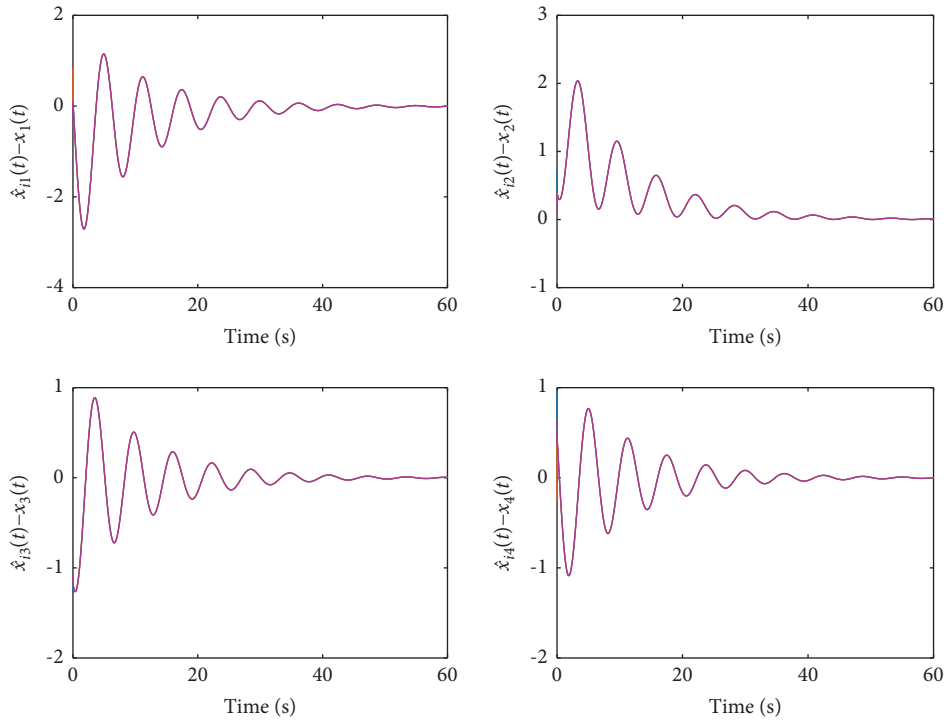


Figure 3: Observer errors $\eta_i(t) = \hat{x}_i(t) - x(t)\,(i = 1, 2, 3)$ for time-varying communication delay.

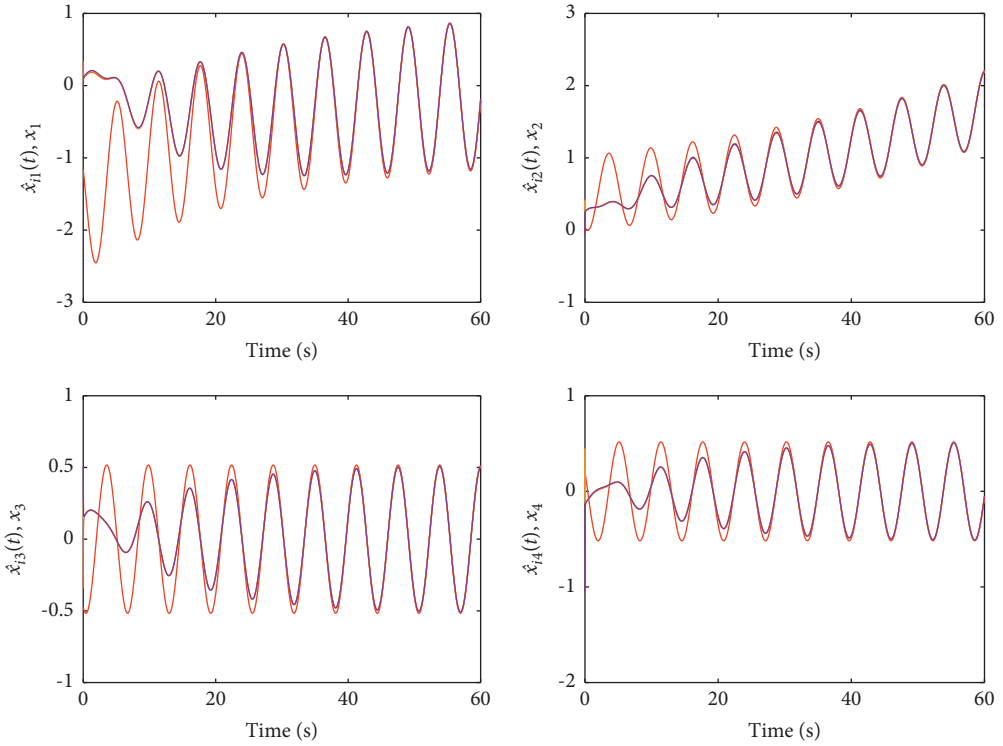


Figure 4: Observer states $\hat{x}_i(t)\,(i = 1, 2, 3)$ and the plant state $x(t)$ for the delay-free case.

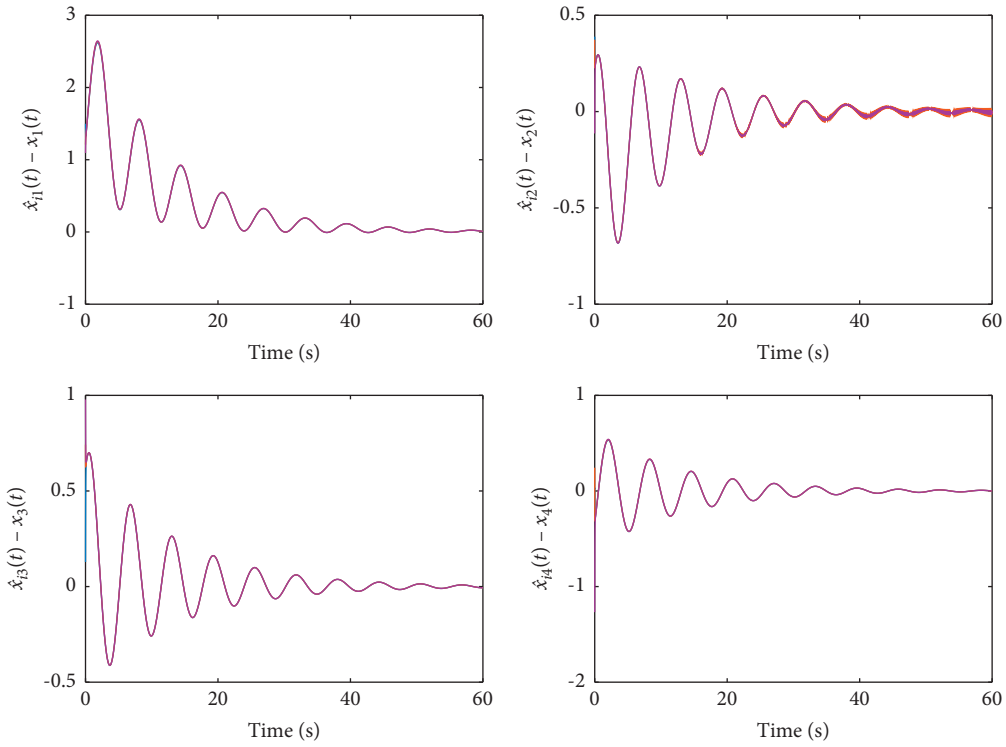


FIGURE 5: Observer errors $\eta_i(t) = \hat{x}_i(t) - x(t)\,(i = 1, 2, 3)$ for the delay-free case.

5. Conclusion

In this paper, we have dealt with the state estimation problem of general continuous-time linear systems, which is solved by the distributed Luenberger observer technique. The structure of the proposed distributed observers comprises networked local observers, in which observers are constructed locally based on its own output measurement and its neighbors according to a given connected communication graph. Remarkably, a network-induced time-varying communication delay has been considered. Two gain parameters, i.e., the gain matrix for local output measurement and the coupling gain for information exchange, have been designed. The local observer gain has been derived by solving a parametric algebraic Riccati equation, while the coupling strength has been determined based on the Lyapunov–Krasovskii stability theorem. Moreover, a bound on the time-varying delay has been also given under which each local observer in the network asymptotically resolves the entire state of the plant.

It should be noted that we focused on linear systems in this paper, and it should be promising to consider some practical issues such as nonlinear plant dynamics, heterogeneous communication delays, transmission errors, and directed communication channels.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China, under Grants 62103115, 11771102, and 61803090, Natural Science Foundation of Guangdong Province, under Grants 2019A1515012109 and 2021A1515011636, Young innovative talents project of general colleges and universities in Guangdong Province, under Grant 2019KQNCX107, Foundation of Key Laboratory of System Control and Information Processing, and Ministry of Education, China.

References

[1] W. Ren and R. W. Beard, *Distributed Consensus in Multi-Vehicle Cooperative Control*, Springer, New York, NY, USA, 2008.

[2] J. Wang, K. Chen, Q. Liu, and Q. Ma, "Observer-based adaptive consensus tracking control for nonlinear multi-agent systems with actuator hysteresis," *Nonlinear Dynamics*, vol. 95, no. 3, pp. 2181–2195, 2019.

[3] D. Luenberger, "Observers for multivariable systems," *IEEE Transactions on Automatic Control*, vol. 11, no. 2, pp. 190–197, 1966.

[4] D. Luenberger, "An introduction to observers," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 596–602, 1971.

[5] Y. Jin, W. Qi, and G. Zong, "Finite-time synchronization of delayed semi-markov neural networks with dynamic event-triggered scheme," *International Journal of Control, Automation and Systems*, vol. 19, no. 6, pp. 2297–2308, 2021.

[6] W. Qi, X. Yang, J. H. Park, J. Cao, and J. Cheng, "Fuzzy smc for quantized nonlinear stochastic switching systems with

semi-markovian process and application," *IEEE Transactions on Cybernetics*, pp. 1–10, 2021.

[7] W. Qi, G. Zong, and W. X. Zheng, "Adaptive event-triggered smc for stochastic switching systems with semi-markov process and application to boost converter circuit model," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 68, no. 2, pp. 786–796, 2020.

[8] J. Wang, H. Zhang, K. Ma, Z. Liu, and C. P. Chen, "Neural adaptive self-triggered control for uncertain nonlinear systems with input hysteresis," *IEEE Transactions on Neural Networks and Learning Systems*, 2021.

[9] W. Qi, Y. Hou, G. Zong, and C. K. Ahn, "Finite-time event-triggered control for semi-markovian switching cyber-physical systems with fdi attacks and applications," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 68, no. 6, pp. 2665–2674, 2021.

[10] H.-N. Wu and H.-X. Li, "Robust adaptive neural observer design for a class of nonlinear parabolic pde systems," *Journal of Process Control*, vol. 21, no. 8, pp. 1172–1182, 2011.

[11] X. Zhang, Y. Han, L. Wu, and Y. Wang, "State estimation for delayed genetic regulatory networks with reaction-diffusion terms," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 2, pp. 299–309, 2018.

[12] A. M. Mohammed, H. Wang, and Y. Tian, "State observer for linear system with unknown input disturbance and sampled and delayed output with measurement noise," *Transactions of the Institute of Measurement and Control*, vol. 41, no. 3, pp. 749–759, 2019.

[13] R. Mohajerpoor, L. Shanmugam, H. Abdi, S. Nahavandi, and J. H. Park, "Delay-dependent functional observer design for linear systems with unknown time-varying state delays," *IEEE Transactions on Cybernetics*, vol. 48, no. 7, pp. 2036–2048, 2018.

[14] J. Wang, Y. Huang, T. Wang, C. Zhang, and Y. h. Liu, "Fuzzy finite-time stable compensation control for a building structural vibration system with actuator failures," *Applied Soft Computing*, vol. 93, Article ID 106372, 2020.

[15] X. Zhang and F. L. Lewis, "Cooperative output regulation of heterogeneous multi-agent systems based on passivity," *International Journal of Systems Science*, vol. 49, no. 16, pp. 3418–3430, 2018.

[16] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proceedings of the Fourth International Symposium on Information Processing in Sensor Networks, 2005*, pp. 63–70, Boise, ID, USA, April 2005.

[17] R. Olfati-Saber, "Distributed kalman filtering for sensor networks," in *Proceedings of the 2007 46th IEEE Conference on Decision and Control*, pp. 5492–5498, New Orleans, LA, USA, December 2007.

[18] U. A. Khan and J. M. F. Moura, "Distributing the kalman filter for large-scale systems," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4919–4935, 2008.

[19] M. V. Subbotin and R. S. Smith, "Design of distributed decentralized estimators for formations with fixed and stochastic communication topologies," *Automatica*, vol. 45, no. 11, pp. 2491–2501, 2009.

[20] D. Viegas, P. Batista, P. Oliveira, and C. Silvestre, "Decentralized observers for position and velocity estimation in vehicle formations with fixed topologies," *Systems & Control Letters*, vol. 61, no. 3, pp. 443–453, 2012.

[21] W. Zhou, J. Fu, H. Yan, X. Du, Y. Wang, and H. Zhou, "Event-triggered approximate optimal path-following control for unmanned surface vehicles with state constraints," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–15, 2021.

[22] V. Delouille, R. Neelamani, and R. Baraniuk, "Robust distributed estimation in sensor networks using the embedded polygons algorithm," in *Proceedings of the Third International Symposium on Information Processing in Sensor Networks, 2004*, pp. 405–413, Berkeley, CA, USA, April 2004.

[23] R. Olfati-Saber, "Distributed kalman filter with embedded consensus filters," in *Proceedings of the 44th IEEE Conference on Decision and Control*, pp. 8179–8184, Seville, Spain, December 2005.

[24] S. Park and N. C. Martins, "An augmented observer for the distributed estimation problem for LTI systems," in *Proceedings of the 2012 American Control Conference (ACC)*, pp. 775–6780, Montreal, QC, Canada, June 2012.

[25] S. Park and N. C. Martins, "Design of distributed LTI observers for state omniscience," *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 561–576, 2017.

[26] H. Zhu, K. Liu, J. Lü, Z. Lin, and Y. Chen, "On the cooperative observability of a continuous-time linear system on an undirected network," in *Proceedings of the 2014 International Joint Conference on Neural Networks (IJCNN)*, pp. 2940–2944, Beijing, China, July 2014.

[27] K. Liu, H. Zhu, and J. Lü, "Cooperative stabilization of a class of LTI plants with distributed observers," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 7, pp. 1891–1902, 2017.

[28] K. Liu, Y. Chen, Z. Duan, and J. Lü, "Cooperative output regulation of LTI plant via distributed observers with local measurement," *IEEE Transactions on Cybernetics*, vol. 48, no. 7, pp. 2181–2191, 2018.

[29] K. Chen, J. Wang, Z. Zhao, G. Lai, and Y. Lyu, "Output consensus of heterogeneous multiagent systems: a distributed observer-based approach," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–7, 2020.

[30] T. Kim, H. Shim, and D. D. Cho, "Distributed luenberger observer design," in *Proceedings of the 2016 IEEE 55th Conference on Decision and Control (CDC)*, pp. 6928–6933, Las Vegas, NV, USA, December 2016.

[31] W. Han, H. L. Trentelman, Z. Wang, and Y. Shen, "A simple approach to distributed observer design for linear systems," *IEEE Transactions on Automatic Control*, vol. 64, no. 1, pp. 329–336, 2019.

[32] Z. Zhiwei Gao and S. X. Ding, "State and disturbance estimator for time-delay systems with application to fault estimation and signal compensation," *IEEE Transactions on Signal Processing*, vol. 55, no. 12, pp. 5541–5551, 2007.

[33] Z. Zhiwei Gao, T. Breikin, and H. Hong Wang, "Reliable observer-based control against sensor failures for systems with time delays in both state and input," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 38, no. 5, pp. 1018–1029, 2008.

[34] J. R. Moyne and D. M. Tilbury, "The emergence of industrial control networks for manufacturing control, diagnostics, and safety data," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 29–47, 2007.

[35] F. Liu, H. Gao, J. Qiu, S. Yin, J. Fan, and T. Chai, "Networked multirate output feedback control for setpoints compensation and its application to rougher flotation process," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 1, pp. 460–468, 2014.

[36] H. J. Savino, C. R. P. dos Santos, F. O. Souza, L. C. A. Pimenta, M. de Oliveira, and R. M. Palhares, "Conditions for consensus of multi-agent systems with time-delays and uncertain

switching topology," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 2, pp. 1258–1267, 2016.

[37] K. Chen, J. Wang, Y. Zhang, and Z. Liu, "Leader-following consensus for a class of nonlinear strick-feedback multiagent systems with state time-delays," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 7, pp. 2351–2361, 2020.

[38] Z. Gao, "Estimation and compensation for lipschitz nonlinear discrete-time systems subjected to unknown measurement delays," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 9, pp. 5950–5961, 2015.

[39] K. Liu, J. Lü, and Z. Lin, "Design of distributed observers in the presence of arbitrarily large communication delays," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 9, pp. 4447–4461, 2018.

[40] H. Basu and S. Y. Yoon, "Distributed state estimation by a network of observers under communication and measurement delays," *Systems & Control Letters*, vol. 133, 2019.

[41] C. Godsil and G. F. Royle, *Algebraic Graph Theory*, Springer Science & Business Media, Berlin, Germany, 2001.

[42] B. Bin Zhou, Z. Zongli Lin, and G.-R. Guang-Ren Duan, "Properties of the parametric Lyapunov equation-based low-gain design with applications in stabilization of time-delay systems," *IEEE Transactions on Automatic Control*, vol. 54, no. 7, pp. 1698–1704, 2009.

[43] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM, Philadelphia, PA, USA, 1994.

[44] K. Gu, J. Chen, and V. L. Kharitonov, *Stability of Time-Delay Systems*, Springer Science & Business Media, Berlin, Germany, 2003.

[45] E. Fridman, *Introduction to Time-Delay Systems: Analysis and Control*, Springer, Berlin, Germany, 2014.

[46] Z. Li, W. Ren, X. Liu, and M. Fu, "Consensus of multi-agent systems with general linear and Lipschitz nonlinear dynamics using distributed adaptive protocols," *IEEE Transactions on Automatic Control*, vol. 58, no. 7, pp. 1786–1791, 2013.

[47] W. Yu, W. Ren, W. X. Zheng, G. Chen, and J. Lü, "Distributed control gains design for consensus in multi-agent systems with second-order nonlinear dynamics," *Automatica*, vol. 49, no. 7, pp. 2107–2115, 2013.

Hindawi
Complexity
Volume 2021, Article ID 8553689, 12 pages
https://doi.org/10.1155/2021/8553689

Research Article

Optimal Liquidation Behaviour Analysis for Stochastic Linear and Nonlinear Systems of Self-Exciting Model with Decay

Jiangming Ma [1] **and Xiankang Luo** [2]

[1]*School of Economics, Xihua University, Chengdu, Sichuan 610039, China*
[2]*Faculty of Science, Yibin University, Yibin, Sichuan 644000, China*

Correspondence should be addressed to Xiankang Luo; xkluo1978@163.com

Received 21 July 2021; Revised 20 August 2021; Accepted 21 August 2021; Published 2 September 2021

Academic Editor: Yue Song

When the market environment changes, we extend the self-exciting price impact model and further analysis of investors' liquidation behaviour. It is assumed that the model is accompanied by an exponential decay factor when the temporary impact and its coefficient are linear and nonlinear. Using the optimal control method, we obtain that the optimal liquidation behaviours satisfy the second-order nonlinear ODEs with variable coefficients in the case of linear and nonlinear temporary impact. Next, we solve the ODEs and get the form of the investors' optimal liquidation behaviour in four cases. Furthermore, we prove the decreasing properties of the optimal liquidation behaviour under the linear temporary impact. Through numerical simulation, we further explain the influence of the changed parameters ρ, a, b, x, and α on the investors' liquidation strategy X_t in twelve scenarios. Some interesting properties have been found.

1. Introduction

Bertsimas and Lo [1] proposed the optimal execution model which has a linear and discrete price impact model in a fixed time. They got the optimal liquidation behaviours by an optimal control method in some cases. Two years later, Almgren and Chriss [2] used the minimization of mean-variance function to think the expected costs and risks. They obtained an optimal closed-form solution of static behaviour. When this model is proposed, it is quickly studied by a large number of scholars in different situations. Almgren [3] further extended the previous model which is assumed that the temporary impact is nonlinear and got the form of the optimal liquidation behaviour.

Huberman and Stanzl [4] studied the linear and nonlinear fixed impact and found that there exists no arbitrage phenomenon only when the market impact is linear. Gatheral and Schied [5] provided the behaviour of optimal expectation transaction cost which is satisfied with a HJB equation. In this paper, the optimal behaviours of two benchmark models are obtained, and it is found that the optimal behaviours of the two models are irrelevant unless there are extremely unrealistic situations and parameters. Gatheral et al. [6] researched the optimal trading strategy in which the temporary market impact has a decay factor. They proved some properties of the optimal investment liquidation behaviour and gave some specific form of the optimal behaviour under some special decay functions. Lehalle and Neuman [7] introduced Markovian signals under the optimal trading framework proposed by Gatheral et al. [6] and got the existence and uniqueness of optimal trading behaviour. Moreover, the authors gave the optimal investment behaviour of investors under the condition of Ornstein Uhlenbeck signal and temporary market impact with exponential decay. Based on the premise of no dynamic arbitrage, Gatheral [8] studied the optimal investment behaviour under the condition of market impact with decay factors. He gave the specific form of the optimal investment behaviour from the case of fixed and temporary market impact with different decay factors. Curato et al. [9] investigated the optimal execution of a large trade when the model has nonlinear transient impact and decay factors.

They got the specific forms of the optimal strategy in some special cases. Barger and Lorig [10] investigated the optimal liquidation behaviours when the parameters of price impact are stochastic processes. They obtained the form of optimal liquidation behaviours by solving the HJB equation. Based on the Bayesian learning and dynamic programming techniques, Bismuth et al. [11] researched the optimal portfolio choice, portfolio liquidation, and portfolio transition problems. Through the optimal control method, they found that the optimal behaviour is the solution of the HJB equation. When the Almgren–Chriss model has n risk-averse agents, Schied and Zhang [12] discussed the Nash equilibria and closed-form solutions of optimal liquidation by solving the ODE equation. Based on the Almgren–Chriss framework, Vaes and Hauser [13] considered a volume uncertainty model and proved that a risk-averse trader has benefit when he delays trades. Meanwhile, Ekren and Nadtochiy [14] researched the utility-based optimal hedging behaviour for a European-type option. From solving the solution of PDE, they obtained the optimal behaviours.

Based on Almgren–Chriss framework [2], a self-exciting price impact model is proposed by Caye and Muhle-Karbe [15]. They considered that a large sell order creates persistent selling pressure which not only incurs price impact but also increases the execution costs. However, they only discussed that the temporary impact is a linear function. When there exists a decay factor in the market, temporary impact and its coefficient are nonlinear functions, and the relevant problems have not been studied. Following the relevant literature, we add the exponential decay factor. Furthermore, when the temporary impact and its coefficient are linear and nonlinear, the optimal liquidation behaviours are obtained by the second-order nonlinear ODEs with variable coefficients.

The paper is organized as follows. In Section 2, we introduce the unaffected price process, the self-exciting price model, execution cost, and the objective function. In Section 3, we present our main results: utilizing the optimal control method, we give the form of optimal liquidation behaviour when temporary impact and its coefficient are linear and nonlinear and discuss the properties of optimal liquidation behaviours. In Section 4, from seven different numerical examples, we discuss the effect of parameter changes on liquidation behaviours and show the corresponding financial interpretations. In Section 5, we conclude our research and give the future research direction.

2. Statement of Background

The following assumptions of the self-exciting price impact model are given. The trading time is fixed in $[0,T]$. X_t represents the holdings at time t, where $X_0 = x$ and $X_T = 0$. Meanwhile, X_t which is absolutely continuous and bounded with derivative $\dot{X}_t$ and $X_t = x + \int_0^t \dot{X}_t \mathrm{d}t$ satisfies $\int_0^T (\dot{X}_t)^2 \mathrm{d}t < \infty$.

The self-exciting price impact model contains unaffected price process as follows:

$$S_t^0 \coloneqq \sigma dW_t, \tag{1}$$

where W_t is the standard Brownian motion. Thus, the specific form of the self-exciting price impact is assumed to be

$$S_t \coloneqq S_t^0 + (a + b(x - X_t))\dot{X}_t, \quad t \in [0,T], \tag{2}$$

where $a > 0$ and $b > 0$.

Therefore, the total execution cost of investors is

$$\mathscr{C}(X) \coloneqq xS_0 + \int_0^T \dot{X}_t \, S_t \mathrm{d}t. \tag{3}$$

Only when the execution cost is the minimum, investors can get the maximum returns. So, the objective function of investors is

$$\text{minimize}\, \mathbb{E}[\mathscr{C}(X)]. \tag{4}$$

3. Main Results

Caye and Muhle-Karbe [15] only discussed liquidation behaviours when the coefficient of temporary impact is linear. When there exists a decay factor in the market, the relevant literature has not been studied. In this paper, we introduce the classical exponential decay factor in the self-exciting price impact model. Next, we study the optimal liquidation behaviour of investors when the market has exponential decay under the four cases.

Theorem 1. *When there is equation (2) with exponential decay $e^{-\rho t}$, the mean optimization strategy is the unique solution of the following nonlinear ODE with variable coefficients:*

$$2[a + b(x - X_t)]\ddot{X}_t - b\dot{X}_t^2 - 2\rho(a + b(x - X_t))\dot{X}_t = 0, \tag{5}$$

with boundary conditions as follows:

$$\begin{aligned} X_0 &= x, \\ X_T &= 0. \end{aligned} \tag{6}$$

The solution of equation (5) is

$$X_t = \frac{a}{b} + x - \frac{C_1}{b} e^{\int_0^t \left[C_2 + (1/2) \int_0^t e^{\rho\nu} \mathrm{d}\nu \right]^{-1} \mathrm{d}\nu}. \tag{7}$$

Proof. When the market has an exponential decay, equation (2) is

$$S_t \coloneqq S_t^0 + e^{-\rho t}(a + b(x - X_t))\dot{X}_t, \quad t \in [0,T]. \tag{8}$$

From equations (3) and (8), we get

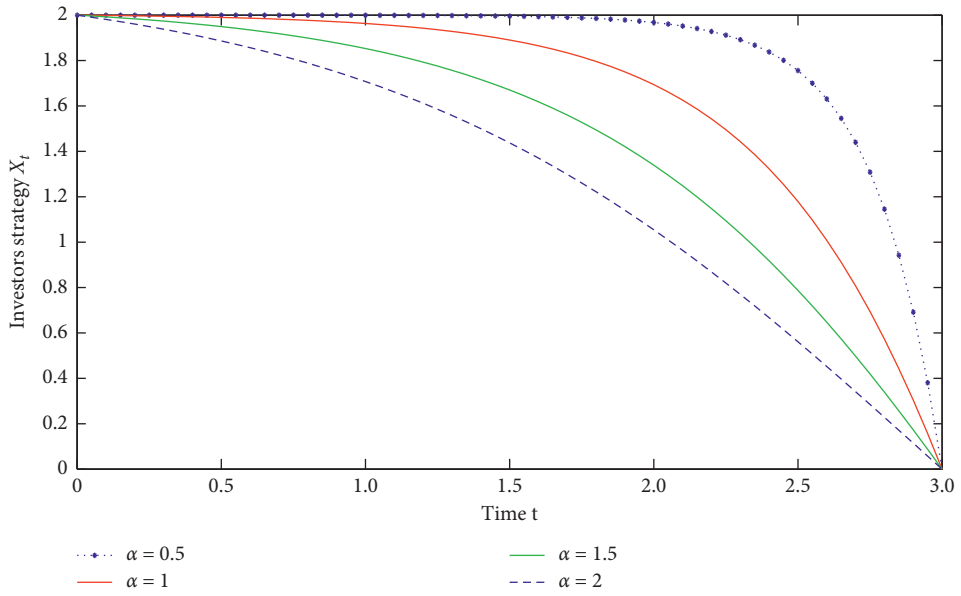


FIGURE 1: Influence of α change on equation (26) when $\rho = 2$, $a = 0.6$, $b = 0.5$, and $x = 2$.

$$\begin{aligned}\mathscr{C}(X) &:= xS_0 + \int_0^T \dot{X}_t\, S_t \mathrm{d}t = xS_0 \\ &+ \int_0^T \dot{X}_t \left(S_t^0 + e^{-\rho t}(a + b(x - X_t))\dot{X}_t\right)\mathrm{d}t \\ &= \int_0^T \sigma X_t \mathrm{d}W_t + \int_0^T e^{-\rho t}(a + b(x - X_t))\dot{X}_t^{\,2}\mathrm{d}t.\end{aligned} \tag{9}$$

Combined with the properties of Ito integral and equation (4), the above equation is changed to be

$$\text{minimize}\, \mathbb{E}[\mathscr{C}(X)] = \int_0^T e^{-\rho t}(a + b(x - X_t))\dot{X}_t^2\, \mathrm{d}t. \tag{10}$$

In order to get the solution of equation (10), we use the Euler–Lagrange equation to get the following second-order nonlinear ODE with variable coefficients:

$$2[a + b(x - X_t)]\ddot{X}_t - b\dot{X}_t^2 - 2\rho(a + b(x - X_t))\dot{X}_t = 0. \tag{11}$$

Let $Y_t = a + b(x - X_t)$, equation (11) is changed as follows:

$$2Y_t\ddot{Y}_t - \dot{Y}_t^2 - 2\rho Y_t\dot{Y}_t = 0. \tag{12}$$

From the [16], the solution of equation (12) is

$$Y_t = C_1 e^{\int_0^t e^{\rho\nu}\left[C_2 + (1/2)\int_0^t e^{\rho\nu}\mathrm{d}\nu\right]^{-1}\mathrm{d}\nu}. \tag{13}$$

Thus, the solution of equation (5) is

$$X_t = \frac{a}{b} + x - \frac{C_1}{b} e^{\int_0^t e^{\rho\nu}\left[C_2 + (1/2)\int_0^t e^{\rho\nu}\mathrm{d}\nu\right]^{-1}\mathrm{d}\nu}. \tag{14}$$

□

Remark 1. In Theorem 1, we address the optimal liquidation behaviour when there is a self-exciting model with exponential decay and give the specific form of optimal liquidation behaviour. That is, the coefficient of temporary impact is multiplied by the exponential function. In fact, when $\rho = 0$, our model is reduced to the price model of Caye and Muhle-Karbe [15].

Theorem 2. *Under the assumption of Theorem 1, the optimal strategy X_t from equation (14) is deterministic, absolutely continuous, and decreasing.*

Proof

$$\begin{aligned}\mathscr{C}(X) &= \int_0^T e^{-\rho t}(a + b(x - X_t))\dot{X}_t^2\, \mathrm{d}t \\ &= (a + bx)\int_0^T e^{-\rho t}\dot{X}_t^2\, \mathrm{d}t - b\int_0^T e^{-\rho t}X_t\dot{X}_t^2\, \mathrm{d}t \\ &= \mathscr{F}(X).\end{aligned} \tag{15}$$

Let $Y = X - X^*$, then we get

$$
\begin{aligned}
\mathscr{F}(X) &= \mathscr{F}\left(Y + X^*\right) \\
&= (a+bx)\int_0^T e^{-\rho t}\left(\dot{Y}_t + \dot{X}_t^*\right)^2 \mathrm{d}t - b\int_0^T e^{-\rho t}\left(Y_t + X^*\right)\left(\dot{Y}_t + \dot{X}_t^*\right)^2 \mathrm{d}t \\
&= (a+bx)\int_0^T e^{-\rho t}\left(\dot{Y}_t{}^2 + 2\dot{Y}_t\,\dot{X}_t^* + \left(\dot{X}_t^*\right)^2\right)\mathrm{d}t - b\int_0^T e^{-\rho t}\left(Y_t + X^*\right)\left(\dot{Y}_t{}^2 + 2\dot{Y}_t\,\dot{X}_t^* + \left(\dot{X}_t^*\right)^2\right)\mathrm{d}t \\
&= (a+bx)\int_0^T e^{-\rho t}\left(\dot{Y}_t{}^2 + 2\dot{Y}_t\,\dot{X}_t^* + \left(\dot{X}_t^*\right)^2\right)\mathrm{d}t \\
&\quad - b\int_0^T e^{-\rho t}\left(Y_t\dot{Y}_t{}^2 + 2Y_t\dot{Y}_t\,\dot{X}_t^* + Y\left(\dot{X}_t^*\right)^2 + X_t^*\dot{Y}_t{}^2 + 2X_t^*\dot{Y}_t\,\dot{X}_t^* + X_t^*\left(\dot{X}_t^*\right)^2\right)\mathrm{d}t \\
&= \int_0^T e^{-\rho t}\left(a+bx-X_t^*\right)\left(\dot{X}_t^*\right)^2 \mathrm{d}t + \int_0^T e^{-\rho t}\left(a+bx-bX_t^*-bY_t\right)\dot{Y}_t{}^2 \mathrm{d}t \\
&\quad + \int_0^T e^{-\rho t}\left[2(a+bx)\dot{Y}_t\,\dot{X}_t^* - 2bX_t^*\dot{Y}_t\,\dot{X}_t^* - bY_t\left(\dot{X}_t^*\right)^2 - 2bY_t\dot{Y}_t\,\dot{X}_t^*\right]\mathrm{d}t \\
&= \int_0^T e^{-\rho t}\left(a+bx-X_t^*\right)\left(\dot{X}_t^*\right)^2 \mathrm{d}t + \int_0^T e^{-\rho t}\left(a+bx-bX_t^*-bY_t\right)\dot{Y}_t{}^2 \mathrm{d}t \\
&\quad + \int_0^T e^{-\rho t}Y_t\left[\begin{array}{c} 2\rho(a+bx)\dot{X}_t^* - 2(a+bx)\ddot{X}_t^* - 2b\rho X_t^*\dot{X}_t^* \\ +2b\left(\dot{X}_t^*\right)^2 + 2bX_t^*\ddot{X}_t^* - b\left(\dot{X}_t^*\right)^2 \end{array}\right]\mathrm{d}t \\
&\quad - 2b\int_0^T e^{-\rho t}Y_t\dot{Y}_t\,\dot{X}_t^*\,\mathrm{d}t = \int_0^T e^{-\rho t}\left(a+bx-X_t^*\right)\left(\dot{X}_t^*\right)^2 \mathrm{d}t \\
&\quad + \int_0^T e^{-\rho t}\left(a+bx-bX_t^*-bY_t\right)\dot{Y}_t{}^2 \mathrm{d}t \\
&\quad - \int_0^T e^{-\rho t}Y_t\left[2\left(a+b\left(x-X_t^*\right)\right)\ddot{X}_t^* - 2\rho\left(a+b\left(x-X_t^*\right)\right)\dot{X}_t^* - b\left(\dot{X}_t^*\right)^2\right]\mathrm{d}t \\
&\quad - 2b\rho\int_0^T e^{-\rho t}Y_tX_t^*\dot{X}_t^*\,\mathrm{d}t + 2b\int_0^T e^{-\rho t}Y_t\left(\dot{X}_t^*\right)^2 \mathrm{d}t + 2b\int_0^T e^{-\rho t}Y_tX_t^*\ddot{X}_t^*\,\mathrm{d}t \\
&= \int_0^T e^{-\rho t}\left(a+bx-X_t^*\right)\left(\dot{X}_t^*\right)^2 \mathrm{d}t + \int_0^T e^{-\rho t}\left(a+bx-bX_t^*-bY_t\right)\dot{Y}_t{}^2 \mathrm{d}t \\
&\quad - b\rho\int_0^T e^{-\rho t}Y_t^2X_t^*\dot{X}_t^*\,\mathrm{d}t + b\int_0^T e^{-\rho t}Y_t^2\left(\dot{X}_t^*\right)^2 \mathrm{d}t + b\int_0^T e^{-\rho t}Y_t^2\ddot{X}_t^*\,\mathrm{d}t \geq \mathscr{F}\left(X^*\right).
\end{aligned}
\tag{16}
$$

According to the basic assumptions of the model, combined with the following numerical simulation, we find that $\dot{X}_t$ satisfies $\dot{X}_t < 0$. The sixth step is to get through integration by parts and $Y_0 = Y_T = 0$. Because X_t^* satisfies equation (5), the third part of seventh step is removed. Using $x - X_t^* - Y_t = x - X_t \geq 0$, the third part of the seventh step is positive. We obtain the eighth step from integration by parts. Through equation (14) and numerical simulation, $X_t^* \geqslant 0$ is verified. Thus, X_t is decreasing. □

Remark 2. In Theorem 2, we find that the optimal liquidation behaviour still has the same properties as the liquidation behaviour without decay in Caye and Muhle-Karbe [15] when the market has decayed.

In the last part, we discuss the optimal liquidation behaviour of investors when the temporary market impact is a linear function. Meanwhile, Almgren [3], Gatheral [8], Curato et al. [9], and Horst and Naujokat [17] studied the optimal liquidation behaviour of investors when the temporary impact is nonlinear in different cases. When the market is accompanied by a decay factor and the temporary market impact is nonlinear, the optimal liquidation behaviour of investors is not studied by Caye and Muhle-Karbe [15]. Through the study of the existing literature, we assume that the temporary impact is a power function. Thus, when market has an exponential decay, the new price impact model is

$$
S_t := S_t^0 + e^{-\rho t}\left(a + b\left(x - X_t\right)\right)\left(\dot{X}_t\right)^{\alpha}, \quad t \in [0, T]. \tag{17}
$$

Theorem 3. *When there is equation (17) with exponential decay $e^{-\rho t}$, the mean optimization behaviour is the unique*

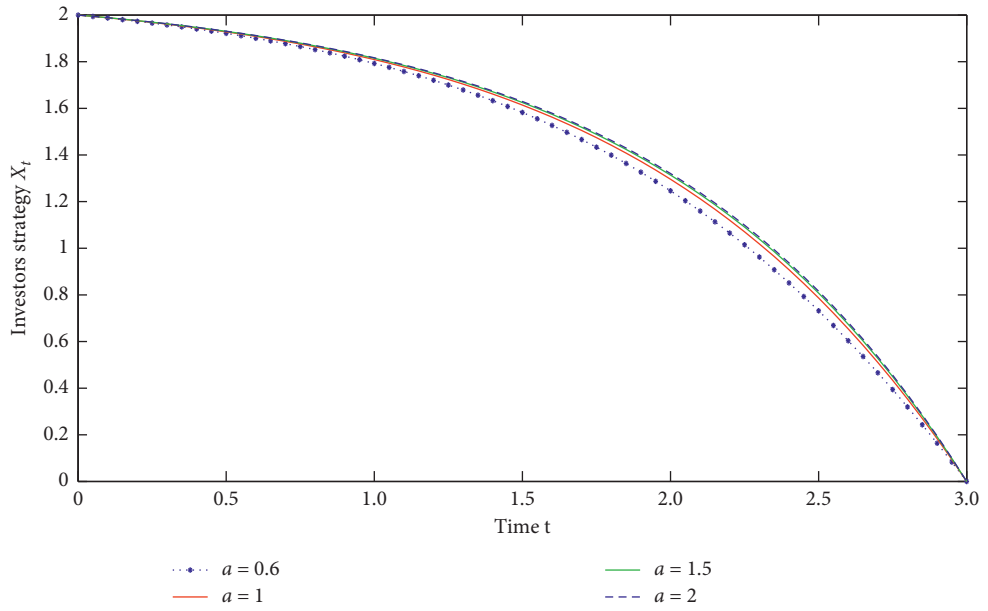


Figure 2: Influence of a change on equation (14) when$\alpha = 1$, $\rho = 1$, $b = 0.5$, and $x = 2$.

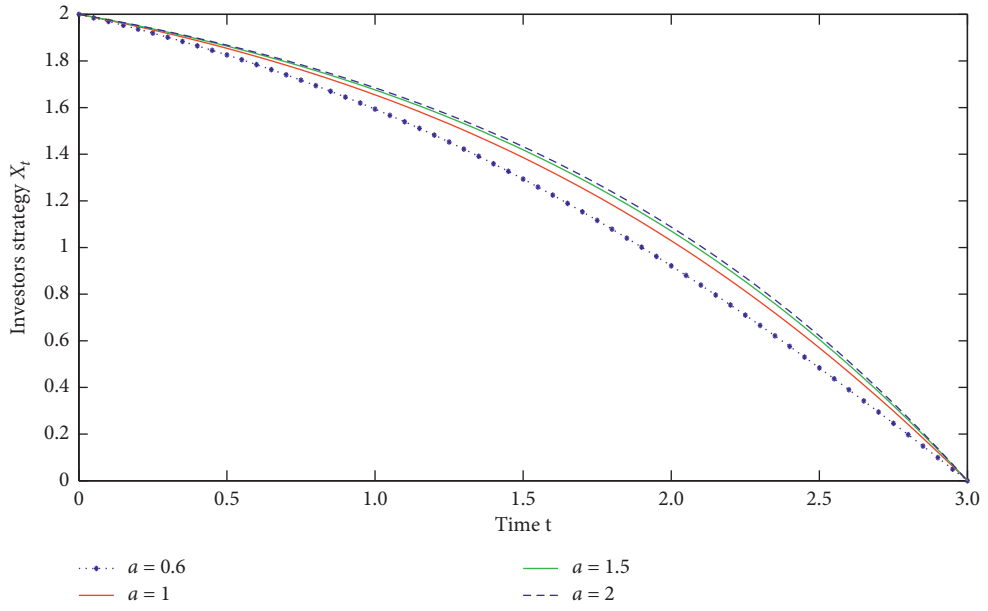


Figure 3: Influence of a change on equation (26) when $\alpha = 1.5$, $\rho = 1$, $b = 0.5$, and $x = 2$.

solution of the following nonlinear ODE with variable coefficients:

$$\begin{aligned}&\alpha(\alpha+1)[a+b(x-X_t)]\ddot{X}_t - b\alpha\dot{X}_t^2\\&-\rho(\alpha+1)(a+b(x-X_t))\dot{X}_t = 0,\end{aligned} \tag{18}$$

with boundary conditions as follows:

$$\begin{aligned}X_0 &= x,\\X_T &= 0.\end{aligned} \tag{19}$$

The solution of equation (18) is

$$X_t = \frac{a}{b} + x - \frac{C_1}{b}e^{\int_0^t e^{(\rho/\alpha)\nu}\left[C_2 + (\alpha/\alpha+1)\int_0^t e^{(\rho/\alpha)\nu}d\nu\right]^{-1}d\nu}. \tag{20}$$

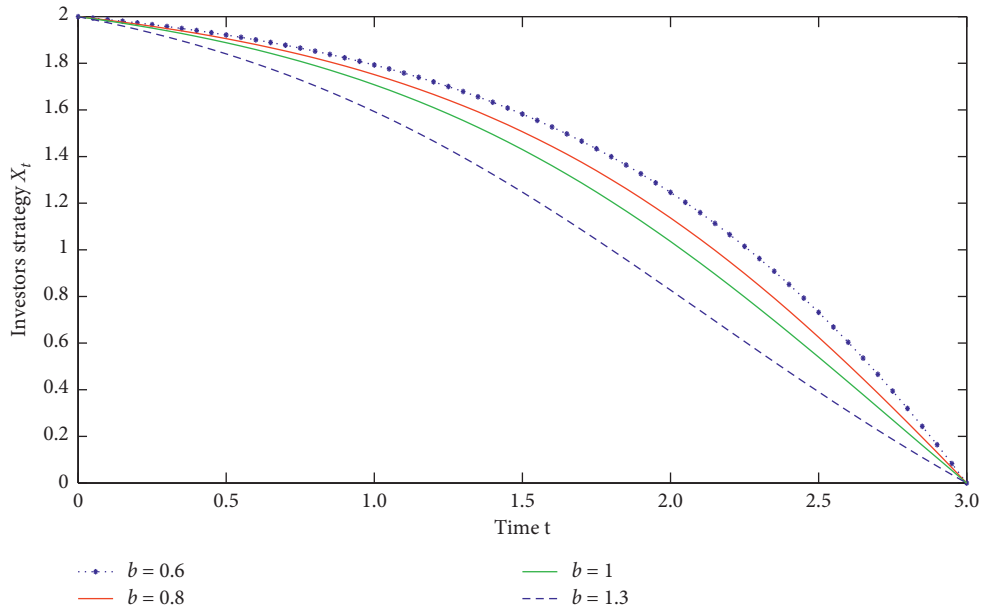


FIGURE 4: Influence of b change on equation (14) when $\alpha = 1$, $\rho = 1$, $a = 0.6$, and $x = 2$.

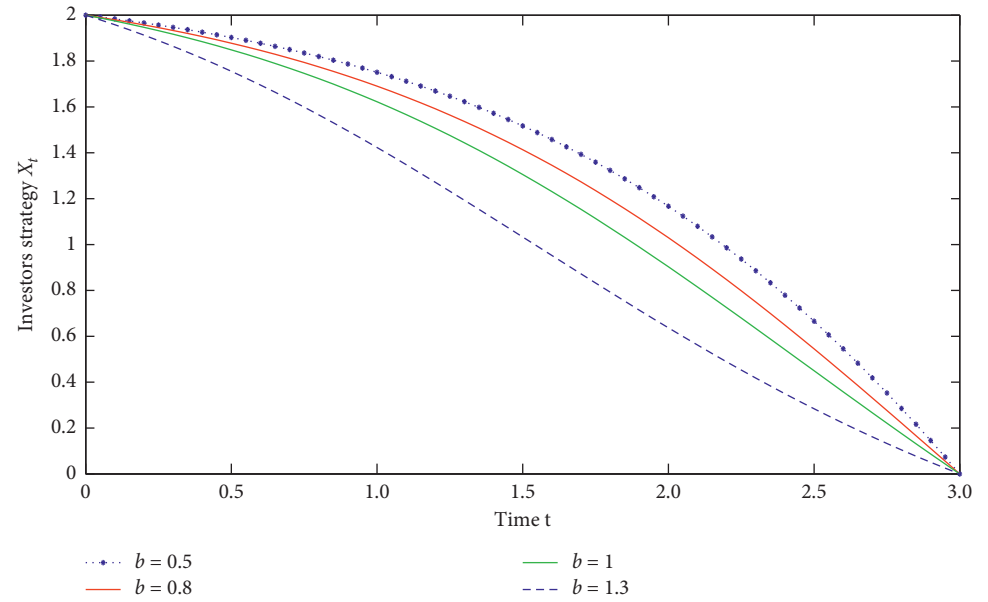


FIGURE 5: Influence of b change on equation (26) when $\alpha = 1.1$, $\rho = 1$, $a = 0.6$, and $x = 2$.

Proof. From equations (3) and (17), we get

$$\begin{aligned}\mathscr{C}(X) &:= xS_0 + \int_0^T \dot{X}_t\, S_t \mathrm{d}t = xS_0 \\ &+ \int_0^T \dot{X}_t \left(S_t^0 + e^{-\rho t}\left(a + b\left(x - X_t\right)\right)\left(\dot{X}_t\right)^{\alpha}\right)\mathrm{d}t \\ &= \int_0^T \sigma X_t \mathrm{d}W_t + \int_0^T e^{-\rho t}\left(a + b\left(x - X_t\right)\right)\dot{X}_t^{\alpha+1}\mathrm{d}t.\end{aligned} \tag{21}$$

According to properties of Ito integral and equation (4), we get

$$\text{minimize}\, \mathbb{E}\left[\mathscr{C}(X)\right] = \int_0^T e^{-\rho t}\left(a + b\left(x - X_t\right)\right)\dot{X}_t^{\alpha+1}\mathrm{d}t. \tag{22}$$

Using the Euler–Lagrange equation, the following second-order nonlinear ODE is obtained with variable coefficients:

$$\begin{aligned}&\alpha(\alpha+1)\left[a + b\left(x - X_t\right)\right]\ddot{X}_t - b\alpha\dot{X}_t^2 \\ &-\rho(\alpha+1)\left(a + b\left(x - X_t\right)\right)\dot{X}_t = 0.\end{aligned} \tag{23}$$

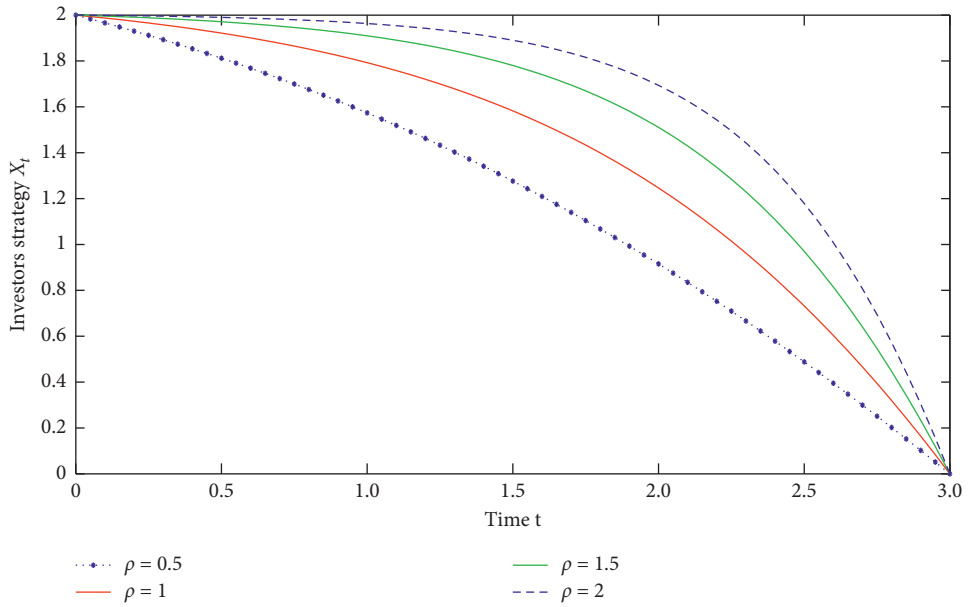


FIGURE 6: Influence of ρ change on equation (14) when $\alpha = 1$, $a = 0.6$, $b = 0.5$, and $x = 2$.

Let $Y_t = a + b(x - X_t)$, equation (23) is changed as follows:

$$\alpha(\alpha+1)Y_t\ddot{Y}_t - \alpha\dot{Y}_t^2 - \rho(\alpha+1)Y_t\dot{Y}_t = 0. \tag{24}$$

From [16], the solution of equation (24) is

$$Y_t = C_1 e^{\int_0^t e^{(\rho/\alpha)\nu}\left[C_2 + (\alpha/\alpha+1)\int_0^t e^{(\rho/\alpha)\nu}d\nu\right]^{-1}d\nu}. \tag{25}$$

Thus, the solution of equation (18) is

$$X_t = \frac{a}{b} + x - \frac{C_1}{b} e^{\int_0^t e^{(\rho/\alpha)\nu}\left[C_2 + (\alpha/\alpha+1)\int_0^t e^{(\rho/\alpha)\nu}d\nu\right]^{-1}d\nu}. \tag{26}$$

From equations (12) and (18), we find that equation (12) is the special form of equation (18) when $\alpha = 1$. □

Remark 3. In Theorem 3, we research the optimal liquidation behaviour when the market has exponential decay and temporary market impact is a nonlinear function. We can easily find that the price impact model of Caye and Muhle-Karbe [15] is our special form in this case of $\alpha = 0$ and $\rho = 0$.

Next, we investigate the liquidation behaviour under the coefficient of temporary impact as $e^{a+b(x-X_t)}$. Thus, equation (4) is changed to be

$$S_t := S_t^0 + e^{a+b(x-X_t)}\dot{X}_t, \quad t \in [0,T]. \tag{27}$$

Theorem 4. *When there is equation (27) with exponential decay $e^{-\rho t}$, the mean optimization behaviour is the unique solution of the following nonlinear ODE with variable coefficients:*

$$2\ddot{X}_t - 2\rho\dot{X}_t - b\dot{X}_t^2 = 0, \tag{28}$$

with boundary conditions as follows:

$$\begin{aligned} X_0 &= x, \\ X_T &= 0. \end{aligned} \tag{29}$$

The solution of equation (28) is

$$\int e^{-(b/2)X_t}dX_t = C_1 e^{\rho t} + C_2. \tag{30}$$

Proof. From equations (3) and (27), we get

$$\begin{aligned} \mathscr{C}(X) &:= xS_0 + \int_0^T \dot{X}_t\, S_t dt \\ &= xS_0 + \int_0^T e^{a+b(x-X_t)}e^{-\rho t}\dot{X}_t^{\,2}dt \\ &= \int_0^T \sigma X_t dW_t + \int_0^T e^{a+b(x-X_t)-\rho t}\dot{X}_t^{\,2}dt. \end{aligned} \tag{31}$$

From the properties of Ito integral and equation (4), we have

$$\text{minimize}\,\mathbb{E}[\mathscr{C}(X)] = \int_0^T e^{a+b(x-X_t)-\rho t}\dot{X}_t^2\,dt. \tag{32}$$

Using the Euler–Lagrange equation, the following nonlinear ODE with variable coefficients is obtained:

$$2\ddot{X}_t - 2\rho\dot{X}_t - b\dot{X}_t^2 = 0. \tag{33}$$

From [16], the solution of equation (33) is

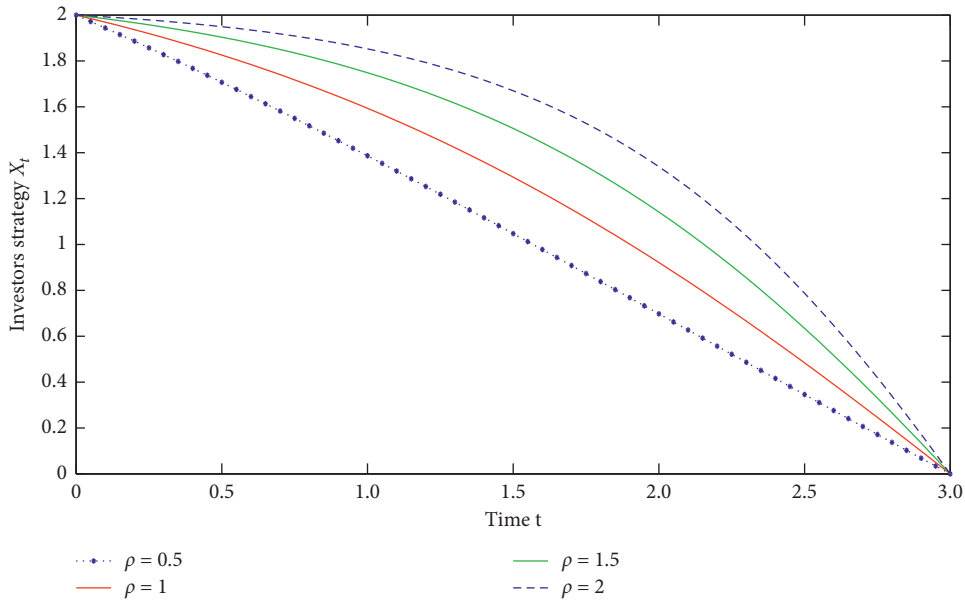


FIGURE 7: Influence of ρ change on equation (26) when $\alpha = 1.5$, $a = 0.6$, $b = 0.5$, and $x = 2$.

$$\int e^{-(b/2)X_t} dX_t = C_1 e^{\rho t} + C_2. \tag{34}$$

□

Remark 4. In Theorem 4, we research the optimal liquidation behaviour when coefficient of temporary impact is a nonlinear function and temporary impact has exponential decay based on Caye and Muhle-Karbe [15].

Theorem 5. *Under the assumption of Theorem 4, the optimal behaviour X_t from equation (34) is deterministic, absolutely continuous, and decreasing.*

Proof

$$\begin{aligned} \mathscr{C}(X) &= \int_0^T e^{a+b(x-X_t)} e^{-\rho t} \dot{X}_t^2 \, dt \\ &= (a+bx) \int_0^T e^{-bX_t - \rho t} \dot{X}_t^2 \, dt = \mathscr{F}(X). \end{aligned} \tag{35}$$

Let $Y = X - X^*$, then we get

$$\begin{aligned} \mathscr{F}(X) &= \mathscr{F}(Y + X^*) \\ &= e^{(a+bx)} \int_0^T e^{(-bX_t^* - bY_t)} e^{-\rho t} \left(\dot{X}_t^* + \dot{Y}_t \right)^2 dt \\ &= e^{(a+bx)} \int_0^T e^{(-bX_t^* - bY_t)} e^{-\rho t} \left(\dot{X}_t^{*\,2} + 2\dot{X}_t^* \dot{Y}_t + \dot{Y}_t^{\,2} \right) dt \\ &= e^{(a+bx)} \int_0^T e^{(-bX_t^* - bY_t)} e^{-\rho t} \dot{X}_t^{*\,2} dt + (a+bx) \int_0^T e^{(-bX_t^* - bY_t)} e^{-\rho t} \left(2\dot{X}_t^* \dot{Y}_t \right) dt \\ &\quad + e^{(a+bx)} \int_0^T e^{(-bX_t^* - bY_t)} e^{-\rho t} \dot{Y}_t^{\,2} dt \cdots {}^{(a+bx)} \int_0^T e^{(-bX_t^* - bY_t)} e^{-\rho t} \dot{X}_t^{*\,3} dt \\ &\geqslant {}^{(a+bx)} \int_0^T e^{-\rho t} e^{-bX_t^*} \dot{X}_t^{*\,3} dt. \end{aligned} \tag{36}$$

From the properties of $\dot{X}_t$, $\dot{X}_t < 0$, and integral, proof of Theorem 5 is easily obtained. □

Remark 5. In Theorem 5, we research a property of the optimal liquidation behaviour when the market has exponential decay and the coefficients of temporary impact are nonlinear functions.

When the temporary impact is a power function, equation (27) is changed to be

$$S_t := S_t^0 + e^{a+b(x-X_t)} \dot{X}_t^{\,\alpha}, \quad t \in [0, T]. \tag{37}$$

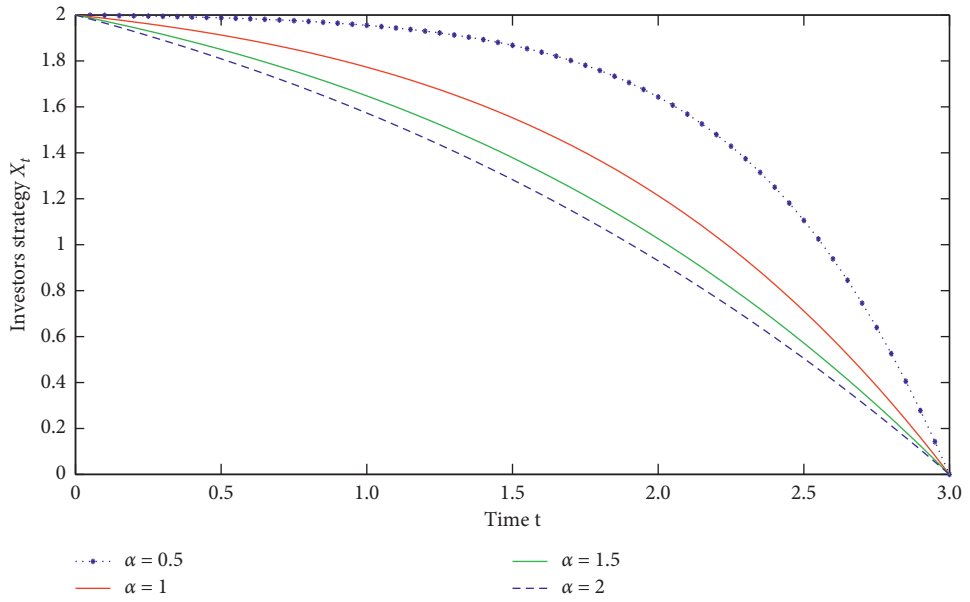


Figure 8: Influence of α change on equation (44) when $\rho = 1$, $b = 0.5$, and $x = 2$.

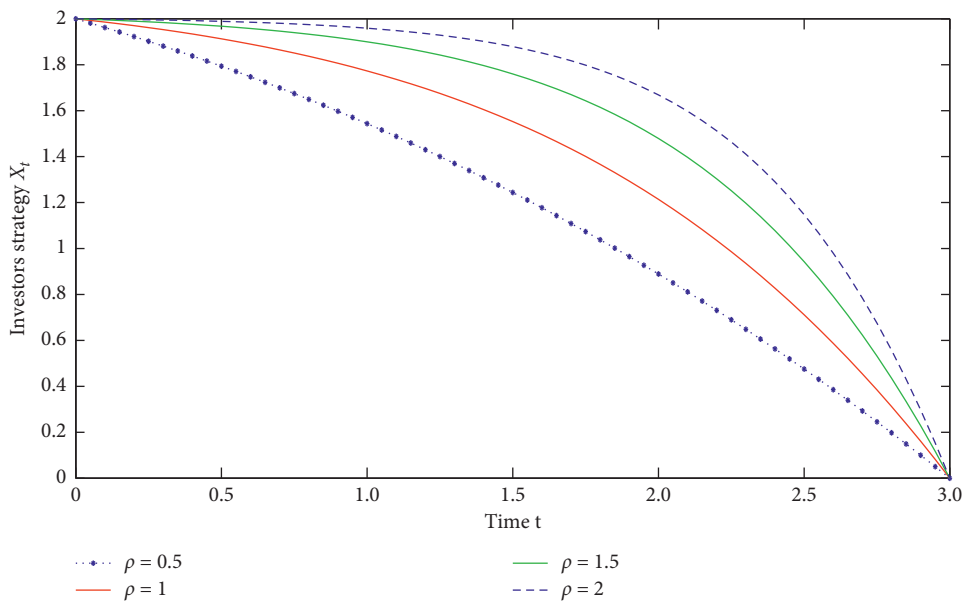


Figure 9: Influence of ρ change on equation (34) when $\alpha = 1$, $b = 0.5$, and $x = 2$.

Theorem 6. *When there is equation (37) with exponential decay $e^{-\rho t}$, the mean optimization behaviour is the unique solution of the following nonlinear ODE with variable coefficients:*

$$\alpha(\alpha+1)\ddot{X}_t - \rho(\alpha+1)\dot{X}_t - b\alpha\dot{X}_t^{\,2} = 0, \tag{38}$$

with two-point boundary conditions as follows:

$$\begin{aligned} X_0 &= x, \\ X_T &= 0. \end{aligned} \tag{39}$$

The solution of equation (38) is

$$\int e^{-(b/\alpha+1)X_t} \mathrm{d}X_t = C_1 e^{(\rho/\alpha)t} + C_2. \tag{40}$$

Proof. From equations (3) and (37), we get

$$\begin{aligned} \mathscr{C}(X) &:= xS_0 + \int_0^T \dot{X}_t\, S_t \mathrm{d}t \\ &= xS_0 + \int_0^T e^{a+b(x-X_t)} e^{-\rho t} \left(\dot{X}_t\right)^{\alpha+1} \mathrm{d}t \\ &= \int_0^T \sigma X_t \mathrm{d}W_t + \int_0^T e^{a+b(x-X_t)-\rho t} \dot{X}_t^{\,\alpha+1} \mathrm{d}t. \end{aligned} \tag{41}$$

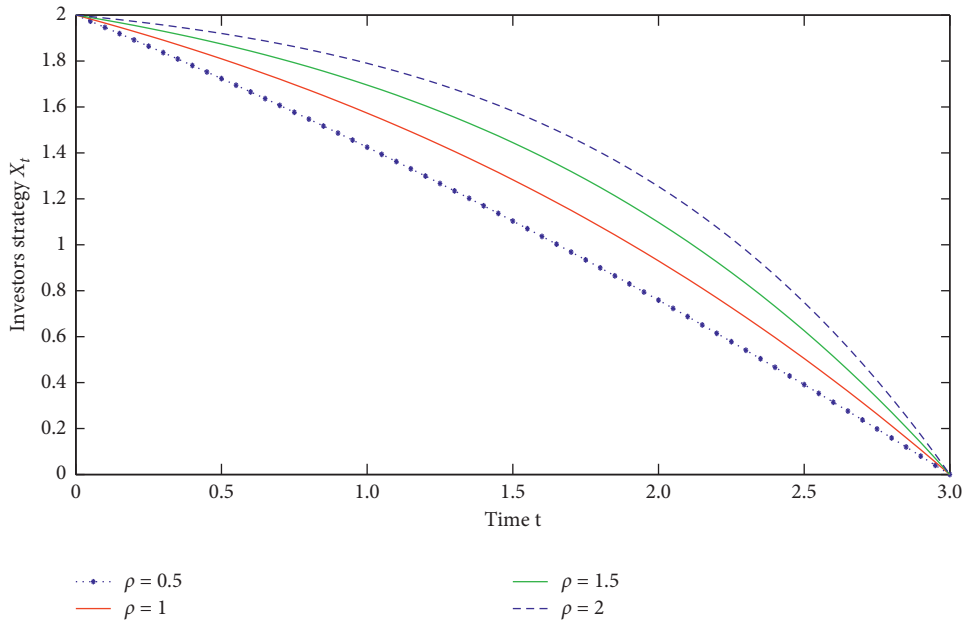


FIGURE 10: Influence of ρ change on equation (44) when $\alpha = 2$, $b = 0.5$, and $x = 2$.

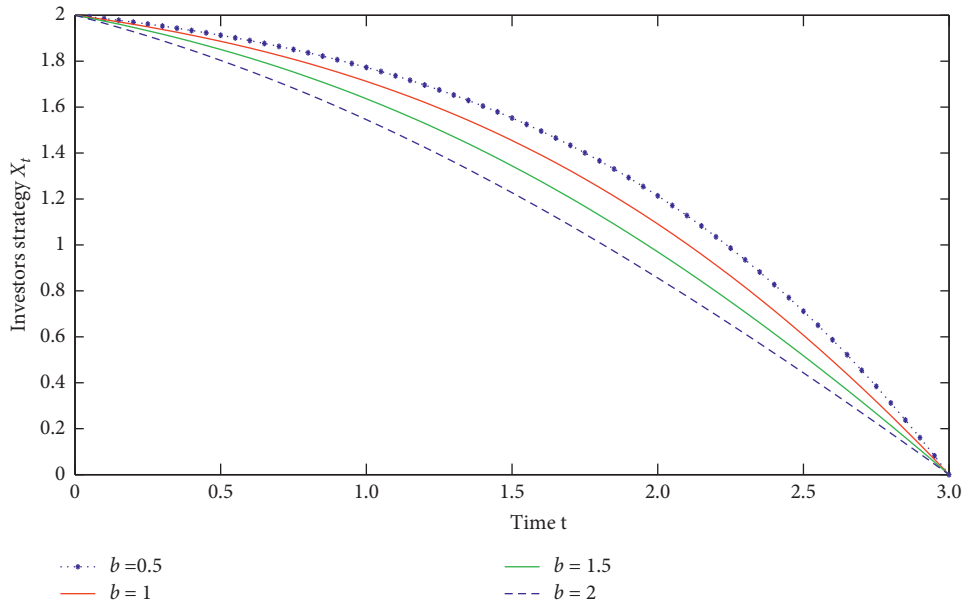


FIGURE 11: Influence of b change on equation (34) when $\alpha = 1$, $\rho = 1$, and $x = 2$.

Combined with the properties of Ito integral and equation (4), we obtain

$$\text{minimize}\,\mathbb{E}[\mathscr{C}(X)] = \int_0^T e^{a+b(x-X_t)-\rho t}\dot{X}_t^{\alpha+1}\mathrm{d}t. \tag{42}$$

Using the Euler–Lagrange equation, the following nonlinear ODE with variable coefficients is obtained:

$$2\ddot{X}_t - 2\rho\dot{X}_t - b\dot{X}_t^2 = 0. \tag{43}$$

From [16], the solution of equation (43) is

$$\int e^{-(b/\alpha+1)X_t}\mathrm{d}X_t = C_1 e^{(\rho/\alpha)t} + C_2. \tag{44}$$

□

Remark 6. In Theorem 6, we research the optimal liquidation behaviour when the market has exponential decay and temporary market impact and its coefficient are nonlinear functions. We can easily find that Theorem 4 is a special form when $\alpha = 1$.

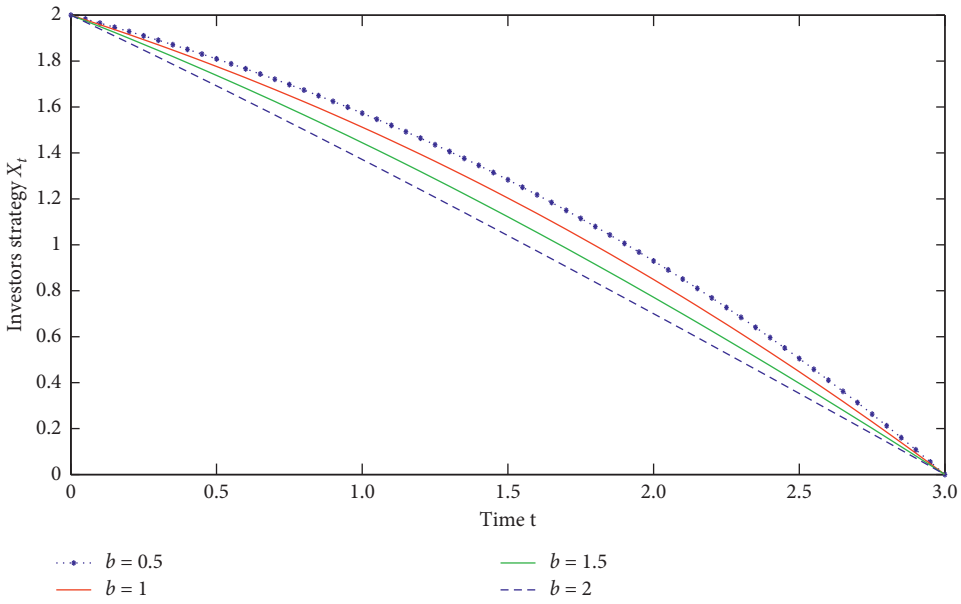


Figure 12: Influence of b change on equation (44) when $\alpha = 2$, $\rho = 1$, and $x = 2$.

4. Numerical Simulation

In Theorems 1, 3, 4, and 6, we give the specific form of the optimal investment behaviour for investors. Next, let $X_0 = 2$, $X_{t=3} = 0$, $t \in [0, 3]$, and we discuss the influence of parameters ρ, a, b, x, and α on the investor's optimal behaviour in four case, respectively.

From Figures 1–7, we find that investors realize that the risk of future trading is reduced when the market has decayed. Therefore, they reduce the amount X_t and speed of liquidation in different situations and speed up the liquidation before the liquidation deadline. It can be found from Figure 1 that when the exponent of temporary impact item increases, investors gradually increase the speed and amount of liquidation in the early stage, reduce the uncertainty risk in the future, and obtain the maximized expected return. Through equations (14) and (26), we find that the optimal behaviour of investors is an increasing function of a. This is verified by Figures 2 and 3. Figures 2 and 3 discuss the change of investors' optimal liquidation behaviour with respect to the linear and nonlinear temporary impact, respectively. With the increase in a, investors reduce the speed of liquidation in the early stage so that they get the maximum return. At the same time, we also find that when a is larger, the influence of a will gradually decrease. Combined with equations (14) and (26) and Figures 4 and 5, we find that no matter whether the temporary market impact is linear or nonlinear, the optimal liquidation behaviour of investors is a decreasing function of b. With the increase in b, investors realize to reduce the liquidation speed in the early period so as to increase the return of future liquidation. Figures 6 and 7 discuss the situation of investors' liquidation behaviour with the change of ρ under linear and nonlinear temporary impact, respectively. When investors find that ρ is increasing, they realize that the holding risk caused by reducing the amount of X_t in the early stage is reduced and increase the speed and amount of liquidation to maximize the return in the future. Therefore, when investors face these liquidation environments, they follow these liquidation strategies to minimize their execution cost in order to get the maximized returns.

Through equations (34) and (44), we see that the change of a has no effect on the liquidation behaviour. From Figures 8–12, we further discuss the properties of the optimal liquidation behaviour about b, α, and ρ. In Figure 8, we find that when a increases, investors realize that the risk of future trading will increase, and they will speed up the early liquidation to reduce the risk so as to obtain greater returns. From Figures 9 and 10, we find that when ρ is larger, investors will reduce the speed and amount of X_t in the early stage in order to reduce the risk of the exchange and get the maximized return. We study the influence of b on X_t under linear and nonlinear temporary impact in Figures 11 and 10. We find that when b becomes larger, investors will accelerate the speed and amount of liquidation in the early stage and reduce the holding risk of assets, so as to maximize the returns.

From Figures 1 and 8, we find that the change of α has the same influence on the liquidation behaviour in two different situations. With the increase in α, investors will speed up the early liquidation speed and amount so as to obtain the maximum returns. From Figures 4, 5, 11, and 12, the change of b has the same effect on liquidation behaviour X_t. When b gets larger, investors will liquidate faster so that they will reduce the risk of holding and get the maximum returns. Figures 6, 7, 4, and 10 investigate the relationship between decay factor ρ and liquidation behaviour X_t. We find that the change of ρ has the same effect on X_t in four

different situations. As ρ gets larger, the investor will maximize the returns by reducing the number of trades made earlier.

5. Conclusion

In this paper, we extend the model of Cay and Muhle-Karbe [15] when there is an exponential decay factor in the market. When temporary market impact and its coefficient are linear and nonlinear, we obtain the optimal liquidation behaviour X_t of investor which satisfies the second-order nonlinear ODE with variable coefficients. At the same time, we get specific forms of X_t in four situations. Through the numerical simulation, the influence of parameter α, a, b, and ρ and initial holdings x change on investors' liquidation strategy X_t are shown. When the market environment meets these conditions, it has certain guiding significance for investors to liquidate assets in the future.

In this paper, we assume that the random factor is the standard Brownian motion and the coefficients are time consistent. In the future, we learn from the relevant research of Zhu [18] and study the optimal liquidation behaviour when the stochastic factor of the price impact model is G-Brownian motion and coefficients of model are time delay. At the same time, when the coefficients are unknown, we refer to the research of Zhu and Wang [19] and use the control method to further discuss the optimal liquidation behaviour for investors.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by Key Research Base of Philosophy and Social Sciences for Colleges and Universities in Sichuan Province (KJJR2019-004), Talent Introduction Project of Xihua University (w202247), Fundamental Research Funds for the Central Universities (JBK2101036), and Scientific Research Project of Yibin University in 2019 (2019QD07).

References

[1] D. Bertsmas and A. Lo, "Optimal execution of execution costs," *Journal of Finance Markets*, vol. 1, pp. 1–50, 1998.

[2] R. Almgren and N. Chriss, "Optimal execution of portfolio transactions," *Journal of Risk*, vol. 3, no. 2, pp. 5–39, 2000.

[3] R. F. Almgren, "Optimal execution with nonlinear impact functions and trading-enhanced risk," *Applied Mathematical Finance*, vol. 10, no. 1, pp. 1–18, 2003.

[4] G. Huberman and W. Stanzl, "Price manipulation and quasi-arbitrage," *Econometrica*, vol. 72, no. 4, pp. 1247–1275, 2004.

[5] J. Gatheral and A. Schied, "Optimal trade execution under geometric brownian motion in the almgren and Chriss framework," *International Journal of Theoretical and Applied Finance*, vol. 14, no. 3, pp. 353–368, 2011.

[6] J. Gatheral, A. Schied, and A. Slynko, "Transient linear price impact and fredholm integral equations," *Mathematical Finance*, vol. 22, no. 3, pp. 445–3474, 2012.

[7] C.-A. Lehalle and E. Neuman, "Incorporating signals into optimal trading," *Finance and Stochastics*, vol. 23, no. 2, pp. 275–311, 2019.

[8] J. Gatheral, "No-dynamic-arbitrage and market impact," *Quantitative Finance*, vol. 10, no. 7, pp. 749–759, 2010.

[9] G. Curato, J. Gatheral, and F. Lillo, "Optimal execution with non-linear transient market impact," *Quantitative Finance*, vol. 17, no. 1, pp. 41–54, 2017.

[10] W. Barger and M. Lorig, "Optimal liquidation under stochastic price impact," *International Journal of Theoretical and Applied Finance*, vol. 21, no. 8, pp. 1850059–1850061, 2018.

[11] A. Bismuth, O. Guéant, and J. Pu, "Portfolio choice, portfolio liquidation, and portfolio transition under drift uncertainty," *Mathematics and Financial Economics*, vol. 13, no. 4, pp. 661–719, 2019.

[12] A. Schied and T. Zhang, "A state-constrained differential game arising in optimal portfolio liquidation," *Mathematical Finance*, vol. 27, no. 3, pp. 779–802, 2017.

[13] J. Vaes and R. Hauser, "Optimal execution strategy with an uncertain volume target," 2020, https://arxiv.org/pdf/1810.11454.pdf.

[14] I. Ekren and S. Nadtochiy, "Utility-based pricing and hedging of contingent claims in almgren-chriss model with temporary price impact," 2020, https://arxiv.org/pdf/1910.01778.pdf.

[15] T. Cayé and J. Muhle-Karbe, "Liquidation with self-exciting price impact," *Mathmatical Finance Economic*, vol. 10, no. 17, pp. 15–28, 2016.

[16] D. P. Andrei and F. Z. Valentin, *Handbook of Exact Solutions for Ordinary Differential Equations*, Chapman Hall/CRC, Boca Raton, FL, USA, 2003.

[17] U. Horst and F. Naujokat, "Illiquidity and derivative valuation," 2018, https://arxiv.org/pdf/0901.0091.pdf.

[18] Q. X. Zhu, "Stabilization of stochastic nonlinear delay systems with exogenous disturbances and the event-triggered feedback control," *IEEE Transactions on Automatic Control*, vol. 64, no. 9, pp. 3764–3771, 2019.

[19] Q. Zhu and H. Wang, "Output feedback stabilization of stochastic feedforward systems with unknown control coefficients and unknown output function," *Automatica*, vol. 87, pp. 166–175, 2018.

Hindawi
Complexity
Volume 2021, Article ID 5545551, 14 pages
https://doi.org/10.1155/2021/5545551

Research Article

Optimal Control through Leadership of the Cucker and Smale Flocking Model with Time Delays

Adsadang Himakalasa[1] and Suttida Wongkaew[2,3]

[1]*PhD Degree Program in Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand*
[2]*Research Center in Mathematics and Applied Mathematics, Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand*
[3]*Data Science Research Center, Department of Mathematics, Faculty of Sciences, Chiang Mai University, Chiang Mai 50200, Thailand*

Correspondence should be addressed to Suttida Wongkaew; suttida.wongkaew@cmu.ac.th

Received 1 March 2021; Revised 26 July 2021; Accepted 13 August 2021; Published 31 August 2021

Academic Editor: Miaomiao Wang

The Cucker and Smale model is a well-known flocking model that describes the emergence of flocks based on alignment. The first part focuses on investigating this model, including the effect of time delay and the presence of a leader. Furthermore, the control function is inserted into the dynamics of a leader to drive a group of agents to target. In the second part of this work, leadership-based optimal control is investigated. Moreover, the existence of the first-order optimality conditions for a delayed optimal control problem is discussed. Furthermore, the Runge–Kutta discretization method and the nonlinear conjugate gradient method are employed to solve the discrete optimality system. Finally, the capacity of the proposed control approach to drive a group of agents to reach the desired places or track the trajectory is demonstrated by numerical experiment results.

1. Introduction

Individual behaviors influence the dynamics of social systems. Studying the behavior of interacting individuals within a group of animals or a community of people is a new mathematical research field. The emergence of a collective level is one interesting issue raised by this type of study, with many natural systems showing collective phenomena, for instance, a flock of birds or a group of fish having coordinated movements, cells, chemical compounds, organisms or crystals, vehicular traffic, crowd dynamics, market economies, opinion formations, distributions of wealth, networks, and artificial intelligence (see, e.g., [1–6]). These collective behaviors are described as pattern configurations. In particular, collective action results from a superimposition of interactions between every possible pair of all agents. In general, the strength of such interaction forces depends on the amount of collaborative distance between agents. For example, birds orient with their nearest one in the group, and people more often agree with others who already conform to their beliefs. Nowadays, the application of these collective features is enormous and has a wide range. It appears in several engineering applications such as those using collective properties to perform complex tasks or in swarm robotics [7]. In addition, crowd models are applied in civil engineering to plan evacuation strategies in buildings (see, e.g., [8, 9]).

In recent years, researchers have been interested in studying collective phenomena, starting with using a few rules of interaction among agents to predict and explain unexpected phenomena or describe the emergence of pattern structure. Despite many works studying the emergence of consensus (see, e.g., [4, 10, 11]), investigating the enforcement and stabilization of pattern formation including the presence of noise or communication with time delays is especially interesting. Communication delays result from traffic congestion or finite speeds of transmission and spread. For example, the flocking model has a processing delay when analyzing information about the location and

velocity of neighboring agents. This fact has motivated an increased interest in studying the consensus problem of a flocking system, including the effect of time delays. Pioneering work on this issue was done in [12–16]. Additionally, when the system's behavior does not realize the desired result, control is essential to force the system to attain given objectives. This problem is related to the control of a self-organized system. Most results focus on the controllability of systems in which the topology of a group communication network is fully connected. In this case, all members of the group are regulated by the same distributed control law. These model settings, however, are limited when dealing with large networks. In particular, efficient control should apply only to a few group members, and this control strategy is known as "sparsity" in the mathematical literature. From the perspective of controllability, the hierarchical leadership concept provides the aggregation states of the system and some forms of group patterns for longtime behavior. The literature on control problems in these systems has recently been documented in [17–23].

The main contribution is divided into two parts. The first part discusses the Cucker and Smale (CS) flocking model, where the position and velocity of agents include the effect of time delay. Furthermore, primarily based on this model, a leader is added into the system to implement control functions. This control action can be applied to other agents through the mechanism of interaction force. The objective of the control strategy in this construction is to drive the evolution of flocking to track the desired trajectory. This control through leadership approach is constructed in the framework of optimal control problem mainly discussed in the second part of this work. In the second part, firstly, the formulation of optimal control problems with time delay is provided. The objective function includes three tracking terms and the cost of control function. The first and second terms account for the tracking error of the position of the leader and the desired target position at the final time and follow the desired path, respectively. The last tracking term is a measure of the distance between the leader and the other agents of the flock. Further, the theoretical and numerical results of optimization problems constrained with delayed CS systems are investigated. The resulting optimal control is obtained by solving an optimality system composed of the CS model's delayed dynamical system, associated adjoint equations, and an optimality condition. Additionally, the aspect of numerical solution and implementation is discussed. An efficient conjugate gradient optimization procedure evaluates optimal control [24, 25]. The gradient of the reduced cost functional is computed using the adjoint framework by solving the forward and backward delayed CS flocking equations that appear in the optimality system. These equations are discretized using an accurate high-order Runge–Kutta scheme (see more details in [26]).

The subsequent sections of our work are organized as follows. In Section 2, the description of the flocking model is provided. This model is based on the Cucker and Smale model, which takes time delays and the presence of a leader into account. Additionally, the existence of solutions and stability for these systems are discussed. Furthermore, numerical simulations are demonstrated to examine the characteristic properties of our constructed model. Section 3 presents the formulation of optimal control of the Cucker and Smale system with time delays. The theoretical issues regarding the existence of first-order necessary optimality systems of delayed optimal control problems are discussed. Section 4 is comprised of two parts. In the first part, the control problem with time delays is discretized by using the Runge–Kutta scheme. Then, using the first-discretize-then-optimize strategy, a discrete gradient formula for the optimization problem is obtained. The second part of Section 4 contains numerical experiments demonstrating the validity and effectiveness of the proposed control strategy. The conclusion is provided in the last section to complete this work.

2. Cucker and Smale Flocking Model with Time Delays and Leadership

In this section, we investigate the flocking model, including the effect of time delays and the presence of a leader. In this work, we focus on the Cucker and Smale (CS) model. This model is one of the well-known flocking models that involve the alignment of the agent; that is, each agent changes its speed in response to an averaging weight of its relative speed to the other agents, which can be seen in [5]. Moreover, we investigate the effect of time delays, which have influenced to converge to state transition and pattern structure. In our case, we consider constant delay function as proposed in [27]. In this context, delays are defined as a lag in the processing of information. To illustrate this phenomenon, consider a situation involving two agents in which the second agent attempts to follow the first but receives information from the first agent after some time. As a result, it requires time to elaborate on its response. We define τ as a positive constant time delay. Additionally, a leader is incorporated into the CS system to facilitate developing a control strategy for guiding the flock dynamical system toward a target (described in the following section). The CS flocking system with one leader and N followers including time delay is presented as follows:

$$\begin{aligned}
\dot{x}_0(t) &= v_0(t),\\
\dot{v}_0(t) &= \frac{1}{N+1}\sum_{j=1}^{N}\Big(a_{0j}(t-\tau)\Big(v_j(t-\tau)-v_0(t)\Big)\Big),\\
\dot{x}_b(t) &= v_b(t),\\
\dot{v}_b(t) &= \frac{1}{N+1}\sum_{j=1}^{N}\Big(a_{bj}(t-\tau)\Big(v_j(t-\tau)-v_b(t)\Big)\Big)+F_b^0,
\end{aligned} \tag{1}$$

for $b = 1, \ldots, N$, with given initial conditions

$$\begin{aligned}
x_i(t) &= x_i^0(t),\\
v_i(t) &= v_i^0(t), \quad \text{for } i = 0, \ldots, N \text{ and } t \in [-\tau, 0],
\end{aligned} \tag{2}$$

where $(x_i(t), v_i(t)) \in \mathbb{R}^d \times \mathbb{R}^d, d \geq 1$, are position and velocity of the i^{th} agent at time t. The initial conditions $x_i^0(t)$ and $v_i^0(t)$ are given continuous functions. The connectivity function a_{ij} measures the interaction strength between agents depending on the distance between i^{th} and j^{th} agents. The connectivity function is assumed to be positive, continuous, bounded, and nonincreasing function.

$$\begin{aligned} &0 < a_{ij}(t) \leq 1, \\ &\forall t \in [-\tau, +\infty), \quad \text{for } i = 0, \ldots, N \text{ and } j = 1, \ldots, N. \end{aligned} \tag{3}$$

In this work, we consider the connectivity function with the following form:

$$a_{bj}(t) := \varphi\left(\left\|x_i(t) - x_j(t)\right\|\right) = \frac{K}{\left(1 + \left\|x_b(t) - x_j(t)\right\|^2\right)^{\sigma}}, \tag{4}$$

where $K > 0$, $\sigma \geq 0$, and the notation $\|\cdot\|$ denotes Euclidean norm in $\mathbb{R}^n$. The force F_b^0 governs the action of the leader with time delay. The leader force can take many forms, such as using short-range repulsion and long-range attraction described as Morse potential function in [18] or considering hierarchical leadership structure in [16]. In this work, the leader's action takes the following form:

$$F_b^0 = \gamma\left(v_0(t - \tau) - v_b(t)\right), \tag{5}$$

where the parameter $\gamma > 0$ corresponds to the strength of the interaction force taken by the leader.

2.1. Stability of the Cucker and Smale Model with Time Delays. The solution of the CS flocking system is influenced by the parameters K and σ. In particular, when $\sigma < (1/2)$, the CS model without time delays exhibits unconditional flocking for all initial configurations, in which the velocities v_i approach asymptotically to the common limit velocity v^*. On the other hand, the flocking is conditional with $\sigma \geq (1/2)$, i.e., the asymptotic behavior of the system is dependent on the value of K and the initial configuration, as detailed in references [5, 14, 27]. To analyze the solution and stability of system (1), we consider the dispersion and disagreement as defined in [16, 19].

Definition 1 (dispersion and disagreement). Given a solution $(x_i(t), v_i(t)) \in \mathbb{R}^d \times \mathbb{R}^d$ of system (1), for $i = 0, \ldots, N$, we define dispersion and disagreement as

$$\begin{aligned} X(t) &= \frac{1}{2(N+1)^2}\left(\sum_{i,j=1}^{N}\left\|x_i(t) - x_j(t)\right\|\right), \\ V(t) &= \frac{1}{2(N+1)^2}\left(\sum_{i,j=1}^{N}\left\|v_i(t) - v_j(t)\right\|\right). \end{aligned} \tag{6}$$

Then, we say that the solution tends to **consensus** if

$$\begin{aligned} &\sup_{t>0} X(t) < +\infty, \\ &\lim_{t \longrightarrow +\infty} V(t) \end{aligned} \tag{7}$$

Next, the numerical simulation results of the CS system with time delays, including the presence of a leader, are demonstrated. Figure 1 depicts the numerical results of system (1) in one dimension, while Figures 2 and 3 illustrate results in two dimensions. For all experiments, we consider one leader and $N = 9$ followers. The corresponding parameters in connectivity function are chosen as $K = 2$ and $\sigma = (1/4)$. The leader interaction strength parameter is given as $\gamma = 10$. Consider 1D case, and Figure 1(a) shows the trajectories of agents with no time delay, whereas the motion of flocking, including the effect of time delay ($\tau = 5$), is presented in Figures 1(b)–1(d). It is observed that due to time delay, agents require more time to adjust their alignments and follow the group.

We also consider one leader and $N = 9$ followers in the experiments for 2D cases. The parameters for the 2D experiments are the same as those for the 1D studies. Agents begin with an initial configuration that places them all in random positions with random velocities, with a leader outside the group. After $t = 2$, each agent begins to adapt and organize its movement to follow the leader. Figure 3 shows that for the longtime behavior $t \longrightarrow \infty$, the value $V(t)$ tends to be constant at zero and the dispersion $X(t)$ is becoming constant. It explains that the consensus is reached after some time. The pattern configuration is formed, and each agent's velocity is synchronized and moves together at the same speed.

3. Optimal Control of Cucker and Smale Model with Time Delays

This section discusses the optimal control problem constrained with the delayed dynamical system of the CS model, including a leader. We note that this control problem is the leadership-based control strategy which is discussed in [18] for the refined flocking model. In addition, the theoretical results corresponding to optimal control problems with delays are investigated. In particular, the existence and characterization properties of solutions and optimality conditions are provided. In the following, we consider the optimal control problem governed by the delayed CS system and the presence of a leader presented as

$$\begin{aligned} \min_{\mathbf{x},u} J(\mathbf{x}, u) = &\frac{1}{2}\left\|x_0(T) - x_{\text{des}}(T)\right\|^2 + \frac{\eta}{2}\int_0^T \left\|x_0(t) - x_{\text{des}}(t)\right\|^2 \mathrm{d}t \\ &+ \frac{\mu}{2}\sum_{b=1}^{N}\int_0^T \left\|x_0(t) - x_b(t)\right\|^4 \mathrm{d}t + \frac{\nu}{2}\int_0^T \|u(t)\|^2 \mathrm{d}t, \end{aligned} \tag{8}$$

together with the delayed differential constraints given by

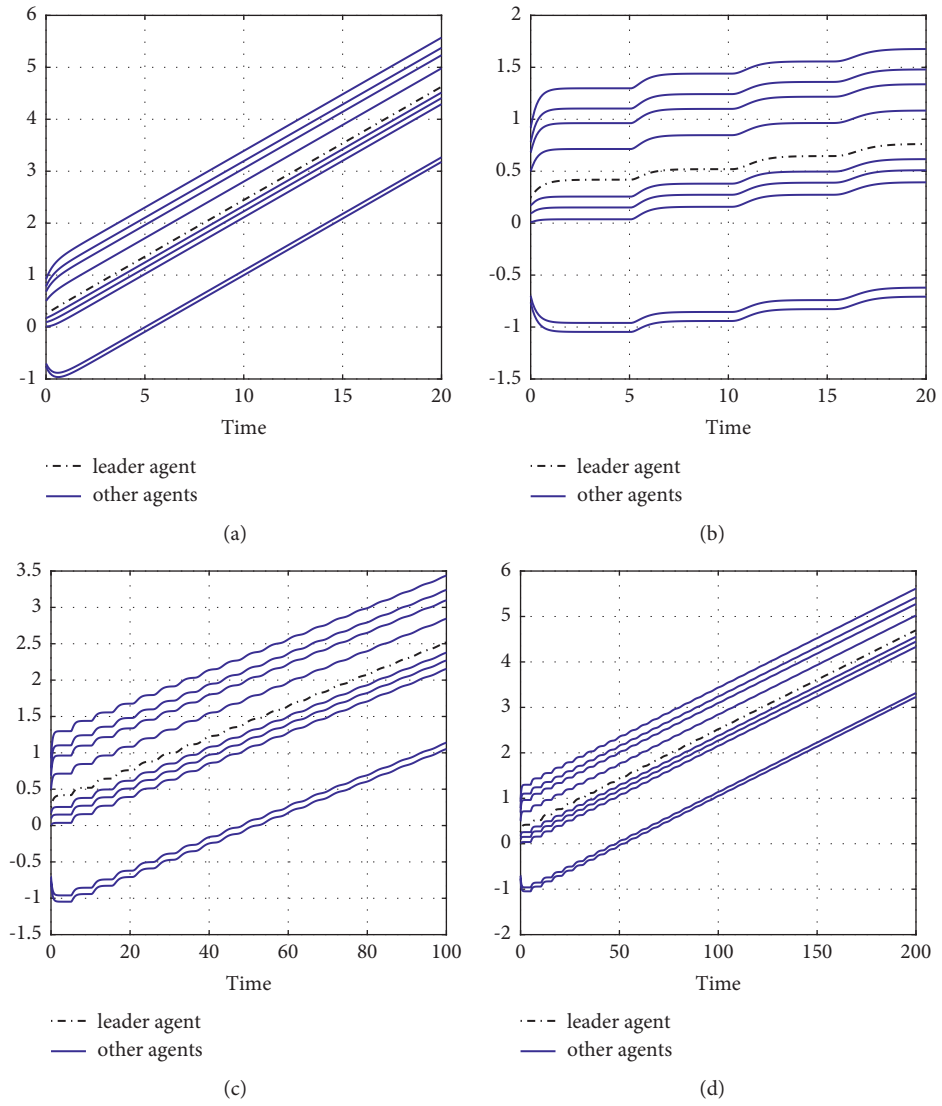


Figure 1: Simulation of the CS model with one leader and nine followers. (a) shows the trajectories of agents where the motion has no delay. (b)–(d) present the trajectories of agents where the movement of flocks have an effect of time delay ($\tau = 5$) in time $t = [0, 20]$, $t = [0, 100]$, and $t = [0, 200]$, respectively. Note that the dashed line represents the trajectory of the leader, while the blue line denotes the trajectory of the follower. (a) Without time delay for $t \in [0, 20]$. (b) With delay ($\tau = 5$) for $t \in [0, 20]$. (c) With delay ($\tau = 5$) for $t \in [0, 100]$. (d) With delay ($\tau = 5$) for $t \in [0, 200]$.

$$
\begin{aligned}
\dot{x}_0(t) &= v_0(t),\\
\dot{v}_0(t) &= \frac{1}{N+1}\sum_{j=1}^{N}\left(a_{0j}(t-\tau)\left(v_j(t-\tau)-v_0(t)\right)\right)+u(t),\\
\dot{x}_b(t) &= v_b(t),\\
\dot{v}_b(t) &= \frac{1}{N+1}\sum_{j=1}^{N}\left(a_{bj}(t-\tau)\left(v_j(t-\tau)-v_b(t)\right)\right)\\
&\quad+\gamma\left(v_0(t-\tau)-v_b(t)\right),
\end{aligned}
\tag{9}
$$

for $b = 1, \ldots, N$, along with their corresponding initial conditions. The state of the system is represented by the notation

$$
\mathbf{x}(t) = \left(x_0(t), \ldots, x_N(t), v_0(t), \ldots, v_N(t)\right) \in \mathbb{R}^{2(N+1)d}. \tag{10}
$$

The control function is denoted by $u \in L^2((0,T);\mathbb{R}^d)$. The parameters η, μ, and ν in the cost functional are positive constants. The first term in objective functional aims to minimize the terminal position of the leader and the given target. The second term is related to track the desired path by the leader agent. Also, the distance between the leader agent and others is measured in the third term. The cost of control is minimized in the last term.

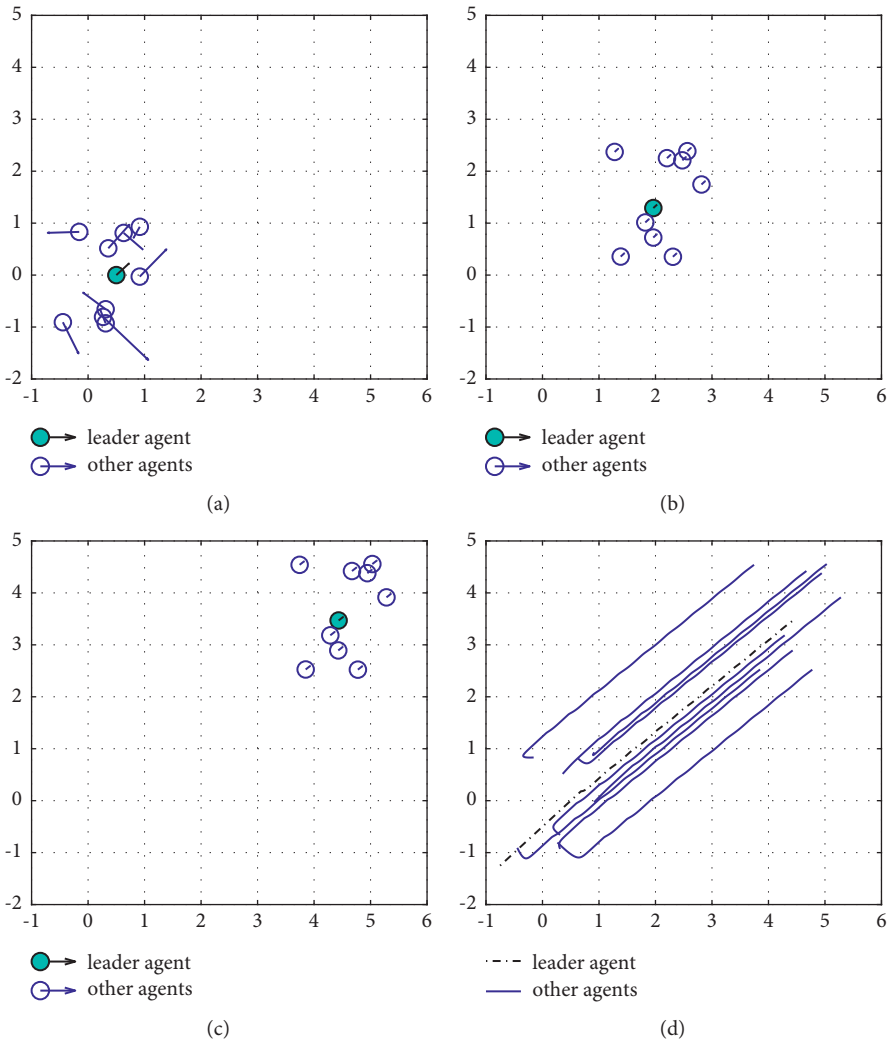


Figure 2: The simulation of the leadership of CS model with time delay $\tau = 2$ for one leader and nine followers in two dimensions. Parameters in the connectivity function are chosen as $K = 2, \sigma = (1/4)$ and leader interaction force is given as $\gamma = 10$. (a)–(c) present the positions of the agents and their velocities at time $t = 0$, $t = 20$, $t = 50$, respectively. The circles represent the positions of each agent and the shaded circle denotes the leader. (d) shows trajectories of agents in $t \in [0, 50]$. (a) Initial configuration. (b) $t = 20$. (c) $t = 50$. (d) Trajectories of agents.

To investigate the solution of our optimal control problems (8), we consider the optimal control problem that state and control variables have an effect of time delays, that is, state variable $\mathbf{x}(t)$ and control variable $\mathbf{u}(t)$ include time delays $\tau_x \geq 0$ and $\tau_u \geq 0$, respectively. The theoretical results of the delayed optimal control problems are investigated in [28, 29]. The general framework of an optimal control problem with delay in state and control is presented as follows:

$$\begin{aligned} &\min J(\mathbf{x}, \mathbf{u}) = \phi(\mathbf{x}(T)), \\ &\qquad \dot{x}(t) = f\left(t, \mathbf{x}(t), \mathbf{x}\left(t - \tau_x\right)\right)\mathbf{u}(t), \mathbf{u}\left(t - \tau_u\right)\Big), \\ &\text{subject to } \mathbf{x}(t) = \mathbf{x}_0(t), \quad t \in \left[-\tau_x, 0\right], \\ &\qquad \mathbf{u}(t) = \mathbf{u}_0(t), \quad t \in \left[-\tau_u, 0\right). \end{aligned} \tag{11}$$

We assume that $\mathbf{x}(t) \in W^{1,\infty}([0, T], \mathbb{R}^{nx})$ and $\mathbf{u}(t) \in L^{\infty}([0, T], \mathbb{R}^{nc})$. The function ϕ: $\mathbb{R}^{nx} \longrightarrow \mathbb{R}$ represents the objective function and is assumed to be continuously differentiable. f: $[0, T] \times \mathbb{R}^{nx} \times \mathbb{R}^{nx} \times \mathbb{R}^{nc} \times \mathbb{R}^{nc} \longrightarrow \mathbb{R}^{nx}$ represents the dynamics of the model and is assumed to be continuously differentiable. The initial conditions $\mathbf{x}_0$: $[-\tau_x, 0] \longrightarrow \mathbb{R}^{nx}$ and $\mathbf{u}_0$: $[-\tau_u, 0] \longrightarrow \mathbb{R}^{nc}$ are continuous. We note that a solution of the dynamical system in (11) is uniquely determined for a given control $\mathbf{u}$ and the initial conditions denoted by $\mathbf{x} = \mathbf{x}(\mathbf{u})$; moreover, $\mathbf{u} \mapsto \mathbf{x}(\mathbf{u})$ is a differentiable map. For all admissible $\mathbf{u}$, we aim to find a control $\mathbf{u}^*$ such that

$$J\left(\mathbf{x}\left(\mathbf{u}^*\right), \mathbf{u}^*\right) \leq J(\mathbf{x}(\mathbf{u}), \mathbf{u}). \tag{12}$$

Furthermore, to obtain the necessary optimality conditions for delayed optimal control (11), we introduce two additional variables $\mathbf{y} \in \mathbb{R}^{nx}$ and $\mathbf{w} \in \mathbb{R}^{nc}$ for delayed state

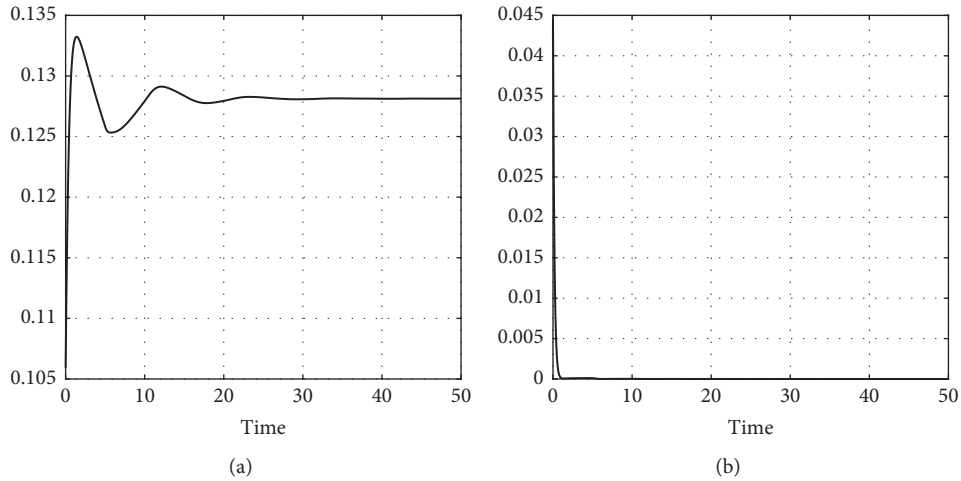


Figure 3: The dispersion and disagreement of the leadership of CS model with time delay $\tau = 2$ for $t \in [0, 50]$. (a) $X(t)$. (b) $V(t)$.

and control variables, respectively, where the functions $\mathbf{y}$ and $\mathbf{w}$ are defined as follows:

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{x}(t - \tau_x), \\ \mathbf{w}(t) &= \mathbf{u}(t - \tau_u). \end{aligned} \tag{13}$$

The first-order optimality system can be obtained by considering the Hamiltonian function that is given by

$$\mathscr{H}(t, \mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{w}, \mathbf{p}) \coloneqq \mathbf{p}^{\top} f(t, \mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{w}), \tag{14}$$

where $\mathbf{p} \in \mathbb{R}^{nx}$.

Theorem 1. *Let $(\mathbf{x}^*, \mathbf{u}^*)$ be locally optimal solution to optimal control problem (11). Assume that there exists an open set $\Omega \subset \mathbb{R}^{nx} \times \mathbb{R}^{nc}$ such that the neighborhood $B_\varepsilon(\mathbf{x}^*, \mathbf{u}^*)$ with a small radius $\epsilon > 0$ of $(\mathbf{x}^*, \mathbf{u}^*)$ is a subset of Ω for every $t \in [0, T]$. We assume that $\nabla_{\mathbf{x}} f$ and $\nabla_{\mathbf{u}} f$ are Lipschitz continuous in Ω. Moreover $\nabla_{\mathbf{x}} J$ is Lipschitz continuous in the neighborhood $B_\varepsilon(\mathbf{x}^*(T))$. Then, there exists an adjoint function $\mathbf{p}^* \in W^{1,\infty}([0, T], \mathbb{R}^{nx})$ that satisfies the first-order optimality conditions at $(\mathbf{x}^*, \mathbf{p}^*, \mathbf{u}^*)$, for a.e. $t \in [0, T]$ such that*

$$\begin{aligned} \dot{\mathbf{x}}(t) &= f(t, \mathbf{x}(t), \mathbf{y}, \mathbf{u}(t), \mathbf{w}(t)), \\ \mathbf{x}(t) &= \mathbf{x}_0(t), \quad \text{for } t \in [-\tau_x, 0], \\ \dot{\mathbf{p}}(t) &= -\nabla_{\mathbf{x}} \mathscr{H}(t) - \chi_{[0, T-\tau_x]}(t) \nabla_{\mathbf{y}} \mathscr{H}(t + \tau_x), \\ \mathbf{p}(T) &= \nabla_{\mathbf{x}} \phi(\mathbf{x}(T)), \\ &\nabla_u \mathscr{H}(t) + \chi_{[0, T-\tau_x]}(t) \nabla_{\mathbf{w}} \mathscr{H}(t + \tau_x) = 0, \end{aligned} \tag{15}$$

where $\nabla_{\mathbf{x}} \mathscr{H}(t)$, $\nabla_{\mathbf{y}} \mathscr{H}(t)$, $\nabla_{\mathbf{u}} \mathscr{H}(t)$, and $\nabla_{\mathbf{w}} \mathscr{H}(t)$ refer to the evaluation of the partial derivatives $\mathscr{H}$ with respect to $\mathbf{x}, \mathbf{y}, \mathbf{u}$, and $\mathbf{w}$, respectively. The function χ is defined as

$$\chi_{[a,b]}(t) = \begin{cases} 1, & \text{if } t \in [a, b], \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

To investigate the solution of our problem (10), we formulate our optimal control in form (11) by introducing the following equation:

$$\begin{aligned} \dot{\hat{x}}(t) &= \frac{\eta}{2} \left\| x_0(t) - x_{\text{des}}(t) \right\|^2 + \frac{\mu}{2} \sum_{b=1}^{N} \left\| x_0(t) - x_b(t) \right\|^4 + \frac{\gamma}{2} \|u(t)\|^2, \\ \hat{x}(0) &= 0. \end{aligned} \tag{17}$$

Then, the problem of optimal control (10) can be transformed into the following problem:

$$\begin{aligned} &\min J(\mathbf{x}, u) \\ &\quad \dot{x}_0(t) = v_0(t), \\ &\quad \dot{v}_0(t) = \frac{1}{N+1} \sum_{j=1}^{N} \left(a_{0j}(t - \tau) \left(v_j(t - \tau) - v_0(t) \right) \right) + u(t), \\ &\text{subject to} \quad \dot{x}_b(t) = v_b(t), \\ &\quad \dot{v}_b(t) = \frac{1}{N+1} \sum_{j=1}^{N} \left(a_{bj}(t - \tau) \left(v_j(t - \tau) - v_b(t) \right) \right) \\ &\qquad + \gamma \left(v_0(t - \tau) - v_b(t) \right), \end{aligned} \tag{18}$$

for $b = 1, \ldots, N$ and

$$\dot{\hat{x}}(t) = \frac{\eta}{2} \left\| x_0(t) - x_{\text{des}}(t) \right\|^2 + \frac{\mu}{2} \sum_{b=1}^{N} \left\| x_0(t) - x_b(t) \right\|^4 + \frac{\gamma}{2} \|u(t)\|^2, \tag{19}$$

together with corresponding initial conditions, and the compact form can be rewritten as follows:

$$\min J(\tilde{\mathbf{x}}, u) = \frac{1}{2}\left\|x_0(T) - x_{\text{des}}(T)\right\|_2^2 + \hat{x}(T)$$
$$\text{subject to } \dot{\tilde{\mathbf{x}}}(t) = F(t, \tilde{\mathbf{x}}, u). \tag{20}$$

Now, the state variable is

$$\tilde{\mathbf{x}}(t) = \left(x_0(t), x_1(t), \ldots, x_N(t), v_0(t), v_1(t), \ldots, v_N(t), \hat{x}(t)\right)^T, \tag{21}$$

and $F(\tilde{\mathbf{x}}, u)$ represents the dynamics of the transformed delayed CS systems.

Further, we discuss the first-order optimality conditions for optimal control problem (20). For this purpose, the delayed system is transformed to nondelayed system by using variables

$$\begin{aligned} s_i(t) &= x_i(t-\tau), \\ z_i(t) &= v_i(t-\tau), \quad \text{for } i = 0, 1, \ldots, N, \end{aligned} \tag{22}$$

and we denote $\mathbf{y}(t) = (s_0(t), s_1(t), \ldots, s_N(t), z_0(t), z_1(t), \ldots, z_N(t))^\top$.

The Hamiltonian function for system (20) is presented as

$$\mathscr{H}(t, \tilde{\mathbf{x}}, \mathbf{y}, u, \mathbf{p}) = \mathbf{p}^\top F(t, \tilde{\mathbf{x}}, \mathbf{y}, u), \tag{23}$$

where $\mathbf{p}(t) = (p_{x_0}(t), p_{x_b}(t), p_{v_0}(t), p_{v_b}(t), p_{\hat{x}}(t))^\top \in \mathbb{R}^{nx}$, $nx = 2(N+1)d + 1$.

Assumption 1. There exist Lagrange multipliers $\mathbf{p} \in W^{1,\infty}([0,T], \mathbb{R}^{nx})$ corresponding to the optimization constraint equation of problem (20). Moreover, the adjoint $\mathbf{p}$ satisfies the following equation:

$$\begin{aligned} \dot{\mathbf{p}}(t) &= -\left(\nabla_{\mathbf{x}} F(t)\right)^\top \mathbf{p}(t) - \chi_{[0,T-\tau]}(t)\left(\nabla_{\mathbf{y}} F(t+\tau)\right)^\top \mathbf{p}(t+\tau), \\ \mathbf{p}(T) &= \nabla_{\mathbf{x}} \phi(\mathbf{x}(T)). \end{aligned} \tag{24}$$

Therefore, a solution of (20) can be characterized by the first-order optimality system as presented in the following:

$$\begin{aligned} \dot{\tilde{\mathbf{x}}}(t) &= F(t, \tilde{x}, u), \\ \tilde{\mathbf{x}}(t) &= \tilde{x}_0(t) \quad \text{for } t \in [-\tau, 0], \\ \dot{\mathbf{p}}(t) &= -\left(\nabla_{\mathbf{x}} F(t)\right)^\top \mathbf{p}(t) - \chi_{[0,T-\tau]}(t)\left(\nabla_{\mathbf{y}} F(t+\tau)\right)^\top \mathbf{p}(t+\tau), \\ \mathbf{p}(T) &= \nabla_{\mathbf{x}} \phi(\mathbf{x}(T)). \end{aligned} \tag{25}$$

In addition, the corresponding gradient can be expressed as

$$\nabla_u J(u) = -\left(\nabla_u F(t)\right)^\top \mathbf{p}(t). \tag{26}$$

In particular, the explicit formulations of adjoint equations (20) can be expressed as follows:

$$\begin{aligned} \dot{p}_{x_0}(t) &= -\eta \langle p_{\hat{x}}(t), (x_0(t) - x_{\text{des}}(t)) \rangle - 2\mu \langle p_{\hat{x}}(t), \sum_{b=1}^{N} (x_0(t) - x_b(t))^3 \rangle \\ &\quad + \chi_{[t_0,T-\tau]}(t)\left(\frac{1}{N+1}\right) \langle p_{v_0}(t+\tau), \sum_{j=1}^{N} a'_{0j}(t)\left(v_j(t) - v_0(t+\tau)\right) \rangle, \\ \dot{p}_{v_0}(t) &= -p_{x_0}(t) + \left[\frac{1}{N+1}\sum_{j=1}^{N} a_{0j}(t-\tau)\right] p_{v_0}(t) - \gamma \chi_{[t_0,T-\tau]}(t) \sum_{b=1}^{N} p_{v_b}(t+\tau), \\ \dot{p}_{x_b}(t) &= 2\mu \langle p_{\hat{x}}(t), (x_0(t) - x_b(t))^3 \rangle \\ &\quad - \chi_{[t_0,T-\tau]}(t) \frac{1}{N+1} \langle p_{v_0}(t+\tau), a'_{0b}(t)\left(v_b(t) - v_0(t+\tau)\right) \rangle \\ &\quad + \chi_{[t_0,T-\tau]}(t) \frac{1}{N+1} \langle p_{v_0}(t+\tau), \sum_{j=1}^{N} a'_{bj}(t)\left(v_j(t) - v_b(t+\tau)\right) \rangle, \\ \dot{p}_{v_b}(t) &= -p_{x_b}(t) + \left[\frac{1}{N+1}\sum_{j\neq b}^{N} a_{bj}(t-\tau)\right] p_{v_b}(t) - \gamma p_{v_b}(t) - \chi_{[t_0,T-\tau]}(t)\left(\frac{1}{N+1}\right) a_{0b}(t) p_{v_b}(t+\tau), \\ \dot{p}_{\hat{x}}(t) &= 0, \end{aligned} \tag{27}$$

with transversality condition

$$\begin{aligned} p_{x_b}(T) &= x_0(T) - x_{\text{des}}(T), \\ p_{x_b}(T) &= p_{v_b}(T) = p_{v_0}(T) = 0, \\ p_{\hat{x}}(T) &= 1. \end{aligned} \tag{28}$$

The numerical solution of problems (20) representing as the considered optimality systems (25) will be shown and discussed in the next section.

4. Numerical and Implementation Aspects of Delayed Control Problems

4.1. First-Discretize-Then-Optimize Strategy and Discrete Optimality Conditions. The discretization of the reduced gradient is a critical step in the numerical method for solving optimal control problems. To obtain an accurate discretization for the reduced gradient, we consider the first-discretize-then-optimize strategy. To carry out this strategy, one can follow the procedures outlined below. The cost functional and corresponding differential constraints representing the optimal control problem are discretized in the first step using the Runge–Kutta method, discussed in detail in [26]. Second, the discrete Lagrangian function corresponding to the discrete optimal control problem must be constructed. The final step is to obtain the first-order optimality system for the discretized problem.

For the discretization of optimal control problem (20), the Runge–Kutta scheme is employed on a uniform mesh in the time intervals $(0, T)$ such that the step-sized h is defined as

$$h = \frac{T}{n}, \tag{29}$$

with the total number of discrete time intervals, n. It is important to match the uniform step size $h > 0$ with positive delay τ; that is, we choose any integer fraction of h to refine the discretization grid. For this task, for any step size $h > 0$, there exist positive integers $n, m_\tau \in \mathbb{N}$ satisfying

$$\begin{aligned} \tau &= m_\tau h, \\ T &= nh. \end{aligned} \tag{30}$$

By this setting, we denote the value of $\widetilde{x}(t)$ at the discrete time t^k by

$$\begin{aligned} \widetilde{x}_k &= \widetilde{x}(t_k), \\ t_k &= kh, \quad \text{for } k = 0, \ldots, n-1. \end{aligned} \tag{31}$$

Therefore, the discretization of the s-stage Runge–Kutta scheme setting for optimal control problem (20) becomes

$$\begin{aligned} \min_{\widetilde{x},u} J(\widetilde{x}_k, u_k) &= \frac{1}{2}\left\| x_0(T) - x_{\text{des}}(T) \right\|_2^2 + \hat{x}(T), \\ \widetilde{x}_{k+1} &= \widetilde{x}_k + h\sum_{i=1}^{s} b_i F\left(t_k, \widetilde{x}_k, \widetilde{x}_{k-m_\tau}, \psi_{ki}, u_{ki}\right), \\ \text{subject to } \widetilde{x}_{-k} &= \widetilde{x}_0(-kh), \quad k = 0, \ldots, m_\tau, \\ \psi_{ki} &= \widetilde{x}_k + h\sum_{j=1}^{s} a_{ij} F\left(t_k, \widetilde{x}_k, \widetilde{x}_{k-m_\tau}, \psi_{kj}, u_{kj}\right), \end{aligned} \tag{32}$$

for $k = 0, \ldots, n-1$ and $1 \le i, j \le s$. Note that the control $u^k \in R^{nc \times s}$ denotes the s stages of the RK discrete control function at the stage k, which can be written in the following form:

$$u_k = (u_{k1}, u_{k2}, \ldots, u_{ks}). \tag{33}$$

We remark that the order of a Runge–Kutta discretization for optimal control problem depends on coefficients a_{ij} and b_i (see more details in [26]).

The discrete optimality system that corresponds to (32) is expressed as the following equations:

$$\begin{aligned} \widetilde{x}_{k+1} &= \widetilde{x}_k + h\sum_{i=1}^{s} b_i F\left(t_k, \widetilde{x}_k, \widetilde{x}_{k-m_\tau}, \psi_{ki}, u_{ki}\right), \\ \widetilde{x}_{-k} &= \widetilde{x}_0(-kh), \\ \psi_{ki} &= \widetilde{x}_k + h\sum_{j=1}^{s} a_{ij} F\left(t_k, \widetilde{x}_k, \widetilde{x}_{k-m_\tau}, \psi_{kj}, u_{kj}\right), \\ P_k &= P_{k+1} + \sum_{i=1}^{s} b_i \zeta_{ki}, \\ P_n &= -\nabla_x \phi(\widetilde{x}_n), \\ \zeta_{ki} &= \left(\nabla_x F\left(t_k, \widetilde{x}_k, \widetilde{x}_{k-m_\tau}, \psi_{kj}, u_{kj}\right)\right)^\top \left(\psi_{k+1} + \sum_{j=1}^{s} \frac{b_j a_{ij}}{b_i} \zeta_{kj}\right) \\ &\quad + \chi_{[0,T-m_\tau]}(k)\left(\nabla_y F\left(t_{k+m_\tau}, \widetilde{x}_{k+m_\tau}, \widetilde{x}_k, \psi_{kj+m_\tau}, u_{kj+m_\tau}\right)\right)^\top \left(\psi_{k+1+m_\tau}\right) \\ &\quad + \chi_{[0,T-m_\tau]}(k)\left(\nabla_y F\left(t_{k+m_\tau}, \widetilde{x}_{k+m_\tau}, \widetilde{x}_k, \psi_{kj+m_\tau}, u_{kj+m_\tau}\right)\right)^\top \left(\sum_{j=1}^{s} \frac{b_j a_{ij}}{b_i} \zeta_{kj+m_\tau}\right). \end{aligned} \tag{34}$$

Input u;
(1) The initial conditions are provided and then solve the forward the discrete CS model;
(2) Compute the terminal condition and solve the discrete adjoint equation in (34);
(3) Evaluate the gradient $\nabla_{\mathbf{u}} J(\mathbf{u})$ using (35);

Algorithm 1: Evaluation of the gradient at u.

Input: $\mathbf{u}_0$ index $k = 0$, maximum $k_{\max}$, tolerance $\text{tol} > 0$.
While ($k < k_{\max}$ and $\|\nabla J(u_k)\| > \text{tol}$) do
(1) Obtain $\widetilde{x}_k$ from step 1 in Algorithm 1 with u_k;
(2) Get P_k from step 1 in Algorithm 1 with $(\widetilde{x}_k, u_k)$;
(3) Evaluate the gradient $\nabla_u J(u_k)$ from step 1 in Algorithm 1;
(4) Compute α_k by using backtracking line search scheme with Armijo's condition [30];
(5) Compute β_k by using Hager and Zhang search direction formula [24];
(6) Update $u_{k+1} = u_k + \alpha_k \beta_k$;
(7) $k = k + 1$
End while

Algorithm 2: NCG with Hager and Zhang scheme.

Further, the gradient is in the following form:

$$\nabla_{\mathbf{u}_{ki}} J(u) = -\left(\nabla_u F(\psi_{ki}, \mathbf{u}_{ki})\right)^{\top} \left(P_{k+1} + \sum_{j=1}^{s} \frac{b_j a_{ij}}{b_i} \zeta_{kj} \right), \quad (35)$$

for $1 \le i, j \le s$, and $0 \le k \le n-1$.

To solve optimization problem (32), the nonlinear conjugate gradient (NCG) strategy with the Hager and Zhang scheme [24, 25] is implemented. First, we use the following algorithm to compute the gradient specified in First, we use the following algorithm to compute the gradient specified in (35).

Following that, we apply the gradient derived from Algorithm 1 to the NCG scheme defined in Algorithm 2.

4.2. Numerical Results. This section presents the numerical results of our flocking models. These results demonstrate the control performance of our leader-based control strategy. The test is divided into two parts as follows:

(i) Test I: reach the desired target point.

(ii) Test II: follow the desired path.

In our experiments, we investigate problem (20) in two dimensions by considering a system composed of nine agents and one leader. As a result, the state variables' total $nx = 20$, while the control variables' total $nc = 1$. In numerical test I, we focus on controlling the leader agent to reach the goal position at the final time. The purpose of numerical test II is to force the flocks to follow the desired trajectories. We provide two examples of the desired trajectory in these tests: the linear and circular paths. We use the same initial conditions in all three examples, where agents' positions are distributed, and their initial velocities are random. Additionally, the initial position of the leader is set to $x0 = (0, 0)$.

In numerical test I, the goal position is given as $x_{\text{des}} = (5, 5)$ and the end time is set to $T = 10$. The interaction force parameters are $K = 2, \sigma = 0.25$ and the strength of leader interaction force is $\gamma = 10$. The objective of this test is to force the system to reach the desired target at the final time; as a consequence, the corresponding parameters in objective functional are $\mu = 0.1, \nu = 1$, and $\eta = 0$. As illustrated in Figure 4, the leader is capable of reaching the desired location and leading the flock there.

In numerical test II, the corresponding parameters and initial data are chosen similarly to those in the numerical test I, except that $\eta = 10$ is used to keep the leader agent tracking the desired trajectories. We illustrate two scenarios using two distinct tracking trajectories. The results of the first example are depicted in Figure 5. The plot of snapshots for ten agents traveling along the specified linear path is shown in Figure 5. Figures 5(a)–5(e) illustrate the solutions to optimal control problems at various time points. As illustrated in Figure 5(f), the leader attempts to move in the direction of a given trajectory while the group of agents follows the leader. In the last example, the desired path is given as a circle. As shown in Figure 6, the leader agent tracks the path, and the other agents follow the leader. It can be observed from three examples that agents keep some distance from the leader because time delay creates some

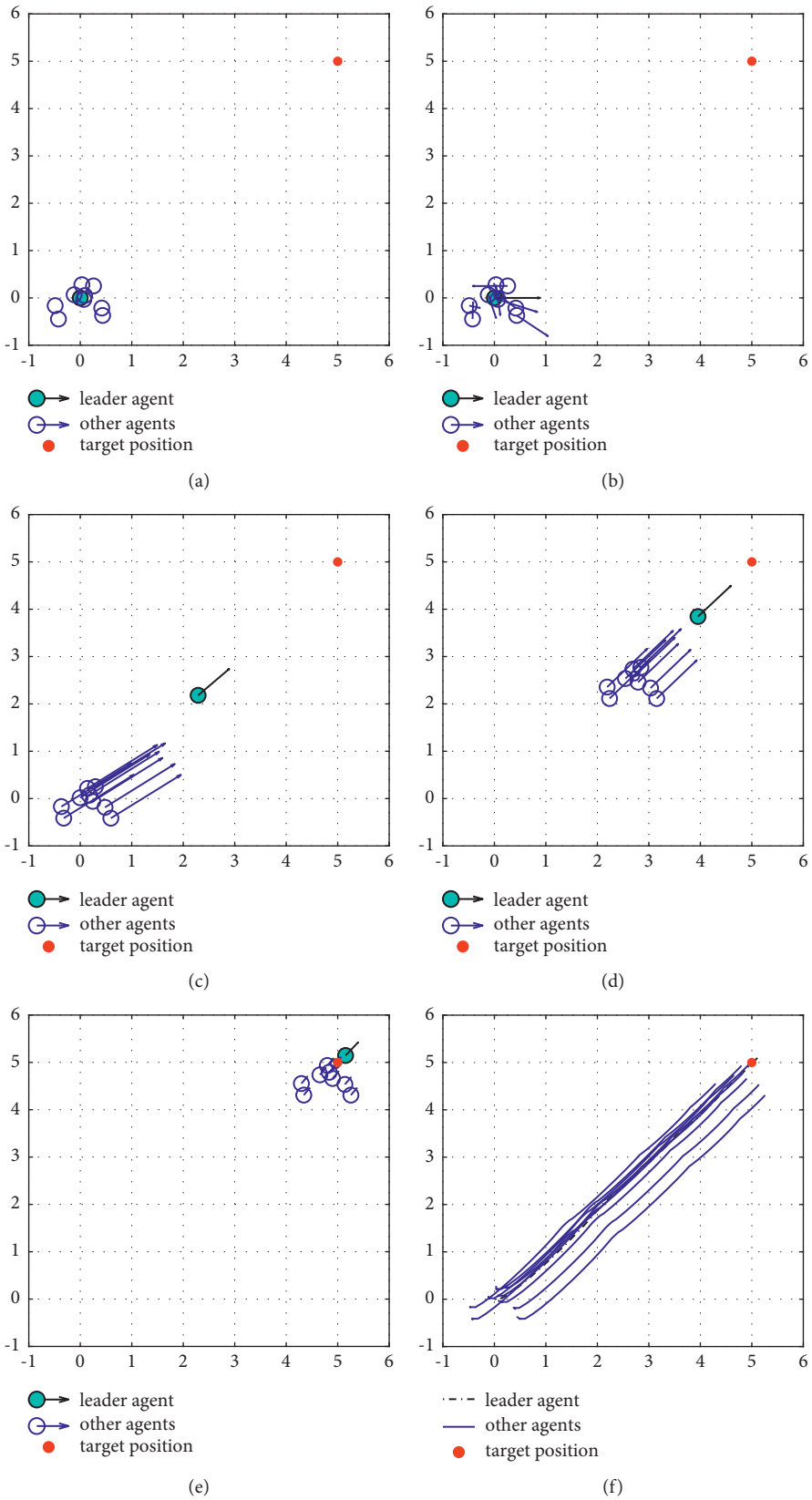


FIGURE 4: The movement of agents reaching final target position governed by leadership of CS model with time delay $\tau = 2$ for $t \in [0, 10]$. (a) Initial configuration. (b) $t = 1$. (c) $t = 2$. (d) $t = 5$. (e) $t = 10$. (f) $t \in [0, 10]$.

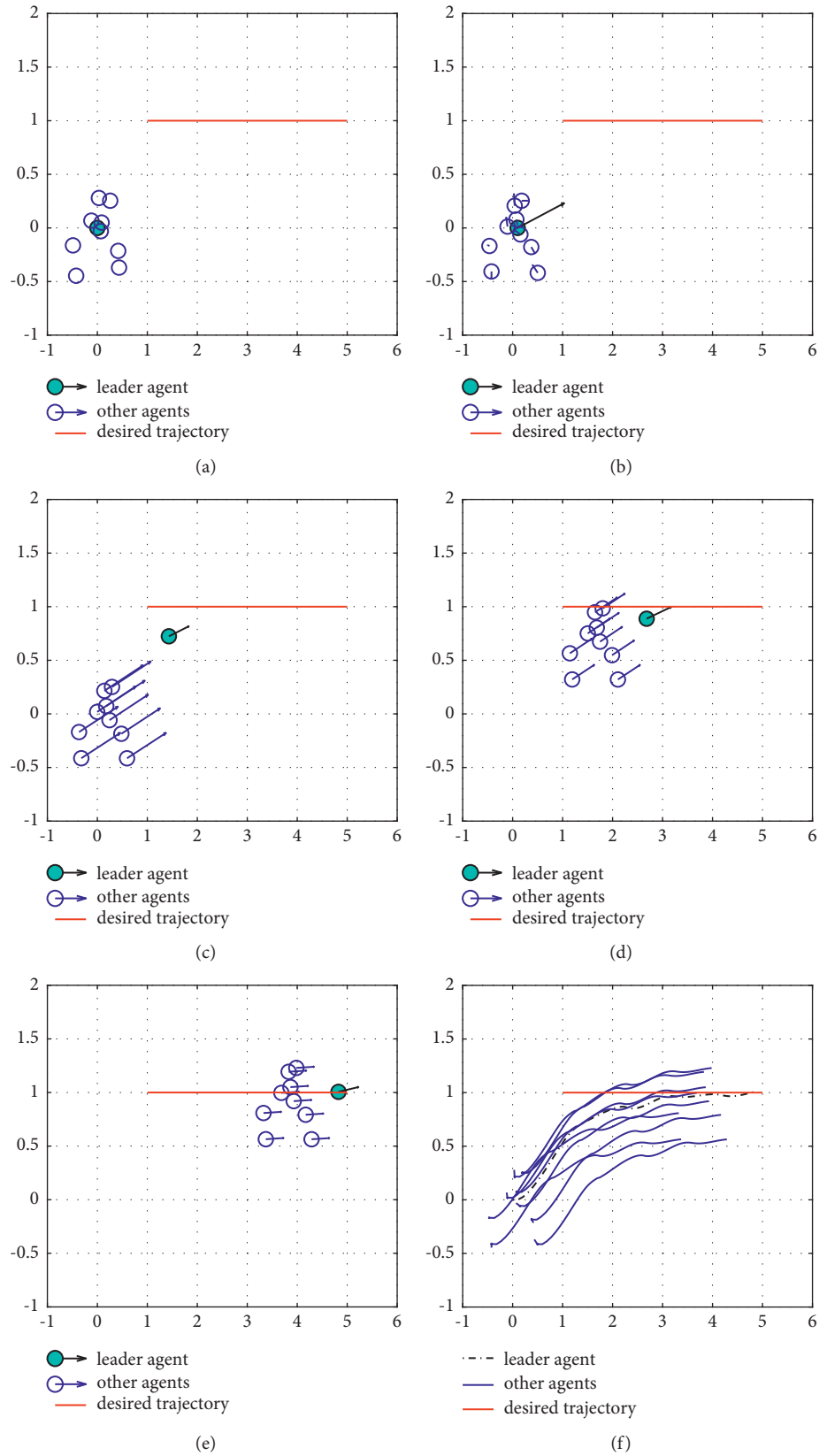


Figure 5: The movement of agents following desired linear path. Time delay is given by $\tau = 2$ for $t \in [0, 10]$. (a) Initial configuration. (b) $t = 1$. (c) $t = 2$. (d) $t = 5$. (e) $t = 10$. (f) $t \in [0, 10]$.

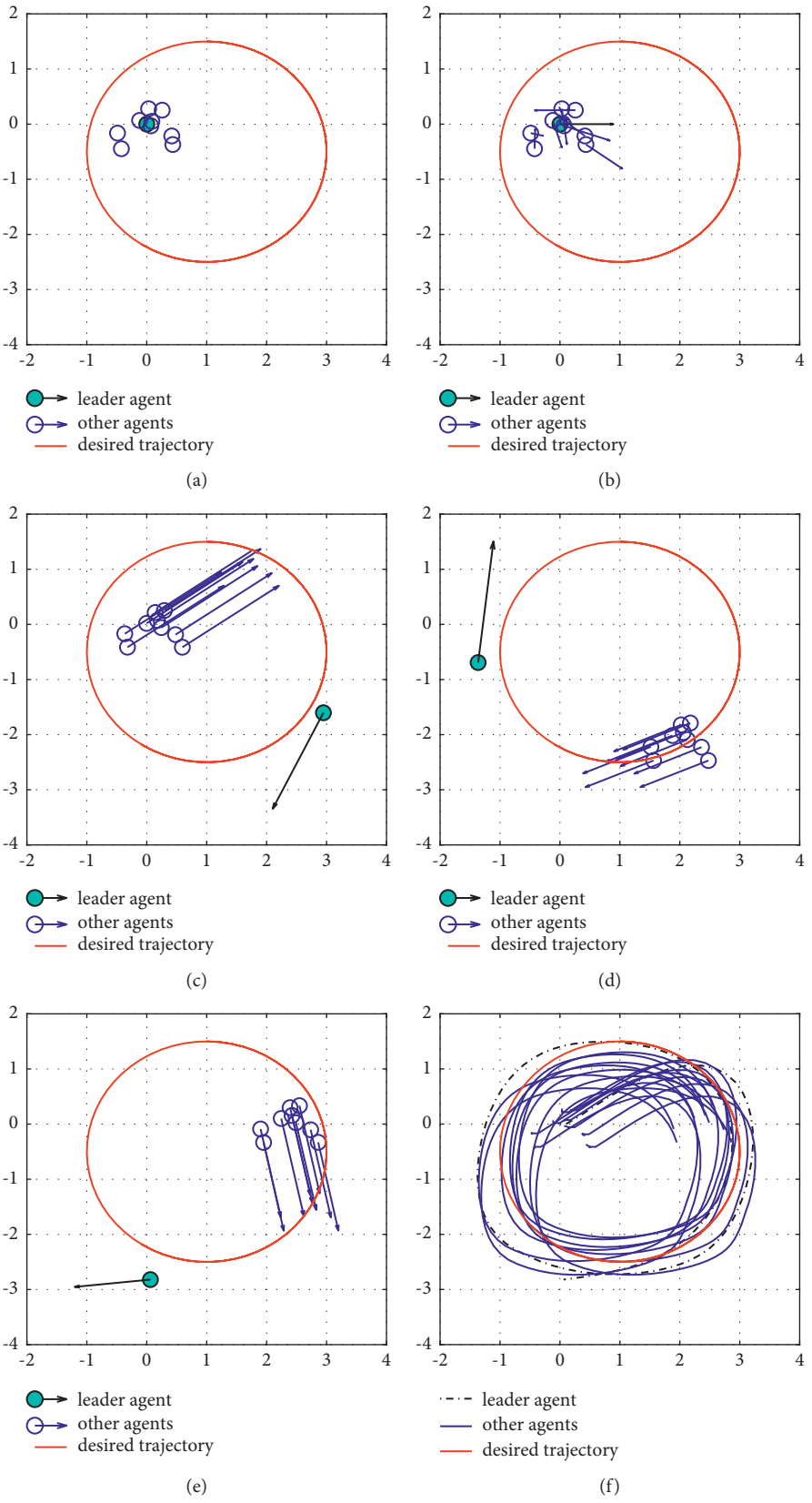


Figure 6: The movement of agents tracking the desired circular trajectory governed by leadership of delayed CS mode. Time delay is given by $\tau = 2$ for $t \in [0, 10]$. (a) Initial configuration. (b) $t = 1$. (c) $t = 2$. (d) $t = 5$. (e) $t = 10$. (f) $t \in [0, 10]$.

absence of leader position. It would take some time for the group to respond to this emotion and attempt to follow the leader; however, the group eventually achieves it.

5. Conclusion

The Cucker and Smale flocking model, including the effect of time delays and the presence of a leader, was studied. A control-based leadership technique was investigated for this model. An optimal control problem was formulated, and the corresponding discretization of an optimal problem was solved using an accurate RK method such that accurate gradients of the reduced objectives can be guaranteed. The nonlinear conjugate gradient scheme was implemented to compute the gradients. The results of the numerical experiment show that the proposed control approach is practical. From a modeling point of view, this flocking system could be enhanced by including the effects of attraction and repulsion and alignment with the vision cone and delay. In addition, the delay in control function would be considered a part of the optimal control problem.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by Chiang Mai University.

References

[1] M. Ballerini, N. Cabibbo, R. Candelier et al., "Interaction ruling animal collective behavior depends on topological rather than metric distance: evidence from a field study," *Proceedings of the National Academy of Sciences*, vol. 105, no. 4, pp. 1232–1237, 2008.

[2] A. B. T. Barbaro, K. Taylor, P. F. Trethewey, L. Youseff, and B. Birnir, "Discrete and continuous models of the dynamics of pelagic fish: application to the capelin," *Mathematics and Computers in Simulation*, vol. 79, no. 12, pp. 3397–3414, 2009.

[3] N. Bellomo and C. Dogbe, "On the modeling of traffic and crowds: a survey of models, speculations, and perspectives," *SIAM Review*, vol. 53, no. 3, pp. 409–463, 2011.

[4] E. Cristiani, B. Piccoli, and A. Tosin, "Modeling self-organization in pedestrians and animal groups from macroscopic and microscopic viewpoints," in *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*, pp. 337–364, Birkhäuser, Basel, Switzerland, 2010.

[5] F. Cucker and S. Smale, "Emergent behavior in flocks," *IEEE Transactions on Automatic Control*, vol. 52, no. 5, pp. 852–862, 2007.

[6] B. Düring, P. Markowich, J.-F. Pietschmann, and M.-T. Wolfram, "Boltzmann and Fokker-Planck equations modelling opinion formation in the presence of strong leaders," *Proceedings of the Royal Society A: Mathematical, Physical & Engineering Sciences*, vol. 465, no. 2112, pp. 3687–3708, 2009.

[7] Y.-L. Chuang, R. H. Yuan, M. R D'Orsogna, and A. L. Bertozzi, "Multi-vehicle flocking: scalability of cooperative control algorithms using pairwise potentials," in *Proceedings of the 2007 IEEE International Conference on Robotics and Automation*, pp. 2292–2299, IEEE, Roma, Italy, April 2007.

[8] E. Cristiani, F. S. Priuli, and A. Tosin, "Modeling rationality to control self-organization of crowds: an environmental approach," *SIAM Journal on Applied Mathematics*, vol. 75, no. 2, pp. 605–629, 2015.

[9] M. Moussaïd, E. G. Guillot, M. Moreau et al., "Traffic instabilities in self-organized pedestrian crowds," *PLoS Computational Biology*, vol. 8, no. 3, p. e1002442, 2012.

[10] J. A. Carrillo, M. Fornasier, J. Rosado, and G. Toscani, "Asymptotic flocking dynamics for the kinetic Cucker-Smale model," *SIAM Journal on Mathematical Analysis*, vol. 42, no. 1, pp. 218–236, 2010.

[11] J. A. Carrillo, M. Fornasier, G. Toscani, and F. Vecil, "Particle, kinetic, and hydrodynamic models of swarming," in *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*, pp. 297–336, Birkhäuser, Basel, Switzerland, 2010.

[12] C.-L. Chen, D.-Y. Sun, and C.-Y. Chang, "Numerical solution of time-delayed optimal control problems by iterative dynamic programming," *Optimal Control Applications and Methods*, vol. 21, no. 3, pp. 91–105, 2000.

[13] P. C. Young and H. Jan, "Cucker-smale model with normalized communication weights and time delay," 2016, https://arxiv.org/pdf/1608.06747.

[14] R. Erban, J. Haškovec, and Y. Sun, "A Cucker--Smale model with noise and delay," *SIAM Journal on Applied Mathematics*, vol. 76, no. 4, pp. 1535–1557, 2016.

[15] Y. Liu and J. Wu, "Flocking and asymptotic velocity of the Cucker-Smale model with processing delay," *Journal of Mathematical Analysis and Applications*, vol. 415, no. 1, pp. 53–61, 2014.

[16] C. Pignotti and I. Reche Vallejo, "Flocking estimates for the Cucker-Smale model with time lag and hierarchical leadership," *Journal of Mathematical Analysis and Applications*, vol. 464, no. 2, pp. 1313–1332, 2018.

[17] M. Aureli and M. Porfiri, "Coordination of self-propelled particles through external leadership," *EPL (Europhysics Letters)*, vol. 92, no. 4, p. 40004, 2010.

[18] A. Borzì and S. Wongkaew, "Modeling and control through leadership of a refined flocking system," *Mathematical Models and Methods in Applied Sciences*, vol. 25, no. 2, pp. 255–282, 2015.

[19] M. Caponigro, M. Fornasier, B. Piccoli, and E. Trélat, "Sparse stabilization and control of the Cucker-Smale model," *Mathematical Models and Methods in Applied Sciences*, vol. 25, no. 3, pp. 521–564, 2015.

[20] J. Shao, W. X. Zheng, L. Shi, and Y. Cheng, "Leader-follower flocking for discrete-time Cucker-Smale models with lossy links and general weight function," *IEEE Transactions on Automatic Control*, 2020.

[21] L. Shi, Y. Cheng, J. Huang, and J. Shao, "Cucker-Smale flocking under rooted leadership and time-varying heterogeneous delays," *Applied Mathematics Letters*, vol. 98, pp. 453–460, 2019.

[22] L. Shi, Y. Xiao, J. Shao, and W. X. Zheng, "Containment control of asynchronous discrete-time general linear multiagent systems with arbitrary network topology," *IEEE transactions on cybernetics*, vol. 50, no. 6, pp. 2546–2556, 2019.

[23] K. Yamamoto and M. Okada, "Continuum model of crossing pedestrian flows and swarm control based on temporal/spatial frequency," in *Proceedings of the 2011 IEEE International Conference on Robotics and Automation*, pp. 3352–3357, IEEE, Shanghai, China, May 2011.

[24] W. W. Hager and H. Zhang, "A new conjugate gradient method with guaranteed descent and an efficient line search," *SIAM Journal on Optimization*, vol. 16, no. 1, pp. 170–192, 2005.

[25] W. W. Hager and H. Zhang, "Algorithm 851," *ACM Transactions on Mathematical Software*, vol. 32, no. 1, pp. 113–137, 2006.

[26] W. W. Hager, "Runge-Kutta methods in optimal control and the transformed adjoint system," *Numerische Mathematik*, vol. 87, no. 2, pp. 247–282, 2000.

[27] P. C. Young and Z. Li, "Emergent behavior of Cucker–Smale flocking particles with heterogeneous time delays," *Applied Mathematics Letters*, vol. 86, pp. 49–56, 2018.

[28] L. Göllmann, D. Kern, and H. Maurer, "Optimal control problems with delays in state and control variables subject to mixed control-state constraints," *Optimal Control Applications and Methods*, vol. 30, no. 4, pp. 341–365, 2009.

[29] L. Göllmann, H. Maurer, and H. Maurer, "Theory and applications of optimal control problems with multiple time-delays," *Journal of Industrial and Management Optimization*, vol. 10, no. 2, pp. 413–441, 2014.

[30] L. Armijo, "Minimization of functions having lipschitz continuous first partial derivatives," *Pacific Journal of Mathematics*, vol. 16, no. 1, pp. 1–3, 1966.

Hindawi
Complexity
Volume 2021, Article ID 9983541, 12 pages
https://doi.org/10.1155/2021/9983541

WILEY | Hindawi

Research Article

Group Polarization Based on Agent Emotional Characteristics and Credibility

Shuyang Jiang and Hu Wang

School of Economics and Management, Southeast University, Nanjing 211189, China

Correspondence should be addressed to Hu Wang; whseu94@163.com

Received 22 March 2021; Revised 30 June 2021; Accepted 14 August 2021; Published 20 August 2021

Academic Editor: Yue Song

The frequent occurrence of group polarization in many internet mass incidents produces adverse impacts on the social stability and poses a significant challenge for social management. We construct an opinion evolution model with agent emotional characteristics and credibility based on the Deffuant model to study the evolutionary mechanism of group polarization, which connects the agent's firmness with the agent's opinion, taking a full account of the agent heterogeneity. We analyze the impacts of initial opinion distribution, network structure, and opinion leader on the group opinion evolution. The results show that group polarization is easier to form when the initial opinions are in a normal distribution, and group polarization will also form under the impact of initial minority agents with extreme opinions. Different network structures will pose different effects on the group opinion evolution, and group polarization is easier to form in the small-world network and BA scale-free network. In addition, opinion leader also affects the group opinion evolution, and it will hinder the generation of group polarization.

1. Introduction

At the early stage of internet public events, certain opinion intensity will be generated in the group, and the initial opinion intensity will be strengthened after the interactions between agents in the group as time goes by, and the consistent opinion intensity will become more and more extreme [1, 2]. Moreover, the progress of the internet and appearing of the virtual community also promote such opinions to develop in an extreme way. Since the agents in the network are more inclined to communicate with agents who have similar opinions to theirs than with those whose positions are opposed to or more different from theirs, agents will strengthen their initial opinions after communicating with those holding similar opinions with them [3]. Due to the collaborative filtering of the internet, agents in the network can only see information that is similar to their own opinions or with common preferences [4]. On the other hand, social network sites and platforms divide the internet users into groups by interest, career, and other characteristics, which aggravates the community socialization of "birds of a feather flock together," so that the internet users easily come into contact with agents who have similar views and positions with their own. Meanwhile, the anonymity of online socializing makes it possible to communicate with others easily, thus promoting the formation of opinion consensus [5]. Group consensus refers to the consistency of agent opinions in a group. To some extent, group polarization can be regarded as a special case of group consensus, i.e., agents show the same extreme views on events, and group polarization usually has a negative impact on social stability. Group polarization is essentially an irrational behavior in the network public opinion development [6]. Since the public generally vents their feelings by expressing their views and opinions on hot social events on the internet, such irrational behavior generally generates an extreme opinion, namely, network public opinion polarization. The canard of the internet, as the descendant of group polarization, develops with the formation of group polarization. Group polarization not only seriously disrupts the network order but also may promote violent aggregate behaviors, bringing a serious negative impact on social stability.

As an important component of social physics, opinion dynamics aims to study the complex group behaviors that emerge from opinion interactions between local agents in the whole scope by defining the opinion states of agents and the basic process of the changes of such opinions [7]. The complex and counterintuitive behavior of social systems and their macrolevel collective dynamics can be better understood by means of complexity science. Therefore, the agent based model of complexity science has been widely used in the study of social group behavior [8]. At the early stage of the network public opinion, there are different opinions; with time going by, multiple opinions will gradually reach a consensus, thus forming a unified opinion intensity. Such group polarization shown in public opinion has aroused the wide concern of scholars. Scholars have proposed various opinion evolution models, including the Ising model [9]; the voter model [10, 11]; the Galam majority regular model [12, 13]; the Sznajd model [14, 15]; the bounded confidence model [16, 17]; the extended opinion evolution model [18–23] based on the above models, which generally includes the following aspects when simulating the group behavior: opinion interval (binary opinion, multiple discrete opinion values, and continuous spectrum); network structure (regular network, random network, small-world network, scale-free network and weighted network); opinion update rules (agents will update their opinions under the impact of the decisions of other agents). Such opinion evolution models simplify the social system and assign different characteristics to agents to study the impact of agent characteristics on the development of social network public opinion, such as the impacts of agent opinion acceptance [24, 25], degree of bigotry [26], and memory effects [27] on the group opinion evolution. In reality, agents are independent but interact with and are affected by the whole social environment, so agent's opinion on an event will be influenced both by the intrinsic factors and the external factor (social environment) [28]. With the development of the complex network, interactions between agents, in reality, can be approximately abstracted as a complex network, which can be used to study the interaction of opinions between agents and the evolutionary process of the group behavior under different network structures to reveal the mechanism of social public opinion evolution [29, 30].

Previous studies on opinion evolution models focused on the simplification of actual situations. Namely, it homogenizes all agents or divides agents into groups based on different characteristics. However, in reality, the agent characteristics and relationships between agents are much more complex. In fact, each person has various agent characteristics, and different agents generate different relationships, so these models cannot reflect the group opinion evolution process in reality. Agents change their opinions according to the current opinions and emotions of the agent and others. In other words, the evolution of agent opinions is affected by various internal and external factors, so it is difficult to accurately describe the group opinion evolution through agent interactions [31]. In our daily life, communication is a way for people to pass on the information and express their attitudes. On the other hand, people will also unconsciously change their own emotional characteristics. For example, to achieve certain purposes, the media and opinion leaders always transmit related information through special expressions to cause emotional resonance of the public, thus influencing the evolution of the group behavior upon the emotional changes. Related studies have shown that emotion plays an important role in group polarization [32, 33]. Agent's emotional characteristics can be reflected by the agent's attitude toward the event, while the change of the agent's opinion depends on his/her own emotional characteristics toward the event. Generally, the clearer an agent's attitude toward the event (namely, when the agent's opinion is in the extreme state, such as agree or disagree), the less likely he/she is to be influenced by others and change his or her opinion. Conversely, the unclear or neutral an agent's attitude toward the event, the more likely he or she is to be influenced by others and change his/her opinion [34]. In addition, the interaction of opinions between different agents is also affected by the relationship between them. Generally, the closer the relationship between agents, the more conducive to the interaction of opinions.

The contributions of this paper are as follows. First, we connect the agent's firmness with the agent's opinions on the event to construct an opinion evolution model with agent emotional characteristics and credibility based on the Deffuant model, in which the agent heterogeneity is fully considered. Second, we analyze the impacts of different network structures on the group opinion evolution from the perspective of the complex network. Finally, we introduce the opinion leaders and analyze the role of opinion leaders in the group opinion evolution.

The structure of this paper is organized as follows. Section 2 introduces the methodology. Section 3 presents the results. Section 4 concludes the paper.

2. Methodology

2.1. Deffuant Model. The Deffuant model is a widely used bounded confidence opinion evolution model, which is an update model of the previous binary opinion model [35, 36]. In this model, opinions are continuous values distributed in a defined spectrum, only the agents with similar opinions will communicate with each other, while those with great different opinions will not communicate and keep their original opinions. Meanwhile, the agents with similar opinions will promote the assimilation of their opinions. The Deffuant model can more accurately simulate the actual process of opinion interaction and information transmission. We assume that the opinions of two random agents i and j in the network at time t are $o_i(t)$ and $o_j(t)$, the opinions at time $t+1$ are governed by the following equations:

If

$$\left|o_i(t) - o_j(t)\right| \le d, o_i(t+1) = o_i(t) + \lambda\left(o_j(t) - o_i(t)\right), o_i(t+1) = o_j(t) + \lambda\left(o_i(t) - o_j(t)\right), \quad (1)$$

If

$$\left|o_i(t) - o_j(t)\right| > d, o_i(t+1) = o_i(t), o_j(t+1) = o_j(t), \quad (2)$$

where d is the tolerance threshold, λ is the convergence coefficient, $\lambda \in [0,1]$, indicating the degree of acceptance of other agents' opinions by the agent. λ determines the convergence rate of the opinions in the group, so the higher the λ is, the closer the updated opinion value of an agent is to the opinion value of another agent in the interaction. Generally, λ is 0.5, indicating that the updated opinion is the average of the opinions of the two agents who communicate with each other. However, in reality, each agent has different characteristics and degree of acceptance of other agents' opinions. Agents with the lowest or highest acceptance of opinions are generally minority, the vast majority of agents' acceptance of opinions is in the middle level. Therefore, in this paper, we will use the normal distribution to reflect the agent's degree of acceptance of opinions.

2.2. RA Model. RA model (The Relative Agreement Model) is an extended bounded confidence model [34]. It introduces the uncertainty of opinion into the opinion evolution process and considers that agent's opinion is not a certain value but a spectrum that extends from the opinion value as the center to both ends, and the agent's attitude can be expressed by this opinion spectrum. The changed opinion value of agent i impacted by agent j is related to the size of the overlap between the opinion intervals of two agents, while the opinion interval between the two agents also influences each other. Meanwhile, if the opinion intervals of the two agents are different, the interaction effect is not symmetric. According to the study, the agent with a small opinion interval is more influential. In reality, agents with higher self-confidence are easier to persuade the agents with lower self-confidence because agents with lower self-confidence have a larger opinion spectrum, with unclear opinions, and are easy to change their opinions under the impact of others, while the agents with higher self-confidence have relatively small opinion interval, with clear opinion, and are not susceptible to external influences to change their opinions.

2.3. Construction of the Opinion Evolution Model Based on Agent Emotional Characteristics and Credibility. According to the above studies on the evolution of the group behavior, the change of an agent's opinion is often affected by the change of emotion, while the agents with extreme opinions always have higher emotional characteristics and self-confidence, with a firm attitude, so they are less likely to be influenced by the external factors and change their opinions. The psychological research [37] also confirms that extremists are more resistant to externally generated than moderates, and extremity does not necessarily beget irrationality, supporting the above hypothesis. For example, as for some hot events, some agents are not interested in it or consider that these events are not related to themselves, they may not express an obvious attitude tendency or keep a neutral attitude toward these events, while some agents are interested in it or consider that these events are closely related to themselves, they may pay much attention to the development of these events and make comments with clear attitude tendency. These agents have strong emotional characteristics and firm attitudes to stick to their opinions and are not easily influenced by external factors and change their attitudes. We assume that the range of agent's opinion is [0 : 1]. With the increase of the opinion value, the agent will have a clearer attitude. When the opinion value is 1, the agent has a clear attitude tendency and emotional characteristics; when the opinion value is 0, the agent has an equivocal attitude tendency, with a neutral attitude and lower emotional characteristics. Similarly, when the range of agent's opinion is [−1 : 0], the opinion value is −1, indicating the agent is clearly opposed to the event, with strong emotional characteristics. In addition, credibility is also a fact that affects the interaction of opinions of different agents. In this paper, we construct an opinion evolution model with agent emotional characteristics and credibility on the basis of the bounded confidence model [38].

$$\begin{aligned}
&\left|o_i(t) - o_j(t)\right| \le R_{ij}d_i, o_i(t+1) = o_i(t) + \lambda\left(o_j(t) - o_i(t)\right),\\
&\left|o_i(t) - o_j(t)\right| \le R_{ij}d_j, o_i(t+1) = o_j(t) + \lambda\left(o_i(t) - o_j(t)\right),\\
&\left|o_i(t) - o_j(t)\right| > R_{ij}d_i, o_i(t+1) = o_i(t),\\
&\left|o_i(t) - o_j(t)\right| > R_{ij}d_j, o_j(t+1) = o_j(t),
\end{aligned} \quad (3)$$

where R_{ij} indicates the credibility degree of agent i to agent j, $R_{ij} \in [0,1]$. The credibility can be calculated by the equation: $R_{ij} = C_{ij} + 1/N_i$, where C_{ij} indicates the number of common nodes of the two agents, N_i indicates the number of nodes of the agent i. The larger the percentage of common nodes with agent j on the connected nodes of agent i, the higher the credibility degree of agent i to agent j. When $C_{ij} = N_i - 1$ and $R_{ij} = 1$, indicating the maximum of credibility degree of agent i to agent j. $d_i = F(|o_i|)$ and $d_j = F(|o_j|)$ indicate the agent's tolerance threshold is a power function of its opinion value, and the tolerance threshold becomes smaller as the opinion value gets larger, this can be expressed by the following equation:

$$d_i = d_m - |o_i|^x (d_m - d_n). \quad (4)$$

When the opinion value $o_i = 0$, $d_i = d_m$; when the opinion value $o_i = 1$, $d_i = d_n$. d_m and d_n represent the agent's tolerance thresholds when opinion value is 0 and 1,

respectively, i.e., d_m is the tolerance of the agents with neutral attitudes and lower emotional characteristics and d_n is the tolerance of the agents with the extreme opinions. In this paper, we assume that the agent with the extreme opinion has a lower tolerance threshold, i.e., d_n is a small positive and constant value ($d_n = 0.001$) and d_m is one of the parameters of the simulations ($d_{\max} = 0.5$). We study the group behavior for different values of x. The larger the value of x is, the larger d_i is for the same value of o_i. This indicates that the agents with intermediate opinions present a larger tolerance threshold and are more likely to be influenced by other agents when the value of x is larger. When parameter x is 1/2, 1 and 2, respectively, the graph of function of the tolerance threshold with the opinion value is shown in Figure 1.

As shown in Figure 1, when $o_i > o_j$, $d_i < d_j$, indicating that during the interaction between agents with different opinion values, agent j has a higher tolerance threshold due to the lack of clear opinion intensity, and is accordingly more susceptible to the influence of others. So agent j is easy to change his/her opinion under the impact of agent i during the interaction, while agent i has a clear opinion and lower tolerance threshold, during the interaction with agent j, the opinion of agent j is outside the tolerance threshold of agent i, so agent i is free from the influence of agent j and keeps his/her initial opinion. Due to the asymmetry of the opinion evolution, the agents with the opinion value close to 0 are more easily affected by the external factors to generate extreme opinions due to his/her larger tolerance threshold, and the agent with extreme opinions is not susceptible to the influence from others due to his/her lower tolerance threshold and keeps his/her initial opinion. Under the impact of this opinion evolution mechanism, every agent may generate extreme opinions, and the network public opinion will eventually form group polarization.

3. Simulation Results

3.1. Effects of the Random Distribution of the Initial Opinions. In this paper, we firstly analyze the group opinion evolution under the condition that the initial opinions are randomly distributed. The initial opinions of the agents are randomly distributed in [0 : 1], and the globally coupled network in the regular network is used as the basic network. We assume the number of agents in the group is $N = 2000$, the agents in the group are fully connected, and any two agents can interact with each other randomly, λ follows a normal distribution with a mean value of 0.5 and a standard deviation of 0.2, the same below. The simulation result in this paper is an average of 200 times. When x is 1/2, 1 and 2, and d_m takes different values, the group opinion evolution is shown in Figures 2–4.

Figures 2–4 show the process of the group opinion distribution evolution under different values of d_m, when x is 1/2, 1 and 2. The abscissa indicates the opinion values of agents, the ordinates indicate the evolution algebra, and the vertical coordinate indicates the proportion of the number of agents holding same opinion value to the total agent in the group. According to Figures 3 and 4, when x is 1 and 2, respectively, d_m takes a smaller value; the various opinion clusters are quickly formed. When the opinion value is close to 1, the tolerance threshold of agents is the minimum, and the

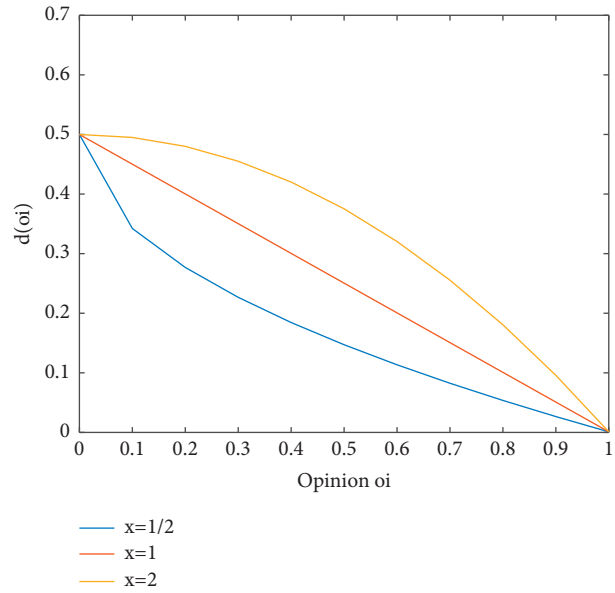


Figure 1: Graph of the function of the tolerance threshold d_i with the opinion value o_i, for $x = 1/2, 1, 2$ and $d_{\max} = 0.5$ and $d_n = 0.001$. The larger the value of x is, the larger d_i is for the same value of o_i.

proportion of agents around this opinion is also minimal. As the gradual increase of d_m, the tolerance threshold of the agents in the group gradually gets larger, and the probability of interaction between agents increases. Finally, the agent opinion values are basically close to 1 under the impact of agents with extreme opinions, thus forming group polarization. Compared with the condition when $x = 1$, the value of d_m is lower when group polarization is formed under $x = 2$ (when $x = 1$, $d_m = 0.55$; when $x = 2$, $d_m = 0.4$), which is because the larger tolerance threshold when x takes on a larger value. When d_m takes the same value, the average interactive threshold of agents in the group with a large x is greater. Agents are more likely to interact opinions with each other. According to Figure 2, when $x = 1/2$, the process of the opinion distribution evolution shows a similar tendency as that when $x = 1$ and $x = 2$. However, after group polarization is formed, d_m is at its maximum ($d_m = 0.9$), because the smaller tolerance threshold when $x = 1/2$ leads to a lower average interactive threshold of agents in the group, thus decreasing the probability of opinion interaction between agents. Only when the tolerance threshold of the agents in the group is large enough, group polarization will be formed under the impact of the extreme opinion.

3.2. Effects of the Normal Distribution of the Initial Opinions. For the purpose of studying the process of the group opinion distribution evolution, the initial opinion values of agents are randomly distributed in the above studies. In fact, the initial opinions of the public on events are not randomly distributed in most cases, and the public will express their own opinions on events when it occurs. Generally, the opinions of most of the public on events will be concentrated in the middle region; only minorities will hold the extreme opinions or keep a neutral attitude. As time goes by, the opinions of most agents will gradually become unified under the impact of the minorities with extreme opinions, thus

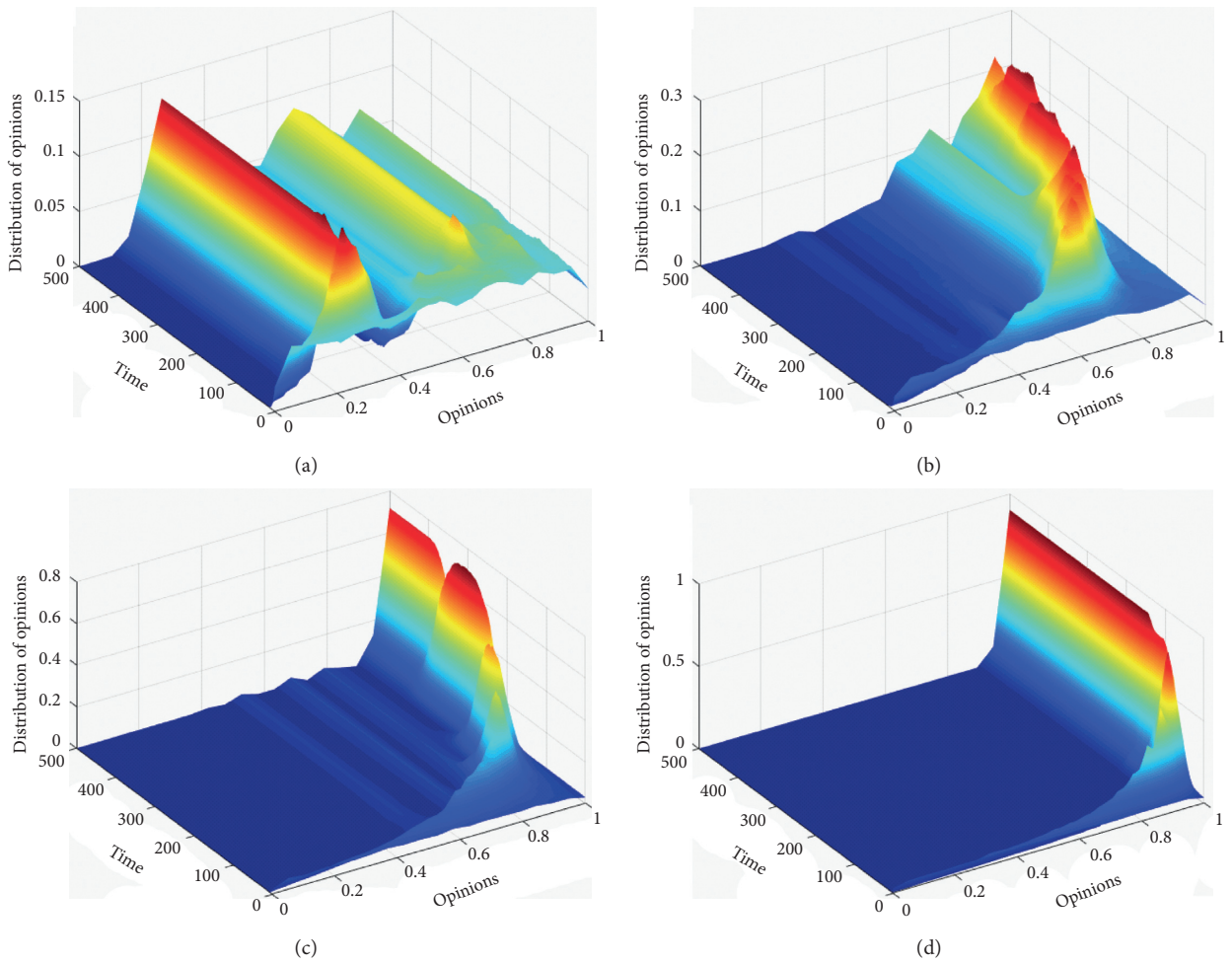


Figure 2: Group opinion distribution evolution when $x = 1/2$. With the increase of d_m, the proportion of agents whose opinion value is close to 1 is increasing. (a) $d_m = 0.25$. (b) $d_m = 0.5$. (c) $d_m = 0.65$. (d) $d_m = 0.9$.

forming group polarization. To study the process of the group opinion distribution evolution under such conditions, we use a normal distribution with a mean value of 0.5 and a standard deviation of 0.2 as the distribution of the initial opinions of agents, in which the initial opinions of the vast majority of agents are between 0 and 1, while the agents with the opinions close to 0 and 1 account for only a small part. When $x = 2$ and d_m takes different values, the group opinion evolution is shown in Figure 5.

According to Figure 5, when the initial opinions are normally distributed, group polarization of opinions will be finally formed, with a smaller value of d_m ($d_m = 0.3$). When d_m takes a small value, there are various opinion clusters concentrated in the middle region, which conforms to the group opinions being normally distributed. With the increase of d_m, the agent opinions gradually cluster around the opinion value of 1 under the impact of the extreme opinion. Since the opinions are normally distributed, the opinion values of most of the agents are concentrated in the middle region. Compared with the condition under the random distribution of the opinions, the opinion values of most agents are more concentrated, which is more conducive to opinion interaction between agents, so group polarization can be formed when d_m takes a small value. This also reflects that in daily life, under the impact of the minority with extreme opinions, the lower proportion of agents with extreme opinions at the early stage will promote the formation of group polarization when most agents have no clear opinion intensity with a higher tolerance threshold of the group, and the higher the proportion of agents with larger tolerance thresholds, the more likely group polarization will be formed.

3.3. Effects of Different Network Structures. Constructing a network model with complex structures to simulate the network relationship in real society is a common way to study the specific social problems. To study the evolution of information dissemination in social networks, scholars have constructed network models with various complex structures, including the regular network, the random network, the small-world network, the BA scale-free network, and others that widely used in researches. In the social network, the information dissemination process is related to the agent characteristics and network structure. To analyze the impacts of different network topology on the information

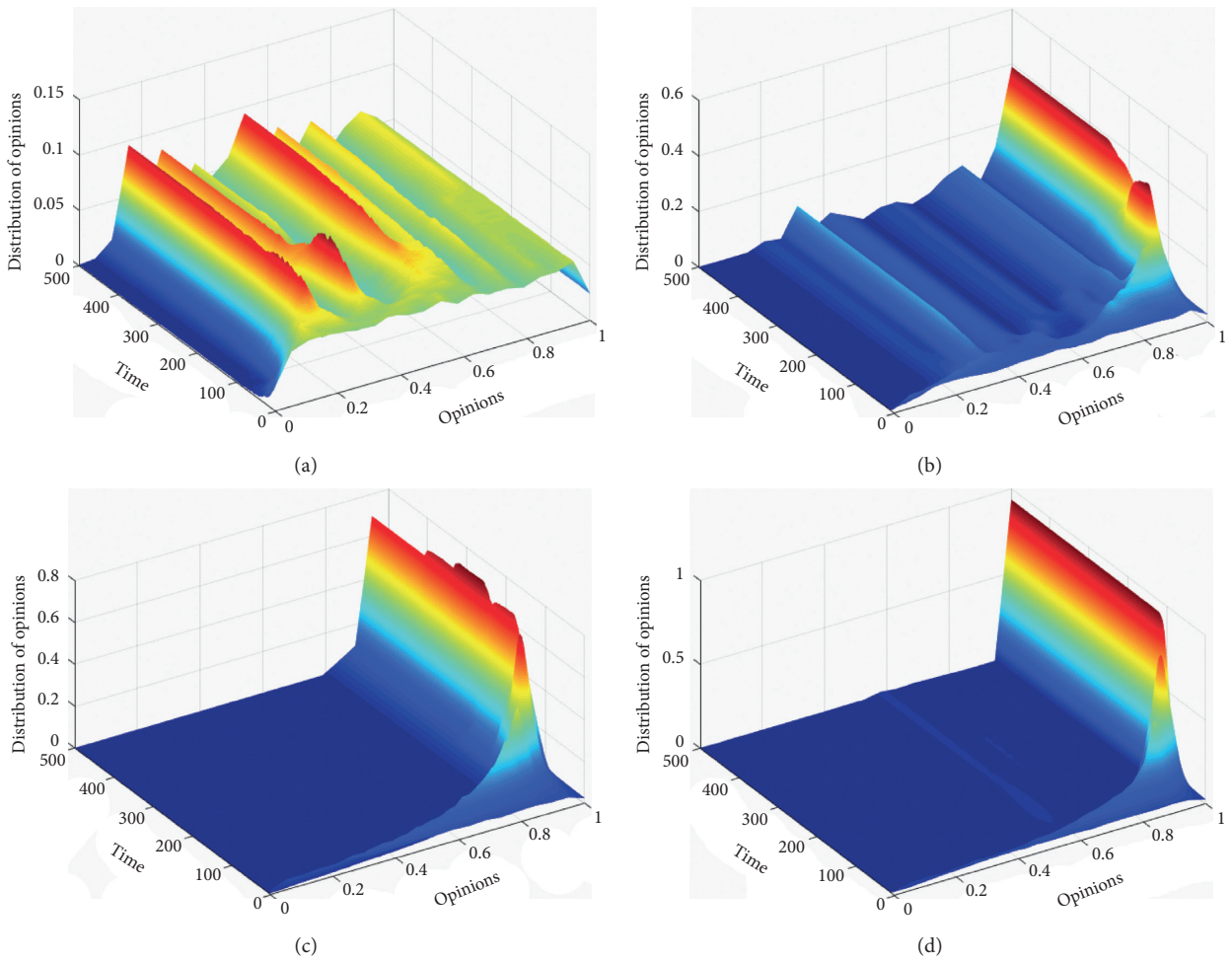


Figure 3: Group opinion distribution evolution when $x = 1$. Compared with the condition when $x = 1/2$, the value of d_m is lower when group polarization is formed under $x = 1$. (a) $d_m = 0.15$. (b) $d_m = 0.3$. (c) $d_m = 0.45$. (d) $d_m = 0.55$.

dissemination process in the social network, we apply the nearest-neighbor coupled network in the regular network, the random network, the small-world network, and the BA scale-free network with the node number of 2000, respectively. The average degree is set to be consistent based on the construction rules of these networks so that the evolution results of network structures have comparability to the greatest extent. The initial opinions of agents are randomly distributed in [0: 1]. When the parameter $x = 2$, and d_m takes different values, the processes of group opinion distribution evolution in different network structures are shown in Figures 6–9.

According to Figures 6–9, in different network structures, group polarization can be formed under different conditions. When d_m takes small value, there are various opinion clusters in these four networks, with the increase of the value of d_m, group polarization will be finally formed. Meanwhile, when opinion polarization is formed, the values of d_m in these four networks are nearest-neighbor coupled network > random network > BA scale-free network > small-world network. It can be seen from the above analysis, when group polarization is formed in the nearest-neighbor coupled network, the value of d_m is at its minimum ($d_m = 0.4$), where the smaller the value of d_m, the easier it is to form group polarization. When the average tolerance threshold is lower, the consensus of the agent's opinions in the group can be reached, thus forming group polarization. Due to the interaction between agents, the globally coupled network, with same node number, has the maximum average degree and clustering coefficients and the minimum average path length, so it is most likely to form group polarization. The small-world network has larger clustering coefficients and a shorter average path length, which increases the probability of interactions between different agents in the group, thus promoting the interaction between agents. The clustering coefficients of BA scale-free network and regular network are lower than that of the small-world network, and the degree of interaction between agents is also lower than that of small-world network. However, due to "priority connection," one of the characteristics of BA scale-free network, there are a few nodes with larger scale, which connect with a large number of nodes, increasing the probability of interaction between other agents and the agent to achieve a full interaction between agents, so the BA scale-free networks are more likely to form opinion polarization than the regular networks. In the nearest-neighbor coupled network, the

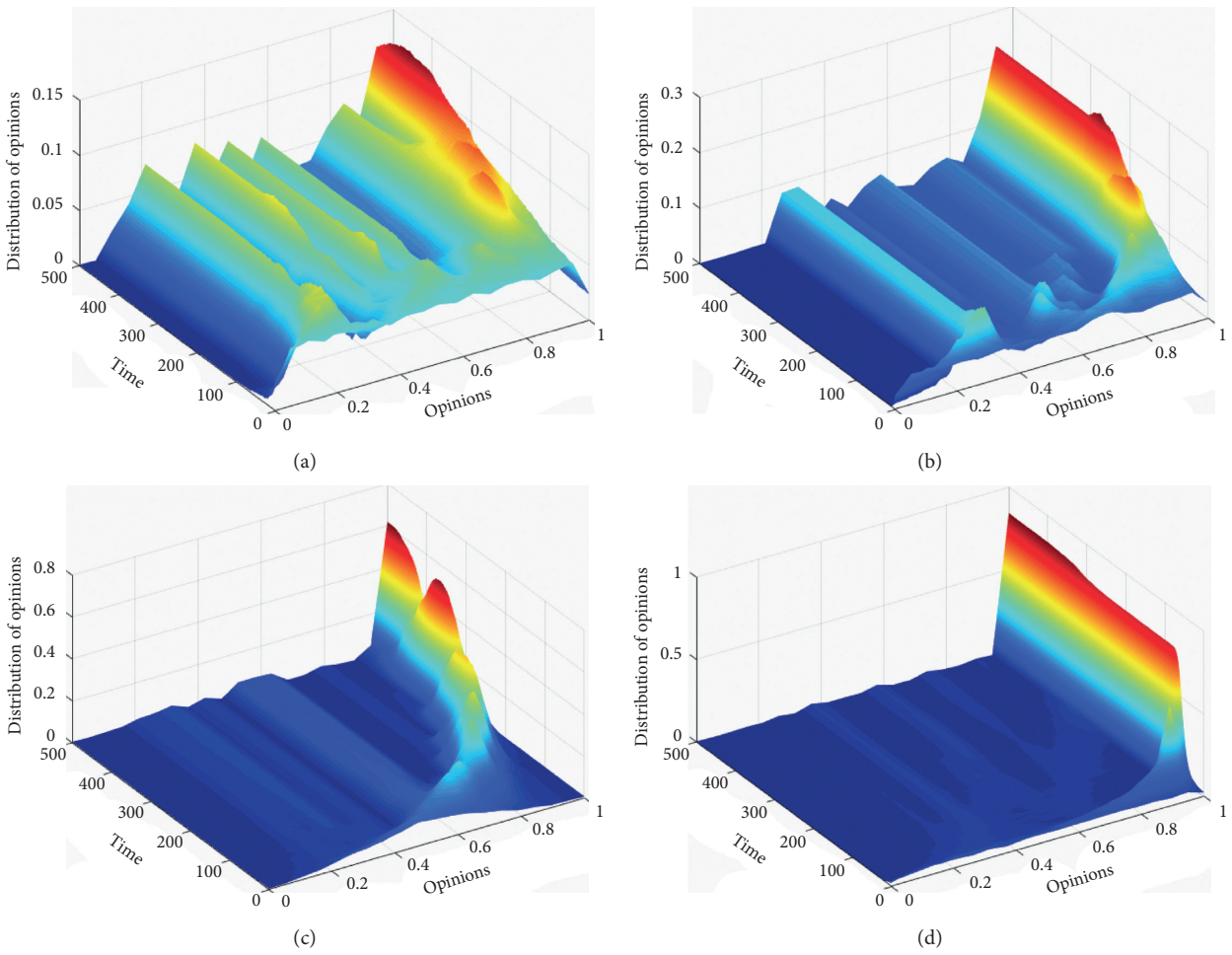


FIGURE 4: Group opinion distribution evolution when $x = 2$. Compared with the conditions when $x = 1/2$ and $x = 1$, the value of d_m is lower when group polarization is formed under $x = 2$. (a) $d_m = 0.1$. (b) $d_m = 0.2$. (c) $d_m = 0.3$. (d) $d_m = 0.4$.

agents only interact with their adjacent agents. Although its larger clustering coefficient can promote the interaction between agents, its larger average path length will confine the interaction between agents to different local areas, making opinion clusters not able to interact with each other, so the value of d_m is at its maximum when group polarization is formed.

3.4. Effects of Opinion Leaders. Opinion leaders, as influential public figures, play an important role in the development of the network public opinion. Generally, they will affect the opinions of the public and lead the development direction of the network public opinion, thus promoting group polarization. Due to the "priority connection" characteristic of scale-free network, there are a few nodes with a larger degree, which accounts for a small proportion of the total number of nodes in the group, but these nodes connect with a large number of other nodes that play a significant role in the whole network. In this paper, we take the BA scale-free network as the basic network, the initial opinions of agents are randomly distributed in [0: 1], and two nodes with the largest degree are selected as the "opinion leaders." Due to the stronger professional competence and knowledge reserve, opinion leaders always firmly insist on their opinion and are less likely to be influenced by the external factors to change their opinions, so the opinion acceptance of the opinion leaders is set as $\lambda \in [0, 0.01]$. Similarly, the range of agent's opinion is extended to [−1: 1], among which, [−1: 0] and [0: 1] indicate the disagreement and agreement toward the event, respectively, and two nodes with larger scale are selected to make their opinion values within the two opinion intervals, respectively. When $x = 2$, and d_m takes different values, the processes of group opinion distribution evolution are shown in Figures 10 and 11.

In Figure 10, two nodes with the largest degree are selected as the opinion leaders, with the initial opinion values of the opinion leaders are 0.5. When $d_m = 0.45$, there are two opinion clusters in the group, among which, the opinion value of the opinion cluster with the larger agent proportion is close to 0.5, while the other is close to 0.8; when $d_m = 0.9$, there is only one opinion cluster in the group, the opinion value of the opinion cluster is close to 0.8. In Figure 4, when $d_m = 0.45$, group polarization has been formed, with the opinion value close to 1, indicating the opinion leaders hinder the formation of group polarization in a certain

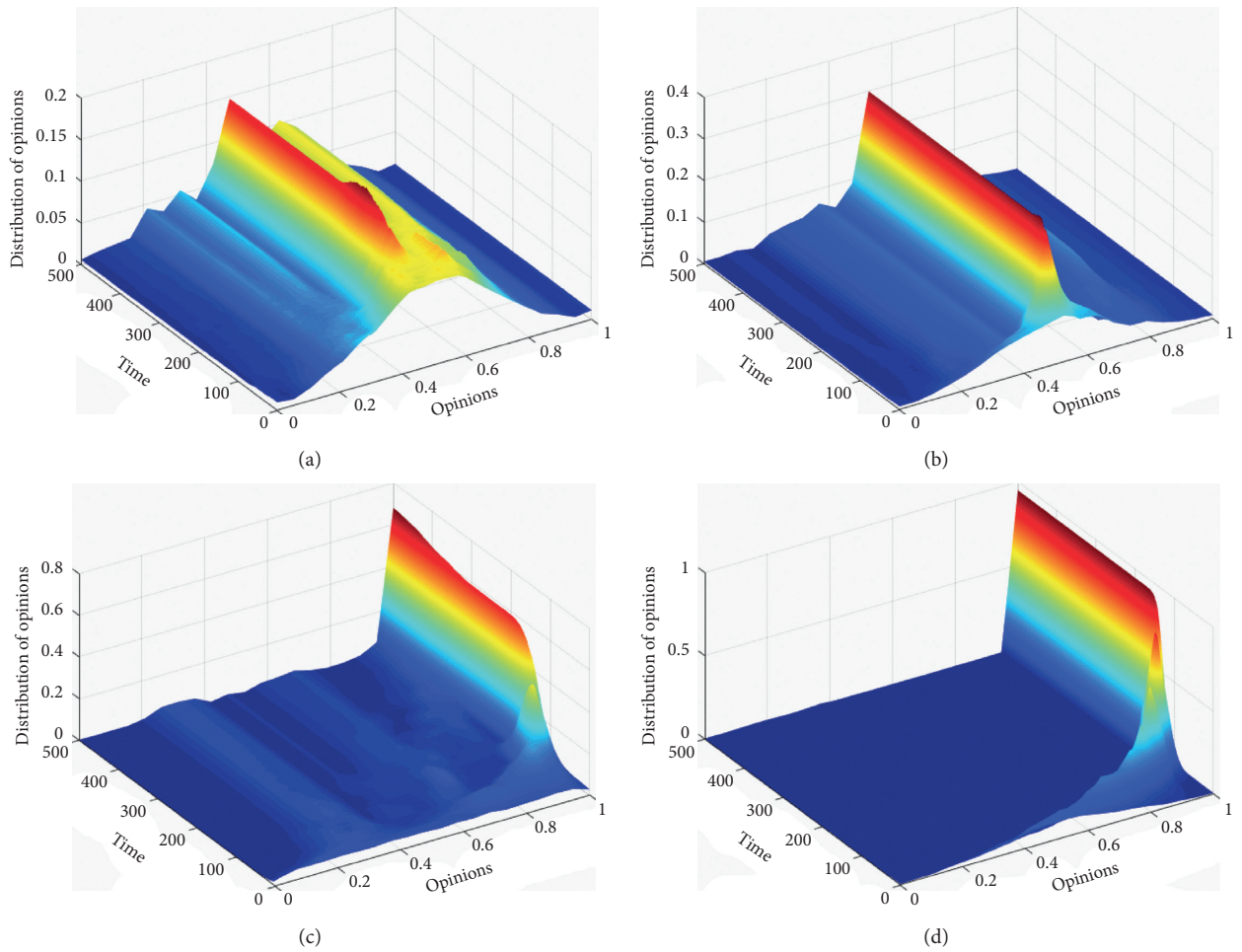


Figure 5: Group opinion distribution evolution under the normal distribution of the initial opinions. Compared with the random distribution of the initial opinions, the value of d_m is lower when group polarization is formed under the normal distribution of the initial opinions. (a) $d_m = 0.1$. (b) $d_m = 0.15$. (c) $d_m = 0.2$. (d) $d_m = 0.3$.

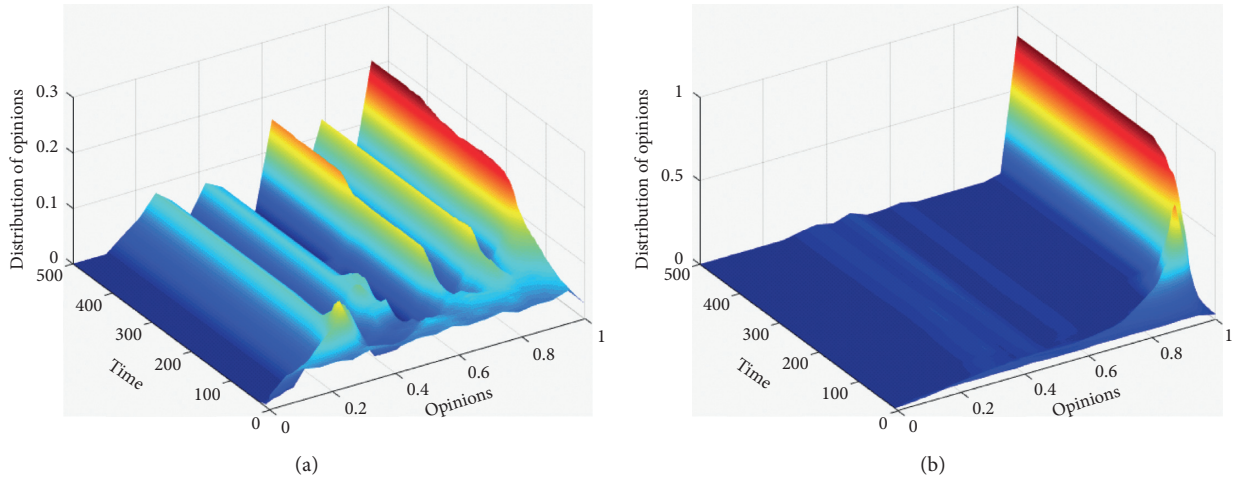


Figure 6: Group opinion distribution evolution in the nearest-neighbor coupled network. Group polarization is formed when $d_m = 0.65$. (a) $d_m = 0.3$. (b) $d_m = 0.65$.

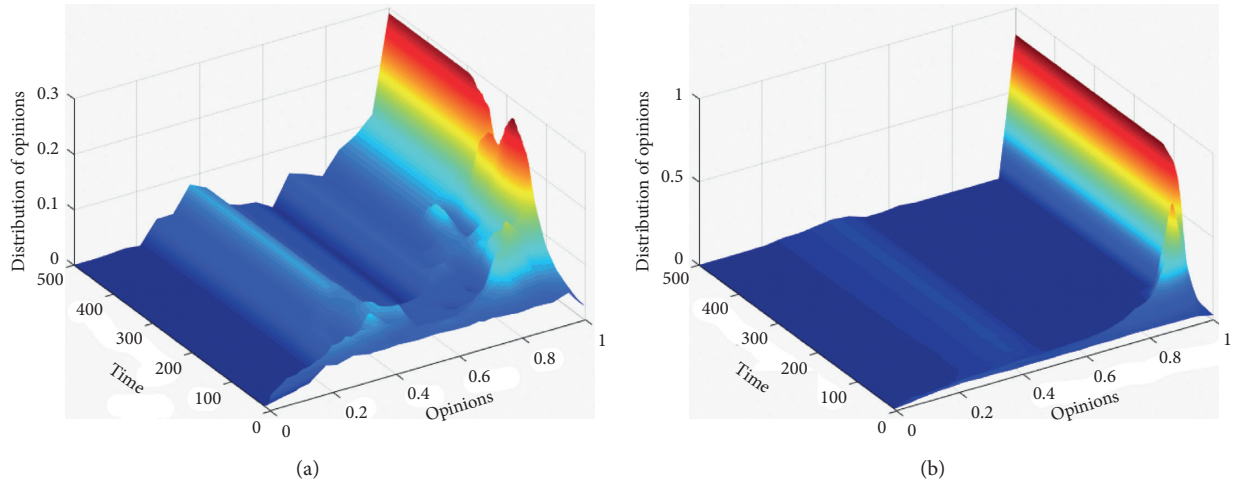


Figure 7: Group opinion distribution evolution in the random network. Group polarization is formed when $d_m = 0.6$. (a) $d_m = 0.3$. (b) $d_m = 0.6$.

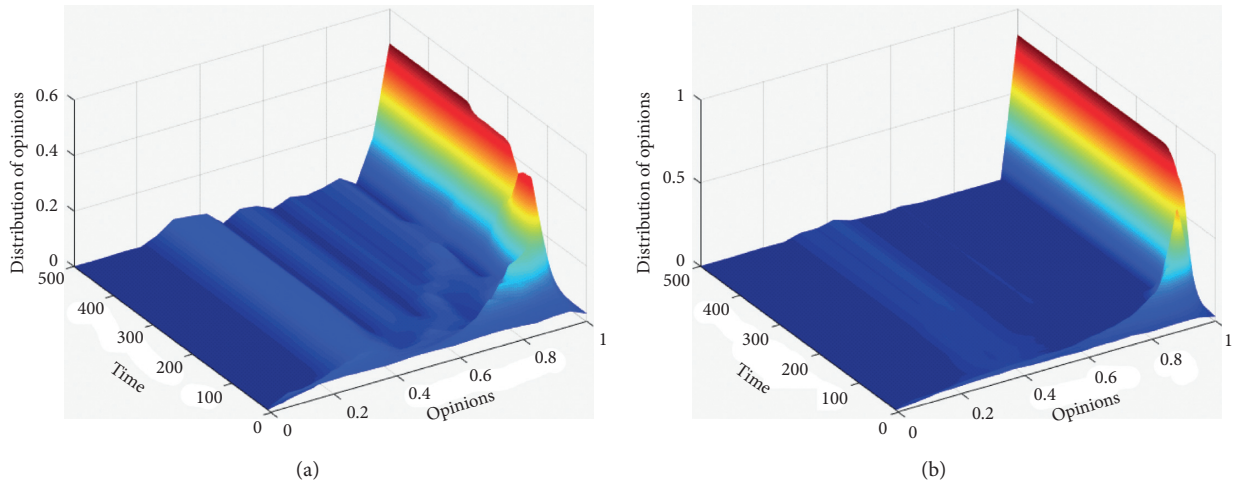


Figure 8: Group opinion distribution evolution in the small-world network. Group polarization is formed when $d_m = 0.52$. (a) $d_m = 0.3$. (b) $d_m = 0.52$.

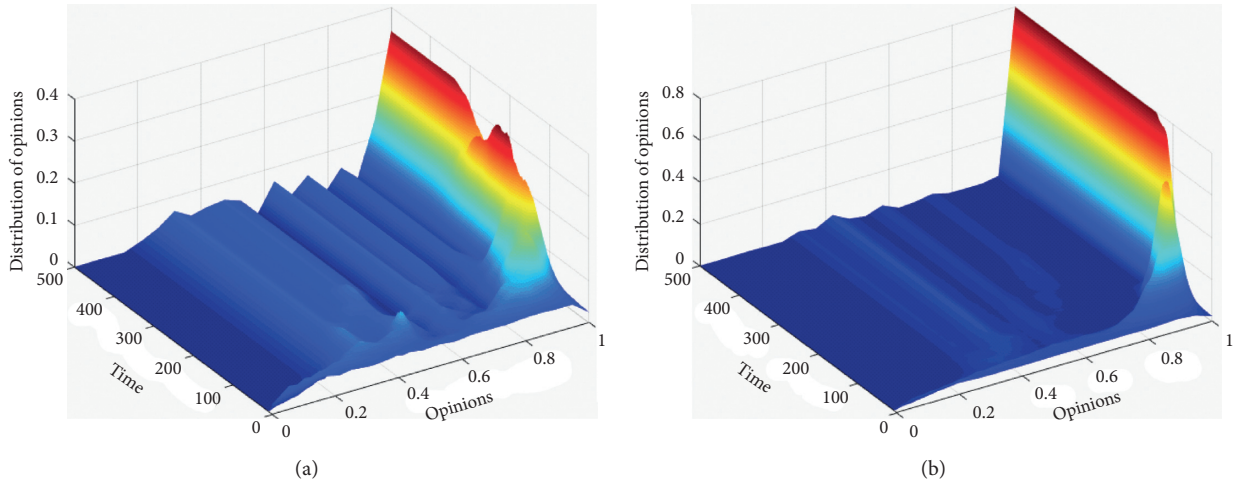


Figure 9: Group opinion distribution evolution in the BA scale-free network. Group polarization is formed when $d_m = 0.58$. (a) $d_m = 0.3$. (b) $d_m = 0.58$.

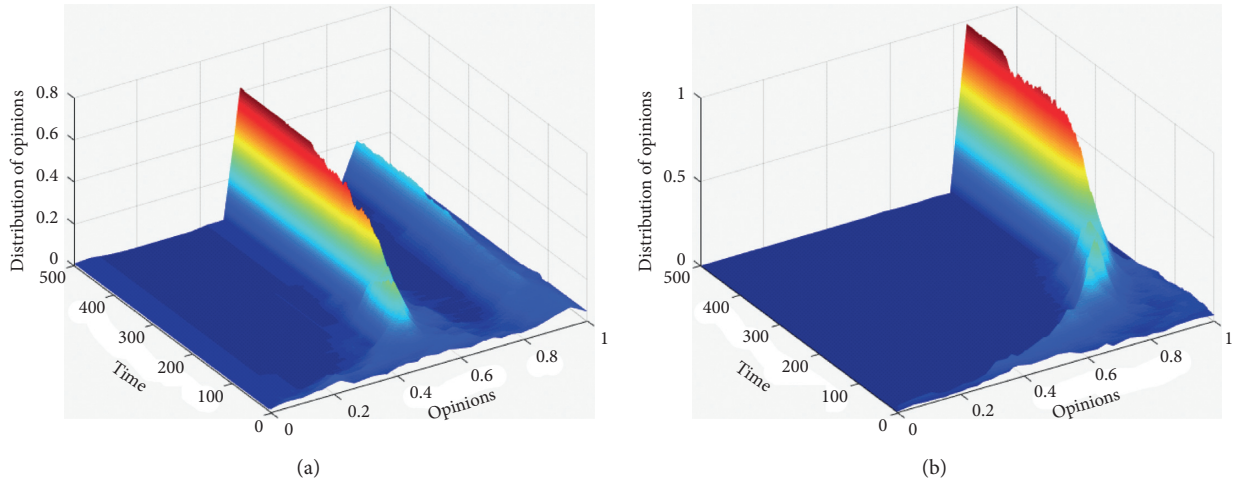


FIGURE 10: Two opinion leaders with positive opinion values. There are two opinion clusters when $d_m = 0.45$, the opinion value of the opinion cluster with the larger agent proportion is close to 0.5, while the other is close to 0.8; there is only one opinion cluster when $d_m = 0.9$, the opinion value of the opinion cluster is close to 0.8. (a) $d_m = 0.45$. (b) $d_m = 0.9$.

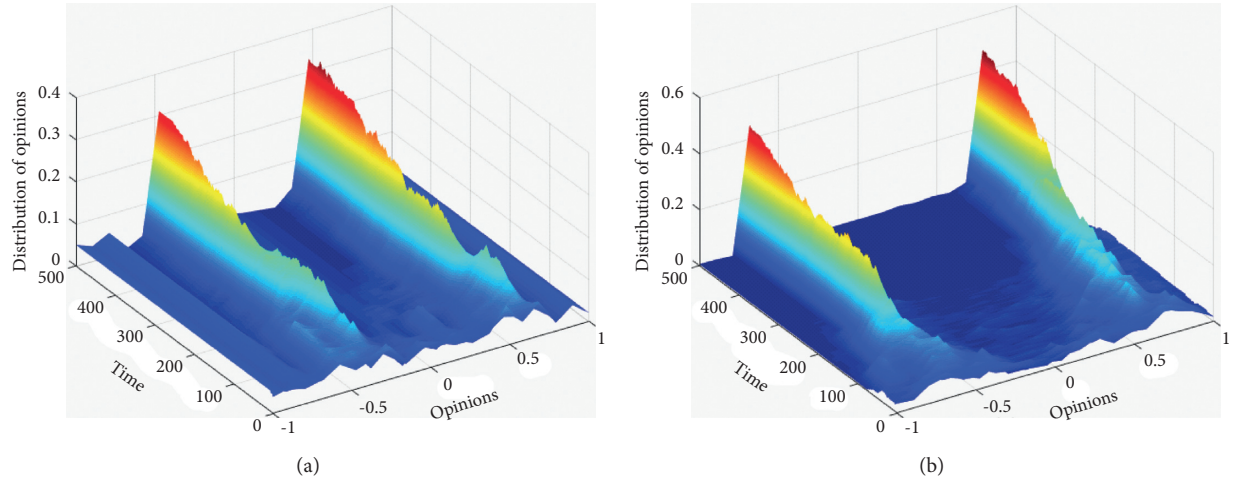


FIGURE 11: Two opinion leaders with opposite opinion values. There are two opinion clusters when $d_m = 0.45$; the opinion values of these two clusters are close to 0.5 and −0.5, respectively; there are two opinion clusters when $d_m = 0.9$; the opinion values of these two clusters are close to 0.8 and −0.8, respectively. (a) $d_m = 0.45$. (b) $d_m = 0.9$.

extent. When d_m takes a smaller value, agents in the group will be impacted by the opinion leaders, forming the opinion cluster with the similar opinion with the opinion leaders. With the increase of d_m, agents in the group have a higher tolerance threshold, and the group opinion will gradually close to the opinion value of 1 under the impact of the agents with extreme opinions. However, combined with the impact of opinion leaders, when group polarization is finally formed, d_m takes a higher value, and finally, the opinion value is around 0.8. In Figure 11, two nodes with the largest degree are selected as the opinion leaders, with the initial opinion values of the opinion leaders are 0.5 and −0.5, respectively. When $d_m = 0.45$, there are two opinion clusters with larger agent proportion of the group, and the opinion values of these two clusters are close to 0.5 and −0.5, respectively. When $d_m = 0.9$, group polarization is formed near the endpoint values of the two opinion intervals, indicating that the opinion values of other agents in the group will move toward the opinion value of the two opinion leaders, and the opinion values of the finally formed opinion polarization are near the 0.8 and −0.8, respectively.

4. Conclusion

In this paper, we construct an opinion evolution model with agent emotional characteristics and credibility based on the extended Deffuant model in the bounded confidence models. We make full use of the theory of opinion interval of the RA model to connect the agent firmness with opinion values through agent emotional characteristics and credibility. The model constructed in this paper qualitatively interprets the reason for group polarization and considers the heterogeneity of agents in reality. In addition, in order to quantitatively analyze group polarization, in reality, we

consider the different initial opinion distribution and several ordinary complex network structures and the impact of opinion leaders on the group opinion evolution.

The results show that compared with that under the random distribution of the initial opinions, group polarization is formed under the normal distribution of initial opinions, and the value of d_m is smaller. Under the impact of the minority with extreme opinions, the lower proportion of agents with extreme opinions at the early stage will promote the formation of group polarization when the most agents have no clear opinion intensity with a higher tolerance threshold of the group, and the higher the proportion of agents with larger tolerance thresholds, the more likely group polarization will be formed. Different network structures have different impacts on the group opinion evolution. The larger clustering coefficients and shorter average path length are more likely to facilitate opinion interactions between agents in a group. Since any two agents in the globally coupled network can interact with each other, it has the largest clustering coefficients and the shortest average patch length, so it is more likely to achieve group polarization. Relying on the "priority connection" characteristic of the BA scale-free network, a large number of nodes in the network conduct the opinion interaction with the node at a higher probability; when the tolerance threshold is lower, group polarization will be formed. Opinion leaders will impact the group opinion evolution. When d_m takes a small value, the agents in the group will be affected by the opinion leaders and form the opinion cluster with similar opinion values as that of opinion leaders. Meanwhile, when group polarization is formed, the value of d_m is larger.

According to the results, we can find that the variable parameter d_m plays an important role in the group opinion evolution. The larger the d_m is, the larger the average tolerance threshold of the agents in the group, the higher the proportion of the agents with a larger tolerance threshold, which is more conducive to opinion interaction among agents. Meanwhile, under the effect of the agents with extreme opinions, the unified opinions are gradually formed, thus forming group polarization. However, the agents with extreme opinions have strong emotional characteristics, lower tolerance threshold, and are not susceptible to the external factors, so it is difficult to change the opinion of these agents with extreme opinions. We can decrease the probability of polarization by lowering the tolerance threshold of the group. For example, at the early stage of the event, governments can guide the public to form clear opinion intensity and fully play the roles of the mainstream media and opinion leaders in the guidance process, thus making the network public opinion develops toward a better direction. The application of social software makes the connections between agents with the structural characters of small-world network and BA scale-free network, which promote group polarization. Therefore, the government should strengthen the supervision of the information dissemination in social software. In addition, regulator can increase the cost of manipulating the public opinion by increasing the punishment for the malicious manipulation behavior, so as to reduce the manipulation behavior of opinion leaders.

The conclusions of the paper provide a certain theoretical foundation for interpreting the reasons for group polarization in our daily life and put forward corresponding countermeasures to cope with the vicious group polarization and to play the positive effect of group polarization. However, there are many shortcomings in this paper. The opinion interaction between agents, in reality, depends on whether they are connected or not, and the strength of this connection relationship also has a significant impact on the tolerance between agents. Therefore, the group opinion evolution based on weighted network structures is more realistic. In addition, the network structures are varied over time. The impact of the network structure with dynamic characteristics on the group opinion evolution is the next research focus.

Data Availability

No data were used in the study.

Disclosure

Shuyang Jiang and Hu Wang are co-first authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Shuyang Jiang and Hu Wang contributed equally to this work.

Acknowledgments

This research was supported by the Fundamental Research Funds for the Central Universities and Scientific Research Foundation of Graduate School of Southeast University (YBPY2146).

References

[1] A. Fang, K. Yuan, J. Geng, and X. Wei, "Opinion dynamics with bayesian learning," *Complexity*, vol. 2020, Article ID 8261392, 5 pages, 2020.

[2] G. Olivares, J. P. Cárdenas, J. C. Losada, and J. Borondo, "Opinion polarization during a dichotomous electoral process," *Complexity*, vol. 2019, Article ID 5854037, 9 pages, 2019.

[3] C. Castellano, S. Fortunato, and V. Loreto, "Statistical physics of social dynamics," *Reviews of Modern Physics*, vol. 81, no. 2, pp. 591–646, 2009.

[4] J. Wu, J. Chang, Q. Cao, and C. Liang, "A trust propagation and collaborative filtering based method for incomplete information in social network group decision making with type-2 linguistic trust," *Computers & Industrial Engineering*, vol. 127, pp. 853–864, 2019.

[5] X. Chen, Z. Wu, H. Wang, and W. Li, "Impact of heterogeneity on opinion dynamics: heterogeneous interaction model," *Complexity*, vol. 2017, Article ID 5802182, 10 pages, 2017.

[6] M. Cao, J. Wu, F. Chiclana, R. Ureña, and E. Herrera-Viedma, "A personalized feedback mechanism based on maximum

harmony degree for consensus in group decision making," *IEEE Transactions on Systems, Man and Cybernetics: Systems*, 2019.

[7] M. Perc, "The social physics collective," *Scientific Reports*, vol. 9, no. 1, p. 16549, 2019.

[8] D. Helbing, D. Brockmann, T. Chadefaux et al., "Saving human lives: what complexity science and information systems can contribute," *Journal of Statistical Physics*, vol. 158, no. 3, pp. 735–781, 2015.

[9] S. Galam, "Rational group decision making: a random field ising model at $T = 0$," *Physica A: Statistical Mechanics and its Applications*, vol. 238, no. 1-4, pp. 66–80, 1997.

[10] M. Mobilia, "Does a single zealot affect an infinite group of voters?" *Physical Review Letters*, vol. 91, no. 2, Article ID 028701, 2003.

[11] M. Mobilia, A. Petersen, and S. Redner, "On the role of zealotry in the voter model," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 8, pp. 266–276, 2007.

[12] S. Galam and F. Jacobs, "The role of inflexible minorities in the breaking of democratic opinion dynamics," *Physica A: Statistical Mechanics and its Applications*, vol. 381, pp. 366–376, 2007.

[13] S. Galam, "Minority opinion spreading in random geometry," *The European Physical Journal B*, vol. 25, no. 4, pp. 403–406, 2002.

[14] K. Sznajd-Weron, M. Tabiszewski, and A. M. Timpanaro, "Phase transition in the sznajd model with independence," *EPL (Europhysics Letters)*, vol. 96, no. 4, p. 48002, 2011.

[15] A. Benatti, H. F. de Arruda, F. N. Silva, C. H. Comin, and L. da Fontoura Costa, "Opinion diversity and social bubbles in adaptive sznajd networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2020, no. 2, Article ID 023407, 2020.

[16] G. Deffuant, "Comparing extremism propagation patterns in continuous opinion models," *Journal of Artificial Societies and Social Simulation*, vol. 9, no. 3, pp. 1–8, 2006.

[17] P. Sobkowicz, "Extremism without extremists: deffuant model with emotions," *Frontiers in Physics*, vol. 3, p. 17, 2015.

[18] R. Ureña, G. Kou, Y. Dong, F. Chiclana, and E. Herrera-Viedma, "A review on trust propagation and opinion dynamics in social networks and group decision making frameworks," *Information Sciences*, vol. 478, pp. 461–475, 2019.

[19] A. Woolcock, C. Connaughton, Y. Merali, and F. Vazquez, "Fitness voter model: damped oscillations and anomalous consensus," *Physical Review E*, vol. 96, no. 3, Article ID 032313, 2017.

[20] E. Hashemi, M. Pirani, A. Khajepour, B. Fidan, A. Kasaiezadeh, and S.-K. Chen, "Opinion dynamics-based vehicle velocity estimation and diagnosis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 7, pp. 2142–2148, 2017.

[21] P. Frasca, S. Tarbouriech, and L. Zaccarian, "Hybrid models of opinion dynamics with opinion-dependent connectivity," *Automatica*, vol. 100, pp. 153–161, 2019.

[22] Y. Dong, Y. Fan, H. Liang, F. Chiclana, and E. Herrera-Viedma, "Preference evolution with deceptive interactions and heterogeneous trust in bounded confidence model: a simulation analysis," *Knowledge-Based Systems*, vol. 175, pp. 87–95, 2019.

[23] C. Altafini and F. Ceragioli, "Signed bounded confidence models for opinion dynamics," *Automatica*, vol. 93, pp. 114–125, 2018.

[24] Z. Cheng, Y. Xiong, and Y. Xu, "An opinion diffusion model with decision-making groups: the influence of the opinion's acceptability," *Physica A: Statistical Mechanics and its Applications*, vol. 461, pp. 429–438, 2016.

[25] K. Li, H. Liang, G. Kou, and Y. Dong, "Opinion dynamics model based on the cognitive dissonance: an agent-based simulation," *Information Fusion*, vol. 56, pp. 1–14, 2020.

[26] K. Burghardt, W. Rand, and M. Girvan, "Competing opinions and stubborness: connecting models to data," *Physical Review E*, vol. 93, no. 3, Article ID 032305, 2016.

[27] A. Jędrzejewski and K. Sznajd-Weron, "Impact of memory on opinion dynamics," *Physica A: Statistical Mechanics and its Applications*, vol. 505, pp. 306–315, 2018.

[28] W. Su, G. Chen, and Y. Hong, "Noise leads to quasi-consensus of Hegselmann-Krause opinion dynamics," *Automatica*, vol. 85, pp. 448–454, 2017.

[29] F. Amblard and G. Deffuant, "The role of network topology on extremism propagation with the relative agreement opinion dynamics," *Physica A: Statistical Mechanics and its Applications*, vol. 343, pp. 725–738, 2004.

[30] J. Ghaderi and R. Srikant, "Opinion dynamics in social networks with stubborn agents: equilibrium and convergence rate," *Automatica*, vol. 50, no. 12, pp. 3209–3215, 2014.

[31] P. Walla, G. Brenner, and M. Koller, "Objective measures of emotion related to brand attitude: a new way to quantify emotion-related aspects relevant to marketing," *PLoS One*, vol. 6, no. 11, Article ID e26782, 2011.

[32] M. Del Vicario, G. Vivaldo, A. Bessi et al., "Echo chambers: emotional contagion and group polarization on facebook," *Scientific Reports*, vol. 6, no. 1, pp. 1–12, 2016.

[33] Y. Kim and Y. Kim, "Incivility on facebook and political polarization: the mediating role of seeking further comments and negative emotion," *Computers in Human Behavior*, vol. 99, pp. 219–227, 2019.

[34] G. Deffuant, F. Amblard, G. Weisbuch, and T. Faure, "How can extremism prevail? a study based on the relative agreement interaction model," *Journal of Artificial Societies and Social Simulation*, vol. 5, no. 4, 2002.

[35] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch, "Mixing beliefs among interacting agents," *Advances in Complex Systems*, vol. 3, no. 01n04, pp. 87–98, 2000.

[36] G. Deffuant, F. Amblard, and G. Weisbuch, "Modelling group opinion shift to extreme: the smooth bounded confidence model," 2004, https://arxiv.org/abs/0410199.

[37] M. J. Brandt, A. M. Evans, and J. T. Crawford, "The unthinking or confident extremist? political extremists are more likely than moderates to reject experimenter-generated anchors," *Psychological Science*, vol. 26, no. 2, pp. 189–202, 2015.

[38] G. Weisbuch, G. Deffuant, F. Amblard, and J.-P. Nadal, "Interacting agents and continuous opinions dynamics," in *Heterogenous Agents, Interactions and Economic Performance*, pp. 225–242, Springer, Berlin, Germany, 2003.

Hindawi
Complexity
Volume 2021, Article ID 3159968, 15 pages
https://doi.org/10.1155/2021/3159968

Research Article

Semantic Recognition and Location of Cracks by Fusing Cracks Segmentation and Deep Learning

Qing An,[1] **Xijiang Chen**,[1,2] **Xiaoyan Du**,[2] **Jiewen Yang**,[2] **Shusen Wu**,[3] **and Ya Ban**[4]

[1]*School of Artificial Intelligence, Wuchang University of Technology, Wuhan, Hubei 430223, China*
[2]*School of Safety and Emergency Management, Wuhan University of Technology, Wuhan, Hubei 430079, China*
[3]*State Key Laboratory of Materials Processing and Die and Mould Technology, School of Materials Science and Engineering, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China*
[4]*Chongqing Measurement Quality Examination Research Institute, Chongqing 404100, China*

Correspondence should be addressed to Xijiang Chen; cxj_0421@163.com and Shusen Wu; ssw636@hust.edu.cn

Received 24 June 2021; Revised 19 July 2021; Accepted 2 August 2021; Published 9 August 2021

Academic Editor: Ning Cai

For a long time, cracks can appear on the surface of concrete, resulting in a number of safety problems. Traditional manual detection methods not only cost money and time but also cannot guarantee high accuracy. Therefore, a recognition method based on the combination of convolutional neural network and cluster segmentation is proposed. The proposed method realizes the accurate identification of concrete surface crack image under complex background and improves the efficiency of concrete surface crack identification. The research results show that the proposed method not only classifies crack and noncrack efficiently but also identifies cracks in complex backgrounds. The proposed method has high accuracy in crack recognition, which is at least 97.3% and even up to 98.6%.

1. Introduction

At present, the use of Chinese concrete structure is fairly wide, involving many engineering industries. But with various cracks generated in the process of construction and using, potential safety hazard has been brought to engineering construction and maintenance. Now, the most common detection method is still manual detection, which is not only inefficient, but also brings safety problems. And the more advanced nondestructive detection methods are costly and cannot realize absolute noncontact detection. In recent years, with the continuous development of computer vision technology, the realization of high-precision crack detection based on computer technology has become a hot spot. The effects of detection methods are mainly focused on the sample-data cluster formation protocol to tackle the leaderless cluster formation problems [1]. They are mainly divided into traditional image processing methods (including edge detection [2], threshold segmentation [3], and region segmentation [4]) and machine learning methods (including deep learning [5] and clustering segmentation [6]).

For the traditional image processing methods, Xu et al. [7] used phase angle and gray distribution to improve the Canny operator, which can identify fine surface cracks, but generate more noise. Rivera et al. [8] tried to detect the cracks on the concrete surface by morphological methods and crack splicing, which can remove noise well, but the splicing effect of the crack's overall skeleton is not very good. Su and Yang [9] proposed an algorithm of image segmentation enhancement, named morphological segmentation based on edge detection-II (MSED-II), to concrete crack segmentation. Lu et al. [10] used a new double-threshold algorithm combined with morphological denoising to realize the identification of strain hardening cementitious composites crack images, but this method cannot identify cracks with complex backgrounds. Based on the edge detection algorithm, Nikravesh and Nezamivand Chegini [11] proposed a detection method using wavelet transform,

which can identify crack images with insignificant grayscale differences better and can identify cracks under a certain complex background, but the effect is not significant enough.

Compared with traditional image processing methods, Rao et al. [12] proposed a faster and simpler single-level convolutional neural network based on real-time target detection technology, which can not only reduce the impact of background on crack identification to a certain extent, but also identify various surface cracks of concrete bridges, but the accuracy is only 66%. Based on deep fully convolutional neural network, Ren et al. [13] proposed a multiscale crack feature extraction network structure for concrete tunnels named CrackSegNet, which has a strong ability to split the entire crack, but it still has trouble in crack extraction with complex background. Dung and Anh [14] introduced "Full Convolutional Neural Network" to realize effective detection of crack targets in a complex background and reduce false marks, but the processing efficiency is still not high enough. Based on the improved and optimized convolutional neural network "GoogLeNet," Ye et al. [15] have realized the high-precision identification of multiple cracks with complex background, but the identification accuracy of the crack data set is not very high.

For the accurate identification of concrete surface cracks, we proposed a method of concrete surface crack identification in this paper. It is based on the combination of convolutional neural network and clustering segmentation algorithm and can achieve accurate identification and size calculation of concrete surface crack images.

2. Preliminary Identification of Cracks

2.1. Crack Identification Network (CIN). According to the diversity of crack shapes and the uncertainty of crack magnitude, this paper proposes a convolution neural network structure called crack identification network (CIN) based on the convolutional neural network, as shown in Figure 1. Set serial number as i, and layer i as L_i; convolution operation i as C_i; and pooling operation i as P_i. L_1, L_3, L_4, L_6, L_8, and L_9 are convolutional layers; L_2, L_5, L_7, and L_{10} are pooling layers; L_{11} is fully connected layer; L_{12} is ReLu layer; L_{13} is dropout layer; and L_{14} is softmax layer.

The basic structure of the model consists of six convolutional layers, four pooling layers, a fully connected layer, an input layer, and an output layer. All convolutional layers and fully connected layers are batch normalized to improve the generalization ability of the model. In order to ensure the images which may be too large or too small not to affect the recognition effect, set the input images size as $256 \times 256 \times 3$. The three dimensions are length, width, and RGB components (red, green, and blue components). In order to perform feature compression better and simplify network complexity, set pooled window size of L_2, L_5, L_7, and L_{10} as 2×2, 4×4, 4×4, and 6×6 and step size as 2 and apply the maximum pooling rule. In order to obtain deeper image features and ensure that the complexity of the network would not increase, set convolution kernel size of convolutional layers L_1, L_3, L_6, and L_8 as 3×3 and convolution kernel size of convolutional layers L_4 and L_9 as 5×5; step size of all convolutional layers is 1. After convolution and pooling, through the Flatten operation, a vector containing 512 elements is output to the ReLu layer. Then, the partial neuron connection is randomly cut off by the dropout layer, and the dependence between network model parameters is reduced, the robustness of the model is improved, and overfitting is effectively prevented. Finally, the softmax layer is used to judge whether the data after the above processing is a crack image.

According to the constructed crack identification network, the cracks can be preliminarily identified through the following steps. First, we established a large number of crack data sets, and the data set was divided into a training set and a verification set. Next, we used the training set and validation set to train and validate the crack identification network model, and last, we used test set to test the model to predict the effect of the model prediction. The overall identification process is shown in Figure 2.

2.2. Data Set Making. The crack acquisition mechanism shown in Figure 3 was constructed to obtain the crack image to make a data set. We used the CCD industrial camera (Basler acA1300–30 gm) mounted on the drone to collect 1000 original images, most of which were crack pictures with complex backgrounds (such as crack pictures with rough surfaces, moss, bumps, depressions, and stains), and each image was different due to different photographing conditions. The size of the collected original images was 4896×3672 pixels. They are too large to input directly, so they are divided into 512×512 subimages in steps of 248, and 1000 original images are divided into 46000 subimages. In order to detect cracks and eliminate noise more effectively, some subimages were rotated. The final data set contained 58,000 subimages, 48,000 of which were used for training and 10,000 for testing and verifying the effectiveness of the network.

In order to train out a crack detection model with high robustness and high accuracy, the generated data set contains a wide range of crack image types. For example, the crack images contain cross cracks, thick and thin cracks, and cracks with debris. The noncrack images contain concrete surfaces with shadows, bumps, and textures, as shown in Figure 4.

It can be seen from Figure 4 that the blue dotted box in the first image of Figure 4(a) marks the intersection of the cross crack, the box in the second image marks the fallen leaves in the crack image, and the last two images are examples of thick and fine cracks. In Figure 4(b), the blue dotted elliptical frame in the first noncracked image marks the shaded part of the concrete surface. The elliptical frame in the second image marks the depression on the concrete surface. The elliptical frame in the third image marks the bumps on the concrete surface, and the oval frame in the last image marks the texture of the concrete surface.

In order to verify the recognition accuracy of the recognition model effectively and ensure the simplicity of the detection method, we used the 5-fold cross-validation [16] to verify and train 5 sets of models to obtain the training results of the 5 sets of models. Then, we selected a training result

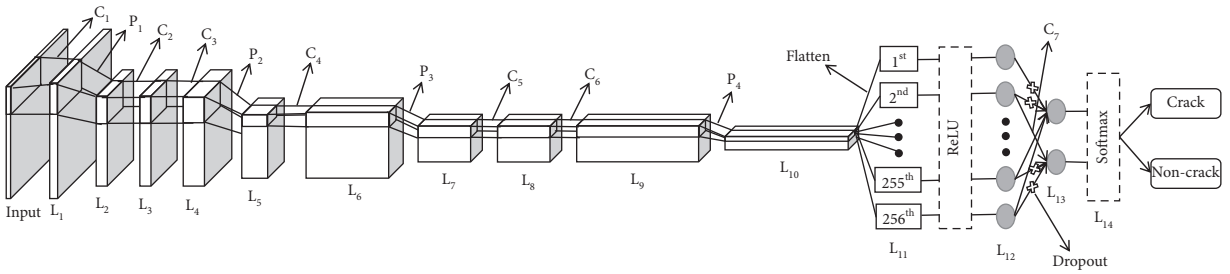


Figure 1: Schematic of crack identification network (CIN).

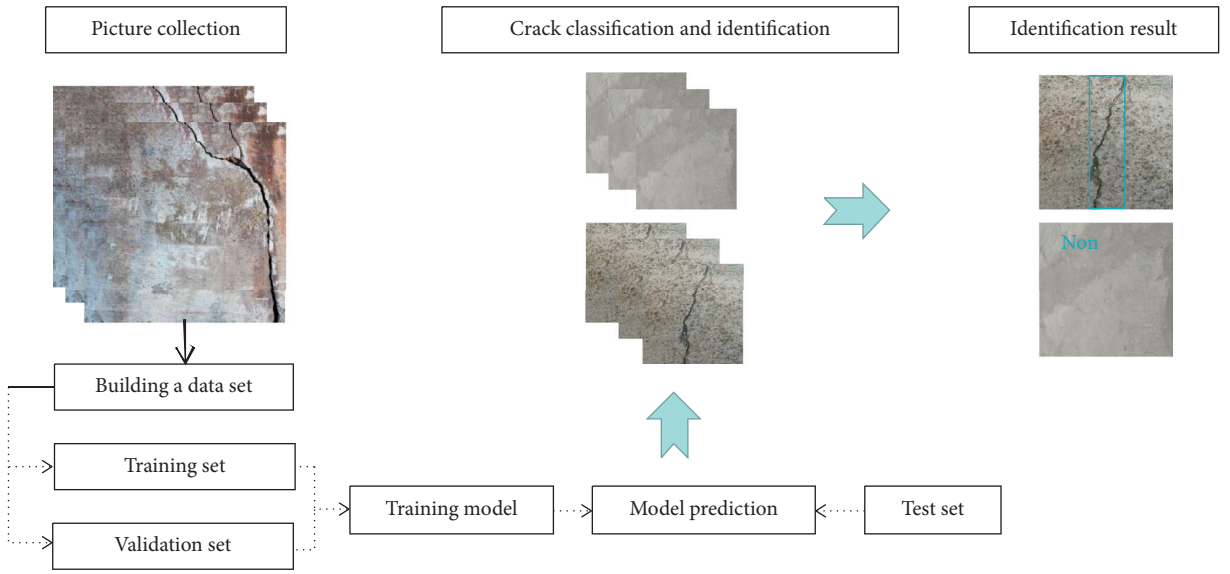
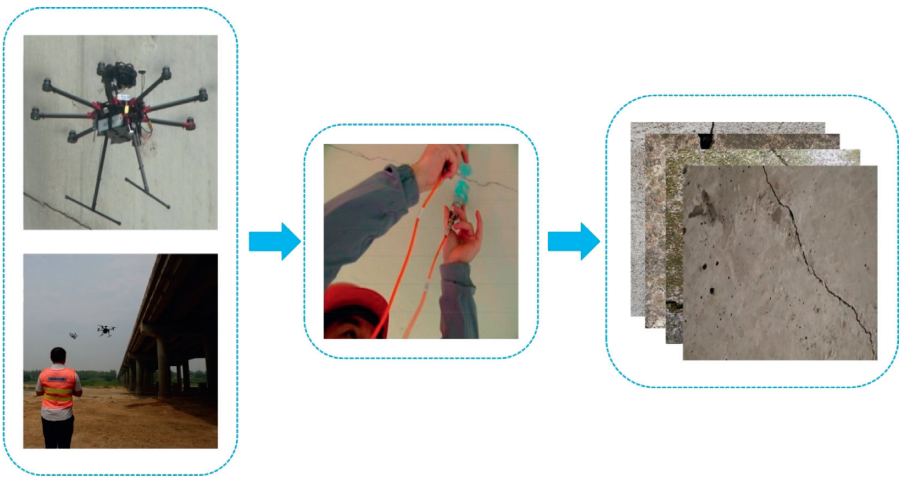


Figure 2: Flowchart of identification.

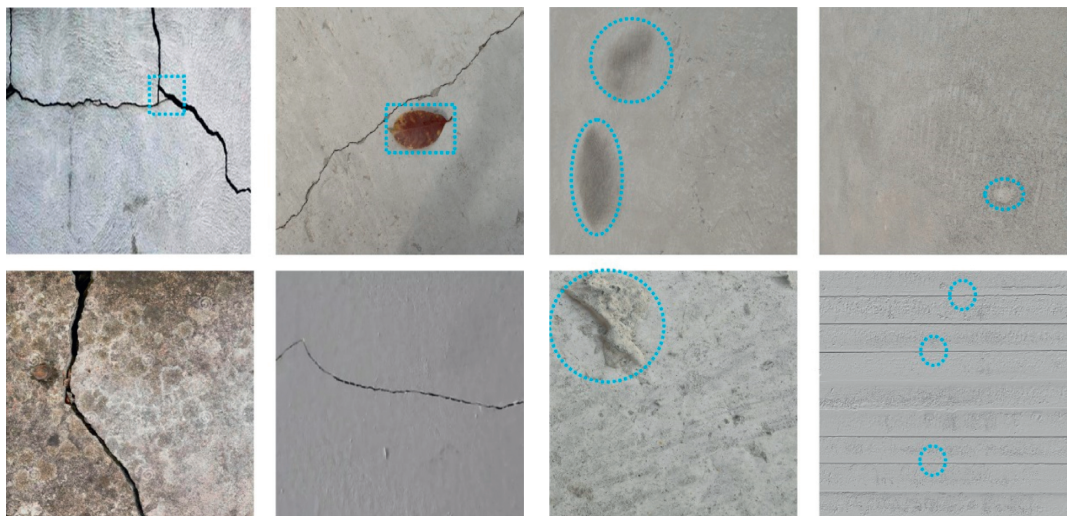


Photographing with UAV Measuring with Instrument Making Crack Data set

Figure 3: Crack collection mechanism.

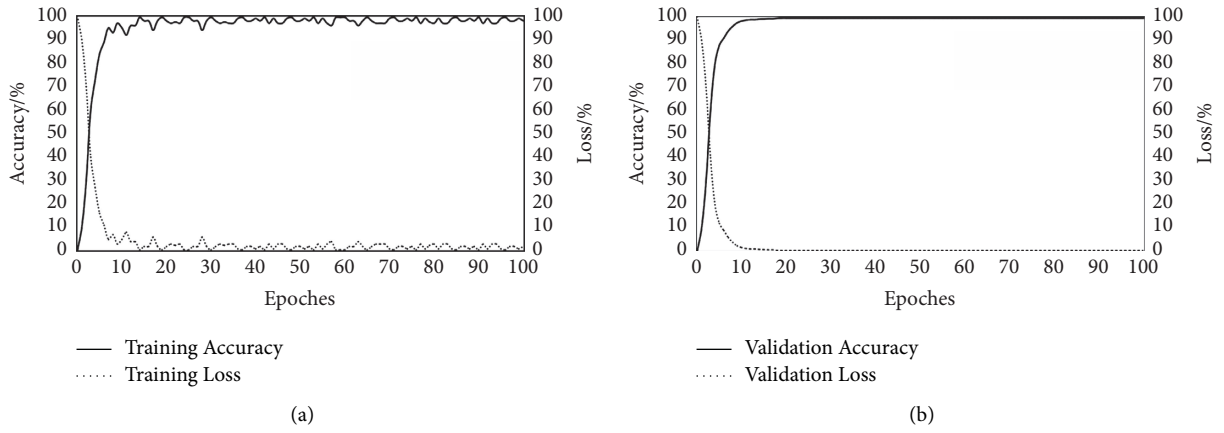
randomly to show the accuracy and loss rate curves of the training set and the verification set, which is shown in Figure 5.

It can be seen from Figure 5(a) that the model convergence speed is very fast. In the case with a small amount of sample data, the loss curve drops fast. The curve tends to be stable at 20 epochs, and it indicates that the sample training is almost completed. It can be seen from Figure 5(b) that the accuracy of the verification set continues to rise, and the loss value continues to decrease. The accuracy of the



(a) (b)

Figure 4: Example of crack and noncrack images. (a) Crack images. (b) Noncrack images.



(a) (b)

Figure 5: Training and validation results. (a) Training results. (b) Validation results.

verification set tends to 100% at 20 epochs; the loss value tends to 0%, and it indicates that training has finished at 20 epochs.

Figure 6 shows the results of the 5-fold cross-validation. The accuracy rates of the five sets of model training sets are all above 99%, and the average accuracy rate of the validation set is 99.1%. The accuracy of the validation set of the five groups of models has a little change, but the range of change is small, which shows that the trained crack network model has good robustness and stability.

Some images were selected from the data set randomly to test the effect of crack identification of the crack identification network, as shown in Figure 7.

From Figure 7(b), it is clearly visible that the crack identification network based on deep learning can classify cracks and noncracks accurately. On the one hand, as shown in Figure 7(b), in which the first and second layers belong to the classified cracks Images, while the third and fourth layers belong to the classified noncrack images. On the other hand, the cracks can be identified and located cracks accurately, as shown in the blue box in Figure 7(b). It can be seen from Figure 7(a) that the resolution of the original image is different. For example, the resolution of the first image is significantly greater than that of the second. However, the first and second images in Figure 7(b) can identify the cracks accurately, indicating that the method is not affected by the resolution. At the same time, there are other miscellanies around the cracks, such as the leaves, as shown in Figure 7(a). The blue boxes in the second and seventh pictures of Figure 7(b) show the cracks around the leaves clearly. It indicates that the method is not affected by the surrounding miscellanies. As shown in the first and seventh pictures of Figure 7(b), the cracks in the first picture of Figure 7(a) are significantly larger than those in the seventh picture of Figure 7(a), but the identification effect is basically the same. The fourth picture in Figure 7(a) shows cross

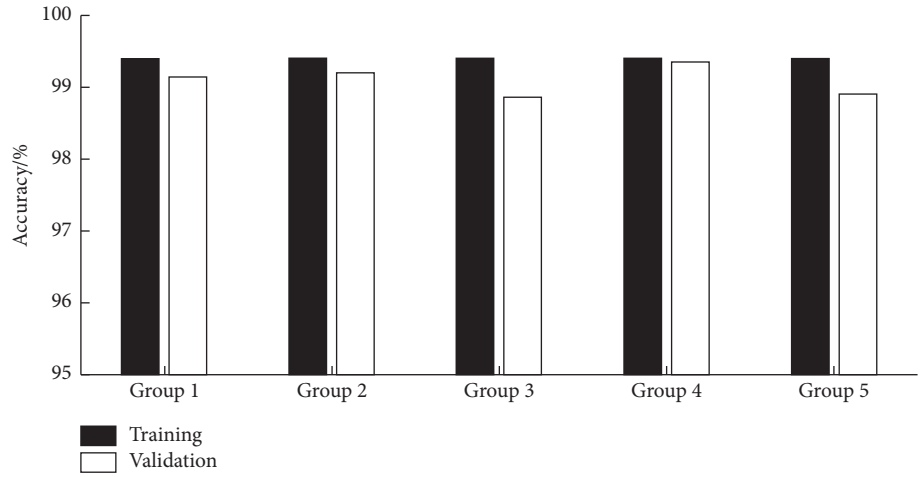
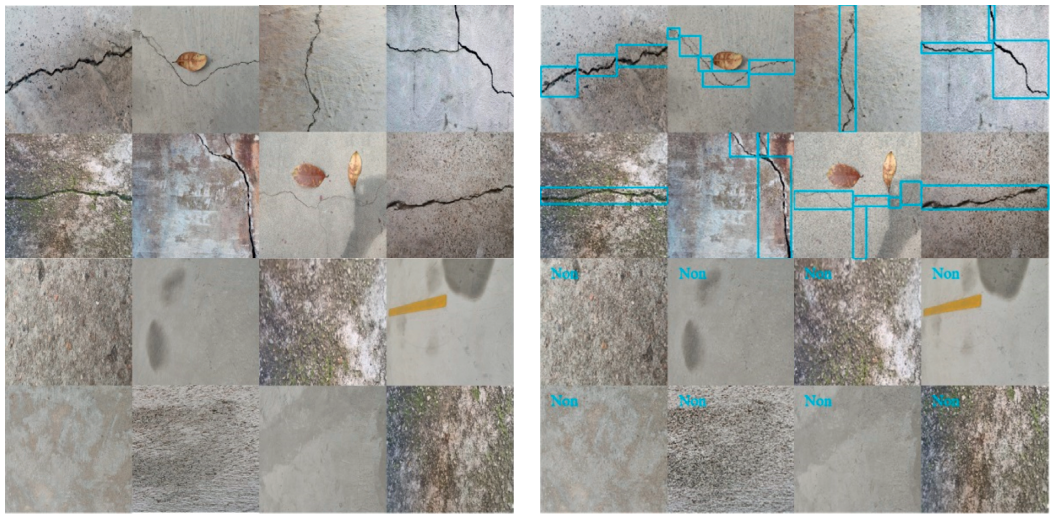


Figure 6: Accuracy rate of training and validation for 5 different groups.



(a) (b)

Figure 7: Example of identification results. (a) Original image. (b) Crack classification and identification results.

cracks, but the fourth picture in Figure 7(b) gives three cross blue box identification results, indicating that the method can automatically decompose the cross crack identification result into straight crack identification result. The noncrack image in the blue dotted oval frame in Figure 7(a) contains large convex and concave concrete surfaces, and the solid blue line in the oval frame shows the concrete surface with large texture. The identification result shown in Figure 7(b) is not affected by those, indicating that the method is not affected by the surface characteristics of the crack. From the above analysis, it can be seen that the crack identification network we proposed is not affected by the image resolution, the surrounding miscellanies, nor the surface characteristics of the crack.

Although the crack identification network accurately positions the determined crack images, the positioning results lack the identification of complete cracks and continuity and cannot eliminate the influence of background miscellanies on the cracks. If we only clip and quantitative calculate according to above results, a large error will occur, so it is necessary to divide the cracks more accurately to achieve quantitative identification of cracks.

3. Accurate Crack Segmentation

After the crack identification in the previous section, the initial crack image was obtained, but the crack in the crack image was not segmented accurately. For this, we achieve accurate segmentation of cracks based on improved *K*-means clustering algorithm.

3.1. Improved K-Means Clustering Segmentation Algorithm. Although the *K*-means clustering algorithm [17] has fast and efficient segmentation characteristics, it has high requirement for high clustering center selection and is very easy to converge to the local optimal solution, thereby missing the global optimal solution. In view of this, we used dynamic particle swarm optimization (DPSO) to improve the

K-means clustering algorithm. The specific steps of improved K-means are as follows:

Step 1. Convert the color image to grayscale image to obtain the initial grayscale image of the crack, and then set the initial cluster number k.

Step 2. Determine the clustering center m according to the K-means algorithm.

Step 3. After the clustering center is determined, calculate the fitness value f_i of the particles i [18], which is shown in the following equation:

$$f_i = \sum_{i=1}^{n} \sum_{j=1}^{k} \left\| x_i - m_j \right\|^2. \tag{1}$$

Here, x_i is the i th data point in the data point set X and m_j is the j th cluster center.

Step 4. Determine the particle inertia coefficient w and learning factor s according to equations (2) and (3) [18]:

$$w = \begin{cases} w_{\min} - \dfrac{(w_{\max} - w_{\min}) \times (f_i - f_{\min})}{f_{\text{avg}} - f_{\min}}, & f_i \le f_{\text{ave}}, \\ w_{\max}, & f_i > f_{\text{ave}}. \end{cases} \tag{2}$$

Here, $w_{\max}$ is the maximum inertia coefficient; $w_{\min}$ is the minimum inertia coefficient; f_i is the current adaptation value of particles; f_{ave} is the current average adaptation value of all particles; and $f_{\min}$ is the minimum adaptation value of all particles.

$$\begin{cases} s_1 = s_{1,\text{int}} + \dfrac{s_{1,\text{fin}} - s_{1,\text{int}}}{t_{\max}} \times t, \\ s_2 = s_{2,\text{int}} + \dfrac{s_{2,\text{fin}} - s_{2,\text{int}}}{t_{\max}} \times t. \end{cases} \tag{3}$$

Here, s_1 and s_2 represent the particle's self-learning ability and the learning ability to the excellent collective; $s_{1,\text{int}}$ and $s_{2,\text{int}}$ represent the particle's initial learning ability; $s_{1,\text{fin}}$ and $s_{1,\text{fin}}$ represent the particle's final learning ability; $s_{1,\text{int}} > s_{2,\text{int}}$; $s_{1,\text{fin}} > s_{2,\text{fin}}$; and $t_{\max}$ represents the maximum time for algorithm operation.

Step 5. Use the inertia coefficient w and learning factor s obtained in step 4 to update the particle velocity V_i [18] and position X_i [18], which are shown in the following equation:

$$\begin{cases} x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1), \quad j = 1, 2, 3, \ldots, d, \\ v_{ij}(t+1) = w v_{ij}(t) + s_1 r_1 \left[p_{ij} - x_{ij}(t) \right] + s_2 r_2 \left[p_{gj} - x_{ij}(t) \right]. \end{cases} \tag{4}$$

Here, r_1 and r_2 are two random numbers which are evenly distributed on the interval $[0, 1]$; t is time; p_{ij} is the position of particle i during the update process; and p_{gi} is the best position that particle i experienced during the update process.

Step 6. When the maximum number of iterations is reached or the particle swarm fitness variance a converges to δ^2 fixed value, the global optimal solution is output to obtain the particle swarm fitness variance δ^2 [18], which is shown in the following equation; otherwise, repeat step two to step four.

$$\delta^2 = \frac{1}{n} \sum_{i=1}^{n} (f_i - f_{\text{ave}})^2. \tag{5}$$

Step 7. Take the optimal solution of step 5 as the optimal clustering center, and then obtain k clusters.

Step 8. Perform criterion function judgment. The index function of the Davidson Burger Index (DBI) [19] is adopted as the criterion function of the final clustering result, and its calculation formula is shown as follows:

$$\text{DBI} = \frac{1}{k} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \max \left\{ \frac{S_i + S_j}{\text{Dis}(i, j)} \right\}. \tag{6}$$

Here, $\text{Dis}(i, j)$ represents the distance between the i th cluster and the j th cluster; S_i represents the sum of the standard error of the Euclidean distance between each data point in the i th cluster and the center m_i of the cluster; and k is the total number of clusters in the data set.

Step 9. If DBI converges, the clustering result at this time is output as the final result; if it does not converge, repeat steps 2 to 8 until it converges, and then output the result. The overall flowchart of the improved K-means algorithm is shown in Figure 8.

3.2. Algorithm Performance Analysis. We used the evaluation index of precision, recall, and F-measure to evaluate the effectiveness of the algorithm. The specific evaluation index [20] is defined as shown in Figure 9.

In Figure 9, P stands for precision rate; R stands for recall rate; F stands for F-measure value; A stands for the number of images that are crack images and correctly identified as crack; B stands for the number of images that are background images but identified as crack; C represents the number of images that are crack images but identified as background; and D represents the number of images that are background images and are correctly identified as background.

3.3. Comparison with Traditional Algorithms. Based on the crack data set in Section 2, taking a few simple crack images as an example, compare the segmentation effects of the improved K-means algorithm with the traditional improved Otsu [21], improved Canny [22] and improved median filtering [23]. Taking the presence and absence of rough surfaces and protrusions as examples, the comparison of segmentation effects is performed as shown in Figures 10 and 11, respectively.

It can be seen from Figure 10 that, for the first crack image with rough surface, although the improved Otsu

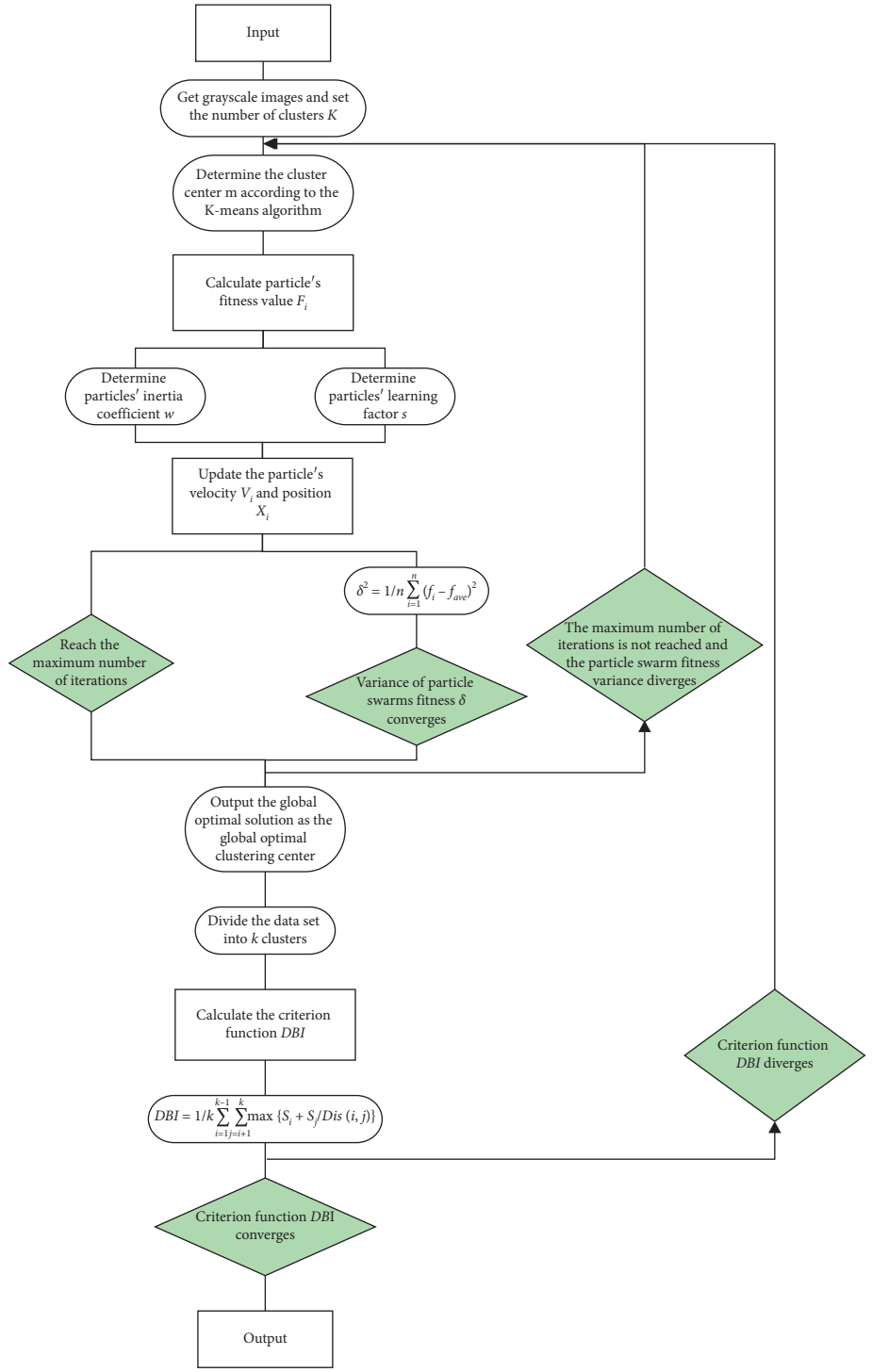


Figure 8: Overall flowchart of the improved *K*-means algorithm.

algorithm, Canny algorithm, and median filter algorithm can all segment the crack, but there is still some noise, but the improved *K*-means algorithm has almost no noise, and the contour of the crack is also clearer. For the second crack image, because the surface is smooth, all algorithms have few noise, but the crack skeleton integrity and continuity of the improved *K*-means algorithm are higher than the traditional algorithms. It can be seen from the above comparison that the precise segmentation of the improved *K*-means algorithm has good noise resistance and accuracy.

As can be seen from Figure 11, for the first image with raised cracks, although the improved Otsu algorithm, Canny algorithm, and median filter algorithm can all segment the cracks, the cracked skeletons are all thick, and the bulge is also segmented out. In addition to almost no noise, the improved *K*-means algorithm can also restore the outline of

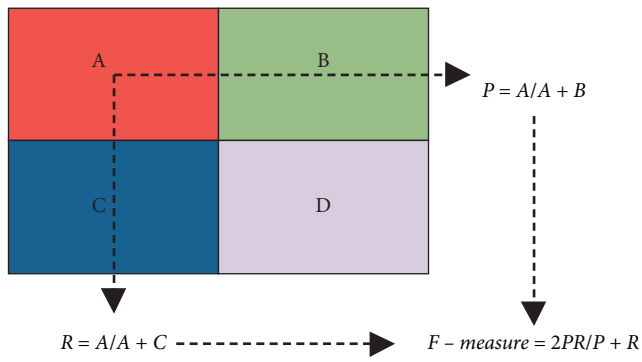


Figure 9: Evaluation index.

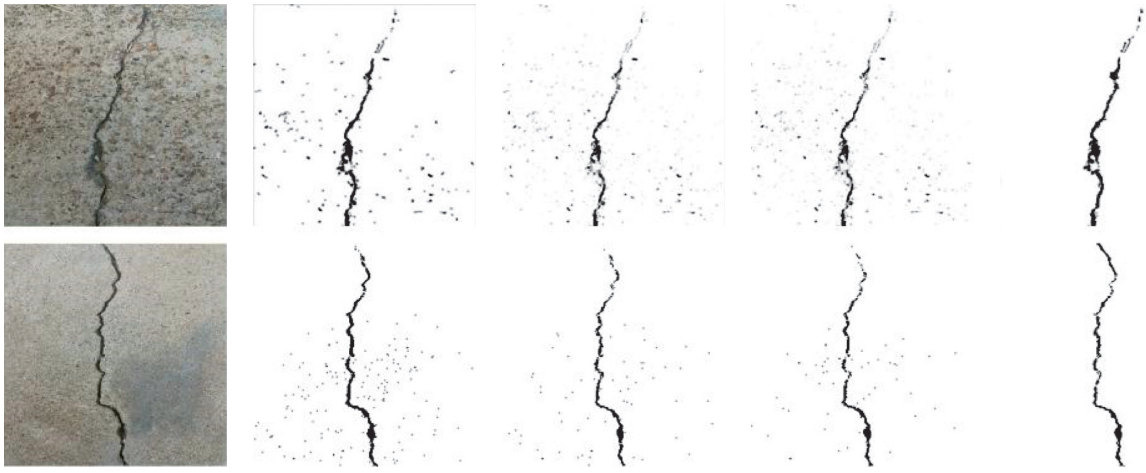


Figure 10: Comparison of rough surface and smooth surface. (a) Original image; (b) segmentation of improved Otsu algorithm; (c) segmentation of improved Canny algorithm; (d) segmentation of improved median filter; and (e) segmentation of this paper's algorithm.

the thin original crack skeleton more accurately. For the second crack image without bumps, because the surface has no bumps, the traditional algorithm only has a small amount of noise, but the crack continuity is poor. But the improved *K*-means algorithm does not have noise, while ensuring integrity and continuity of the crack skeleton well. It can be seen from the above comparison that the precise segmentation of the improved *K*-means algorithm has good noise resistance and accuracy.

Then, combine the algorithms with the crack identification network, and compare and analyze the overall identification effect. The result of evaluation indicators is shown in Figure 12.

As can be seen from Figure 12, based on the data set of this article, the indicators of the improved Otsu algorithm and the improved Canny algorithm are less than 80%, while the improved median filter algorithm only has a precision rate slightly higher than 80%, and the other two indicators are also no more than 80%; and the improved *K*-means algorithm has a recall rate and *F*-measure value of more than 90%, and the accuracy rate has reached 97%. It can be seen from the above that the processing effect of the improved *K*-means algorithm is much higher than the above traditional algorithms.

3.4. Comparison with Clustering Algorithms. Based on the crack data set, the improved *K*-means segmentation algorithm is used to identify crack pictures on the surface of the concrete with moss, fallen leaves, or water stains, and the identification results are combined with the *K*-means algorithm [24], means shift algorithm [25], and fuzzy *C*-means algorithm for comparative analysis, as shown in Figure 13.

It can be seen from Figure 13 that when identifying the first and second cracks with moss in Figure 13(a), the improved *K*-means algorithm removes almost all noise and guarantees the integrity of the crack skeleton well. For the third cracked image with fallen leaves in Figure 13(a) (where the fallen leaves are marked with a blue dotted box), the other three algorithms have failed to eliminate the noise from the fallen leaves. The two cracks connected to the second small segment were clearly identified. For the identification of the fourth crack with fallen leaves and shadows in Figure 13(a) (where the fallen leaves are marked with a blue dotted box and the shadow is marked with a blue solid line), the *K*-means algorithm, mean shift algorithm, and fuzzy *C*-means algorithm can hardly identify cracks because the cracks are too fine, and neither can remove the noise caused by the fallen leaves and shadows. However, the improved *K*-means algorithm identified the entire fine

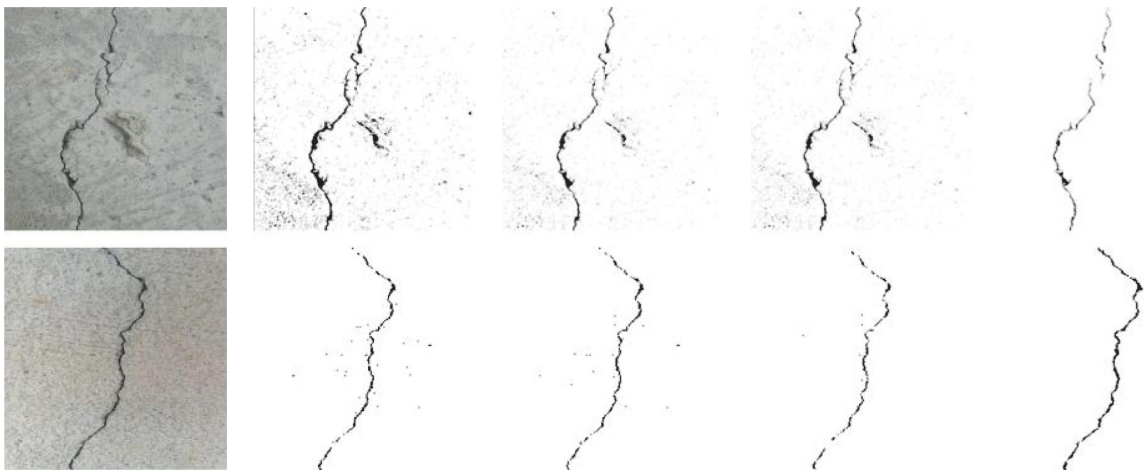


Figure 11: Comparison of salient surface and flat surface. (a) Original image; (b) segmentation of improved Otsu algorithm; (c) segmentation of improved Canny algorithm; (d) segmentation of improved median filter; and (e) segmentation of this paper's algorithm.

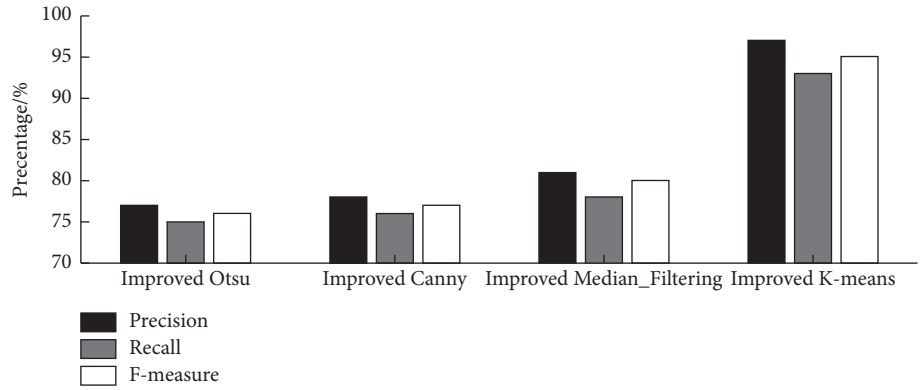


Figure 12: Comparison of evaluation indicators of each algorithm.

cracks well when the fallen leaves and shadow noises are completely removed. For the last picture with water stains in Figure 13(a) (where water stains are marked with solid blue boxes), the other three types of algorithms all have a lot of noise and the upper left part of the crack is missing, and the lower rightmost part of the crack is unable to be identified due to the influence of water stains and the grayscale difference of the concrete surface, but the improved *K*-means algorithm completes the clear identification of the overall crack without noise, and the small arc in the middle of the crack is also clearly identified. The above analysis shows that the crack identification method is not only not affected by moss, shadow, water stains, and miscellanies, but also not affected by the size of the crack and the surface characteristics. Even for such cracks under complex backgrounds, this paper's identification method can still identify and segment the crack skeleton with high accuracy, and it also guarantees the great advantage of no noise and the integrity and continuity of the crack skeleton. And the comparison of the evaluation indexes of the algorithms is shown in Figure 14.

It can be seen from Figure 14 that based on the data set, the recall and *F*-measure values of the improved *K*-means algorithm are both above 90%, and the accuracy rate is 97%, which is still much higher than other three clustering algorithms. It can be inferred that the performance of the improved *K*-means algorithm is better than the other clustering algorithms.

4. Quantitative Identification of Cracks

Use deep learning and improved *K*-means algorithm to accurately extract the crack but cannot determine the size of the crack. For this, this section calculates the physical size of the crack accurately according to the premise determination of the crack pixel size.

4.1. Crack Physical Size. The physical size of the cracks, namely, the area, length, average width, and occupation ratio, needs to be calculated. The calculation method is as follows:

(1) *Crack Area*. Calculate the number of pixels included in each connected domain, according to the actual physical size corresponding to each pixel, as shown

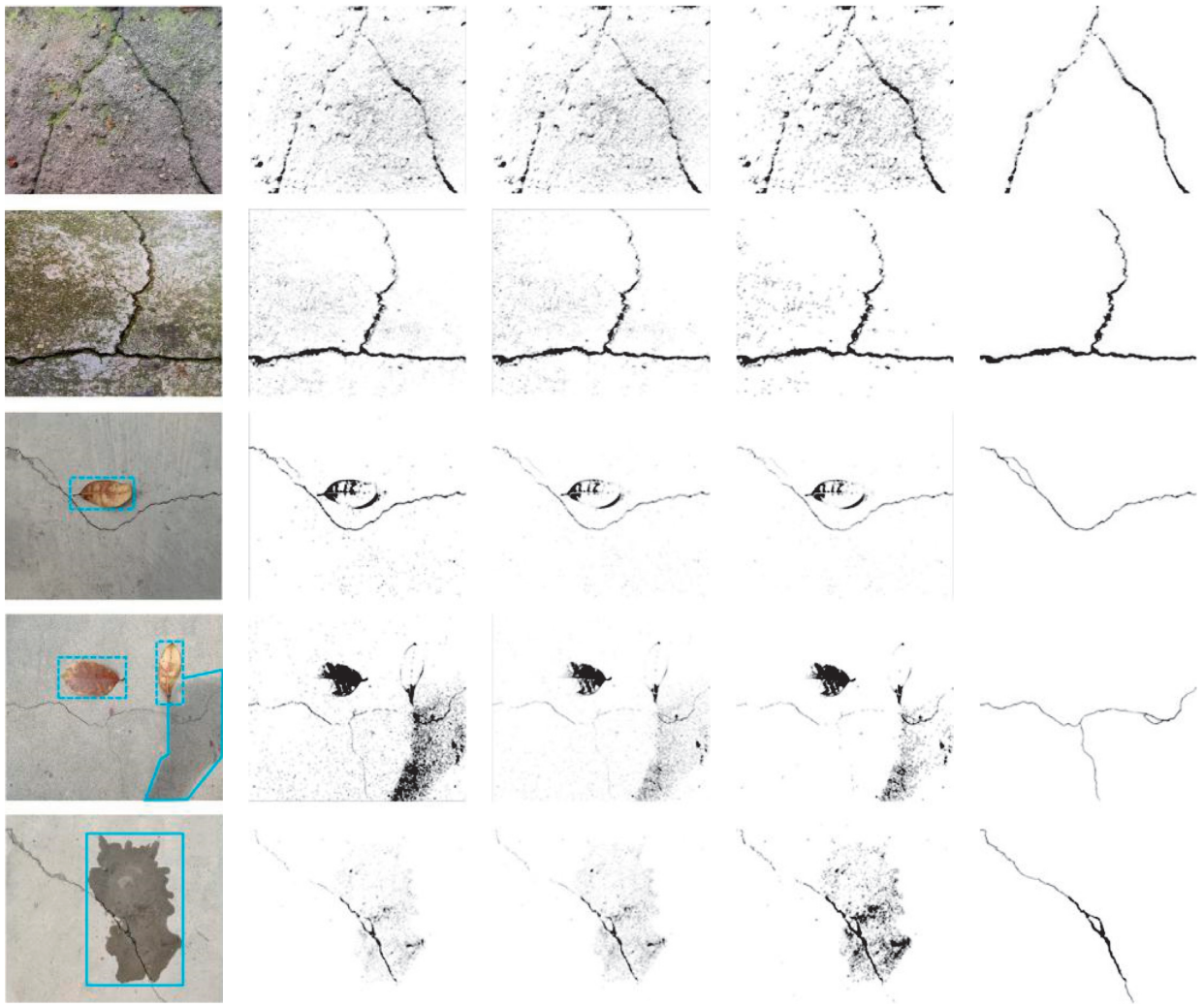


Figure 13: Comparison with traditional algorithms. (a) Original image; (b) segmentation of improved Otsu algorithm; (c) segmentation of improved canny algorithm; (d) segmentation of improved median filter; and (e) segmentation of this paper's algorithm.

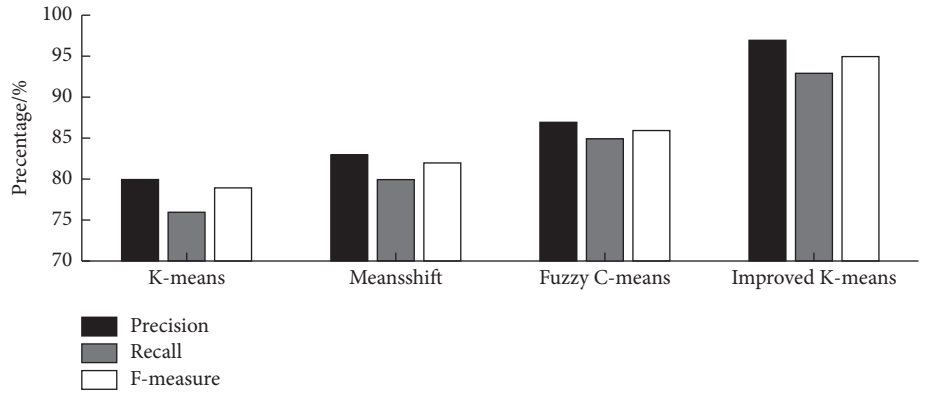


Figure 14: Comparison of evaluation indicators of each algorithm.

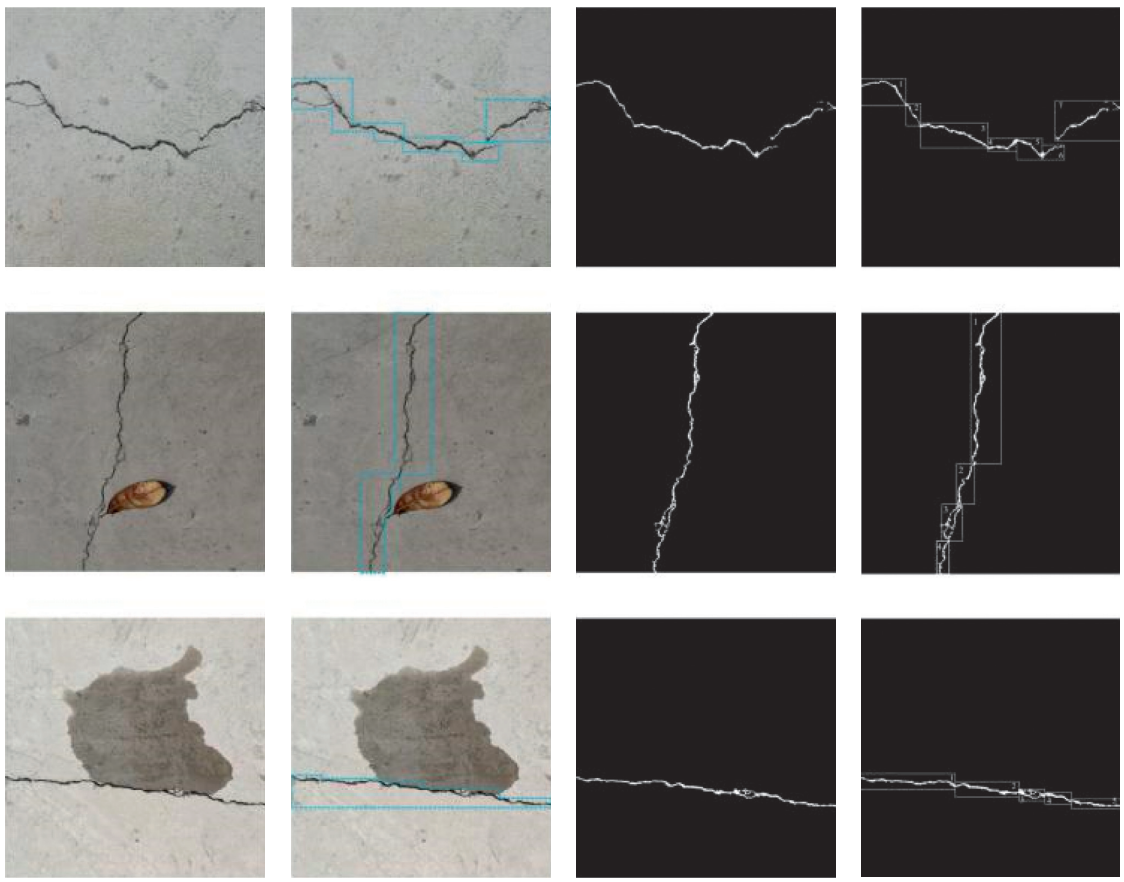


Figure 15: Crack marking. (a) Original image; (b) identification of CIN; (c) segmentation; (d) marking.

in the following equation. Thus, the area of each connected domain is obtained, and the areas of all connected domains are summed.

$$P_a = \alpha \cdot P_i. \tag{7}$$

Here, P_a is the actual physical size of the crack, the unit is mm^2; P_i is the pixel size of the crack, the unit is $pixel^2$; and α is the zoom ratio of the crack, the unit is $mm^2/pixel^2$.

(2) *Crack Length*. Thin the image of the crack skeleton. The crack width after thinning is a single pixel; then calculate half of the perimeter to get the number of pixels of the crack at this time, and get the crack physical length according to formula (7).

(3) *Average Crack Width*. The average crack width is the ratio of the crack area to the length. The calculation formula of the average crack width can be obtained by combining the calculation method of the crack length, which is shown as follows:

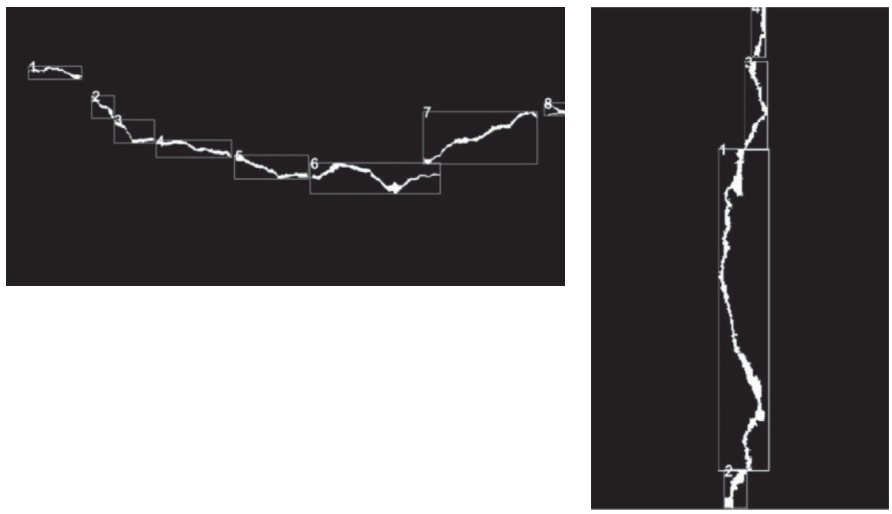
$$d_i = \frac{2S_i}{C_i}, \quad 1 \le i \le n. \tag{8}$$

Here, S_i is the area of the i th crack; C_i is the perimeter of the i th crack; d_i is the average width of the i th crack; and n is the total number of cracks.

(4) *Crack Occupancy Ratio*. It is the ratio of the area of the crack to the total area of the image.

4.2. Crack Marking. To facilitate the calculation of the size of the crack, the acquired crack image needs to be marked. Based on the data set, combined with the output of the CIN network model, the crack pictures were tested for marking through morphological method. The software used is MATLAB 2016b, the processor is Intel® Core™ i5-8300H CPU @2.30 GHz, the memory is 8.00 GB, the graphics card is NVIDIA GeForce GTX1050Ti, and the operating system is Windows 10.

In order to verify the identification performance better, the identification test was carried out by taking the crack shown in Figure 15 as an example. And the crack was marked in sections to facilitate the size calculation; at the same time, the grayscale image was reversed to improve the effect of marking. And, Figure 15(a) is the original image of the crack; the identification result of CIN is shown in



(a) (b)

Figure 16: Example of size calculation. (a) Crack 1. (b) Crack 2.

Table 1: Pixel size of cracks 1 and 2.

Crack number	Segmentation number	Width (pixel)	Length (pixel)	Average width (pixel)	Overall length (pixel)	Area (pixel^2)	Occupation ration (%)
Crack 1	1	4	107	4.63	1193	5994	0.76
	2	4	63				
	3	4	97				
	4	5	152				
	5	6	163				
	6	6	286				
	7	5	258				
	8	3	67				
Crack 2	1	5	913	4.75	1384	6213	1.47
	2	6	106				
	3	4	240				
	4	4	125				

Table 2: Actual size comparison results of cracks 1 and 2.

Crack number	Average width of quantitatively calculation (mm)	Overall length of quantitatively calculation (mm)	Area of quantitatively calculation (mm^2)	Average width of crack gauge measurement (mm)	Overall length of crack gauge measurement (mm)	Area of crack gauge measurement (mm^2)
Crack 1	0.97	250.53	264.33	1.00	251.20	261.52
Crack 2	1.00	290.64	273.99	0.98	288.42	270.26

Figure 15(b); the segmentation of the identified crack is shown in Figure 15(c); and the marking of the split cracks is shown in Figure 15(d).

It can be seen from Figure 15 that the marking of the crack skeleton is very clear and unambiguous due to the precise identification and segmentation, and almost every large turning point is segmented to facilitate the subsequent size calculation.

4.3. Crack Size Calculation. The method of this paper was used to identify and segment cracks and mark them. Then, we chose two images as an example to perform the size calculation, which is shown in Figure 16. The crack size calculation method can achieve quantitative identification of cracks. The shooting equipment used in this article was CCD industrial camera (Basler acA 1300–30 gm), and the zoom ratio of the camera in image shooting is α 0.21 mm/pixel.

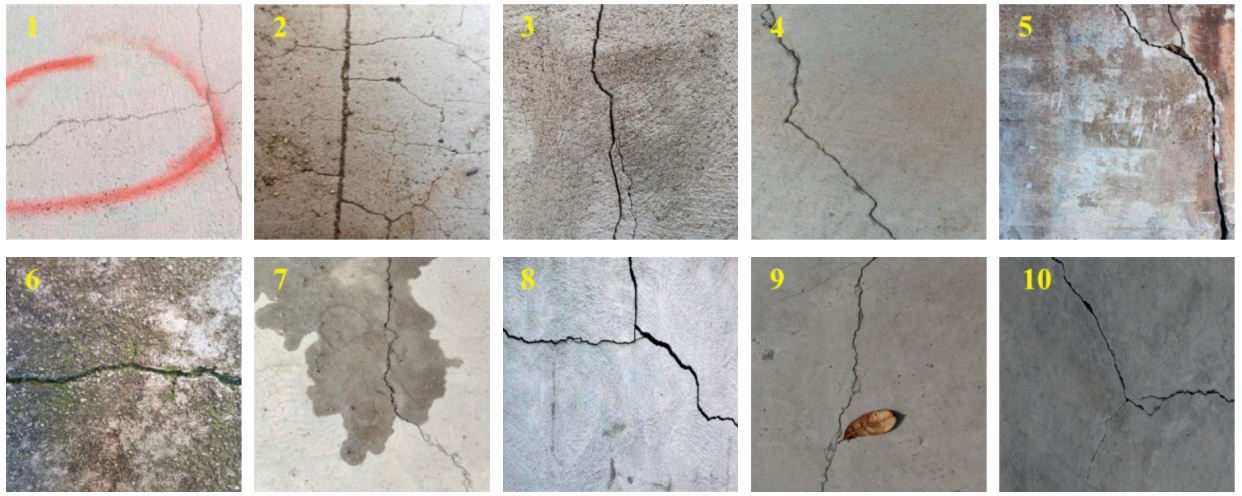


Figure 17: Original crack image for quantitative calculation.

Table 3: Comparison of statistical results.

Group number	Average width of quantitatively calculation (mm)	Average width of crack gauge measurement (mm)	Error (mm)	Overall length of quantitatively calculation (mm)	Overall length of crack gauge measurement (mm)	Error (mm)	Area of quantitatively calculation (mm^2)	Area of crack gauge measurement (mm^2)	Error (mm^2)
1	0.81	0.80	0.01	202.66	205.86	3.20	206.35	203.46	2.89
2	0.76	0.74	0.02	185.13	189.30	4.17	169.14	172.55	3.41
3	0.87	0.86	0.01	221.37	224.75	3.38	249.90	254.73	4.83
4	0.80	0.82	0.02	216.17	212.35	3.82	218.45	222.19	3.74
5	0.78	0.80	0.02	200.65	205.02	4.37	206.40	202.62	3.78
6	0.81	0.80	0.01	210.60	207.68	2.92	212.11	208.19	3.92
7	0.76	0.78	0.02	204.19	200.34	3.85	201.96	197.93	4.03
8	0.80	0.78	0.02	197.57	201.22	3.65	195.80	199.02	3.22
9	0.80	0.82	0.02	207.03	210.75	3.72	217.81	213.78	4.03
10	0.91	0.93	0.02	247.73	243.68	4.05	279.35	284.48	5.13

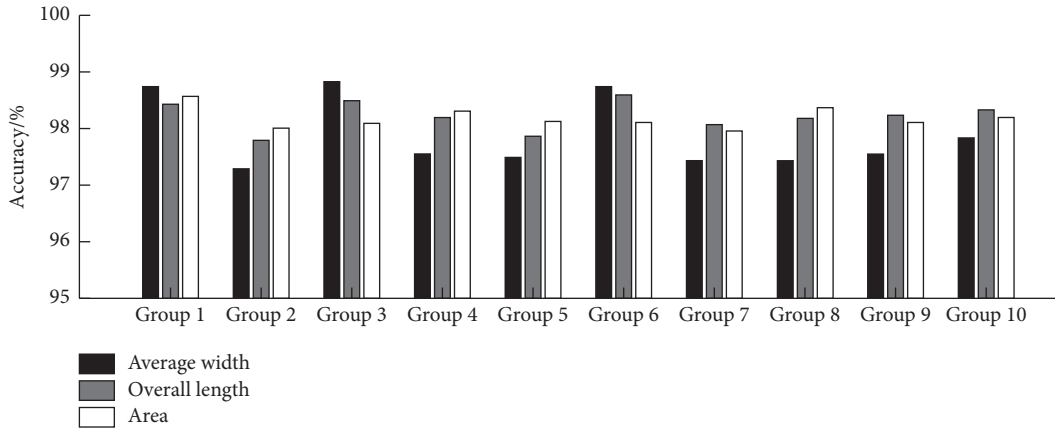


Figure 18: Comparison of accuracy of statistical results.

It can be seen from Figure 16 that segmented marking is adopted due to the different widths of each segment. Crack 1 in Figure 16(a) contains 8 segments, and crack 2 in Figure 16(b) contains 4 segments. Table 1 shows the pixel size and occupancy ratio of each segment of 1 and crack 2 as well as the overall pixel size. Combining Table 1 and the zoom ratio, the actual sizes of the cracks 1 and 2 were obtained and compared with the measurement results of the crack measuring instrument. The results are shown in Table 2.

It can be seen from Tables 1 and 2 that the error between the size obtained by the identification method and the size obtained by the crack measuring instrument is not very large. In order to verify the accuracy more effectively, a part of the image was randomly extracted from the data set. This part of the image was divided into 10 categories, and each category selected a representative image, as shown in Figure 17. The results of these 10 crack images were identified and quantified and compared with the average width, overall length, and area measured by the crack measuring instrument. The statistics and comparison results are shown in Table 3, and the accuracy chart is shown in Figure 18.

It can be seen from Table 3 and Figure 16 that for 10 representative crack images, the accuracy of the average width calculated by the quantitative identification method is not less than 97.3% compared with the average width measured by the crack measuring instrument, and the highest accuracy rate is up to 98.84%; the accuracy of the overall length is no less than 97.8%, and the highest accuracy is even more than 98.6%; the accuracy of the area is no less than 97.9%, and the highest accuracy is up to 98.5%. From the above comparative analysis, we can see that the quantitative identification method has good stability, and the accuracy is extremely high. Compared with the instrument measurement, the error obtained is basically within the controllable range.

5. Conclusion

This paper proposes a concrete surface crack identification and size calculation method combining deep learning convolutional neural network, clustering segmentation, and morphological methods, which can effectively classify concrete cracks and noncrack images and accurately segment and quantitatively identify cracks. The main conclusions are as follows:

(1) Using deep learning convolutional neural network for training, a preliminary crack recognition model is obtained. After verification and testing, the identification accuracy rate reached more than 99%. At the same time, the crack identification model of deep convolutional neural network is combined with clustering segmentation and morphological methods to propose a hybrid multialgorithm crack identification extraction method called CIN.

(2) Compared with traditional segmentation algorithms and similar clustering segmentation algorithms, the improved *K*-means algorithm can identify a variety of concrete surface cracks in complex backgrounds while ensuring extremely high segmentation accuracy.

(3) Based on the improved *K*-means algorithm, combined with morphological methods, the identification of cracks, quantitative identification, and calculation of physical size are realized. Experiments show that the accuracy of quantitative identification is at least 97.3%, and the highest can reach 98.6%.

(4) It provides a certain theoretical basis for ensuring the safety of crack detection personnel on the concrete surface, reducing workload, and maintaining and safety testing of concrete structures. It also provides some research bases for the identification of similar cracks under higher precision and more complex conditions.

Data Availability

All the data generated or analyzed during this study are included within this article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Chongqing Technological Innovation and Application Development (Program SN: cstc2019jscx-msxmX0051), CRSRI Open Research Program (Program SN: CKWV2019758/KY), the Quality and Technical Supervision Bureau Research Project of Chongqing Province (CQZJKY2018004), and Chongqing Administration for Market Regulation Project (CQSJKY2019004).

References

[1] X. Yang, L. Liao, Q. Yang, B. Sun, and J. Xi, "Limited-energy output formation for multiagent systems with intermittent interactions," *Journal of the Franklin Institute*, vol. 16, 2021.

[2] L. Kong, "SIFT feature-based video camera boundary detection algorithm," *Complexity*, vol. 2021, Article ID 5587873, 11 pages, 2021.

[3] Y. Dong, M. Li, and G. Cheng, "Evaluation and comparison of thresholding segmentation techniques," *In Proceedings of the Mechanics & Materials*, vol. 519-520, pp. 689–692, 2014.

[4] X. Zhang, P. Xiao, X. Feng, and G. He, "Another look on region merging procedure from seed region shift for high-resolution remote sensing image segmentation," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 2019, no. 148, pp. 197–207, 2019.

[5] F.-P. An, "Medical image classification algorithm based on weight initialization-sliding window fusion convolutional neural network," *Complexity*, vol. 2019, Article ID 9151670, 15 pages, 2019.

[6] X. Chen, Q. AN, K. Yu, and Y. Ban, "A novel fire identification algorithm based on improved color segmentation and enhanced feature data," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–15, 2021.

[7] H. Xu, Y. Tian, S. Lin et al., "Research of image segmentation algorithm applied to concrete bridge cracks," in *Proceedings of the International Conference on Information Science and Technology*, Yangzhou, China, February 2013.

[8] J. Rivera, G. Josipovic, E. Lejeune et al., "Automated detection and measurement of cracks in reinforced concrete components," *In Proceedings of the Computer Science*, vol. 112, no. 3, pp. 397–405, 2015.

[9] T. Su and M. Yang, "Morphological segmentation based on edge detection-II for automatic concrete crack measurement,"

In Proceedings of the Computer and Concrete, vol. 21, no. 6, pp. 727–739, 2018.

[10] C. Lu, J. Yu, and C. K. Y. Leung, "An improved image processing method for assessing multiple cracking development in strain hardening cementitious composites (SHCC)," *Cement and Concrete Composites*, vol. 74, pp. 191–200, 2016.

[11] S. Nikravesh and S. Nezamivand Chegini, "Crack identification in double-cracked plates using wavelet analysis," *Meccanica*, vol. 48, no. 9, pp. 2075–2098, 2013.

[12] T. Rao, X. Li, H. Zhang, and M. Xu, "Multi-level region-based convolution neural network for image emotion classification," *Neurocomputing*, vol. 333, pp. 429–439, 2019.

[13] Y. Ren, J. Huang, Z. Hong et al., "Image-based concrete crack detection in tunnels using deep fully convolutional networks," *Construction and Building Materials*, vol. 234, 2020.

[14] C. V. Dung and L. D. Anh, "Autonomous concrete crack detection using deep fully convolutional neural network," *Automation in Construction*, vol. 99, pp. 52–58, 2019.

[15] X.-W. Ye, T. Jin, and P.-Y. Chen, "Structural crack detection using deep learning-based fully convolutional networks," *Advances in Structural Engineering*, vol. 22, no. 16, pp. 3412–3419, 2019.

[16] C. Qiu, L. Kong, and S. Yang, "Automatic recognition of flavor types of flue-cured tobacco based on GA-SVM algorithm," *Tobacco Science and Technology*, vol. 52, no. 2, pp. 101–108, 2019.

[17] Y. Zhang, Y. Zhou, X. Guo et al., "Self-adaptive k-means based on a covering algorithm," *Complexity*, vol. 2018, Article ID 7698274, 16 pages, 2018.

[18] D. Nameirakpam and J. Yambem, "An image segmentation approach based on fuzzy C-means and dynamic particle swarm optimization algorithm," *Multimedia Tools and Applications*, vol. 79, no. 25-26, pp. 18839–18858, 2020, prepublish.

[19] L. Tang, Z. Yu, C. Ren et al., "A method for obtaining spatial structure information of stereo intersections based on vehicle GPS tracks," *Journal of Transportation Engineering*, vol. 19, no. 5, pp. 170–179, 2019.

[20] A. K. Bhandari, I. V. Kumar, and K. Srinivas, "Cuttlefish algorithm-based multilevel 3-D otsu function for color image segmentation," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 5, pp. 1871–1880, 2020.

[21] R. Chai, "Otsu's image segmentation algorithm with memory-based fruit fly optimization algorithm," *Complexity*, vol. 2021, Article ID 5564690, 11 pages, 2021.

[22] L. Liu and F. Liang, "Ship infrared image edge detection based on an improved adaptive canny algorithm," *International Journal of Distributed Sensor Networks*, vol. 14, no. 3, pp. 1–7, 2018.

[23] C. Sheela and G. Suganthi, "An efficient denoising of impulse noise from MRI using adaptive switching modified decision based unsymmetric trimmed median filter," *Biomedical Signal Processing and Control*, vol. 2020, Article ID 101657, 55 pages, 2020.

[24] M. Ramalingam and R. Thangarajan, "Mutated k-means algorithm for dynamic clustering to perform effective and intelligent broadcasting in medical surveillance using selective reliable broadcast protocol in VANET," *Computer Communications*, vol. 150, pp. 563–568, 2020.

[25] I. A. Iswanto, T. W. Choa, and B. Li, "Object tracking based on Meanshift and particle-kalman filter algorithm with multi features," *Procedia Computer Science*, vol. 157, pp. 521–529, 2019.

Hindawi
Complexity
Volume 2021, Article ID 1288161, 7 pages
https://doi.org/10.1155/2021/1288161

Research Article

Rotational Flocking with Spontaneous Directional Changes

Xiaolu Liu,[1,2] **Guanbo Shao**,[1] **Yudong Tang**,[2] **and Duxin Chen**[1]

[1]*School of Mathematics, Jiangsu Key Laboratory of Networked Collective Intelligence, Southeast University, Nanjing 211096, China*
[2]*School of Automation, Nanjing Institute of Technology, Nanjing 210096, China*

Correspondence should be addressed to Duxin Chen; chendx@seu.edu.cn

Received 11 April 2021; Accepted 19 July 2021; Published 30 July 2021

Academic Editor: Xiao Ling Wang

Revealing the underlying decision-making strategy governing the high-group polarization accompanied by conflicting individual preferences may play a central part in the lives of social animals. Hereby, we construct a structured spin model in accordance with empirical validation, which shows how distinct individual preferences converge from one consensus homeostasis to another lowest-energy equilibrium. To verify the theoretical derivation, we use high-resolution spatiotemporal GPS data of a flock of thirty pigeons and study the dynamical evolution mechanism of systemic spins. Therein, we find successful rotational direction transitions requiring a sufficient number of supporters. A few initiators trigger the phase transition from one equilibrium to another, where the symmetric transient state indicates a diamond hierarchical network being completed by the intermediates and the rear individuals. By further studying the nature, we reveal that decision-making sequences are strongly triggered and influenced by individual positions and the leader-follower relationship. Thus, we can predict which individual is more likely to make the decision before the initial transition moment and who will draw the complete stop. Consequently, the revealed decision-making strategy facilitates a comprehensive understanding of collective behavioral transition.

1. Introduction

Collective movements of natural animal groups are among the most compelling social manifestations in nature. The underlying interindividual interaction principles, decision-making strategies, and the matters triggering transitions from nonequilibrium to a consensus homeostasis have classically attracted considerable attention [1]. A long-standing question is how gregarious animals with conflicting moving preferences affect the flow of information and give group members distinct weights in making decisions. Generally, collective decisions can be dominated by a single despotic leader, can be determined by the influence of hierarchy, or emerge from a shared democratic process [2]. Evidence collected from schooling fish suggests that the process is democratic, with nearest neighbors and the majority shaping overall collective behavior, and uninformed individuals could promote such democratic consensus [3]. In animal groups with obvious hierarchical social structures, such as elephants [4], dolphins [5], and primates [6], however, such democracy may be replaced by despotism. Thus, revealing the decision-making strategy governing the high polarization in collective behavior remains a core challenge for understanding the social complexity of animals. Following the research line, diverse modeling studies have been developed to elucidate complex collective decision-making processes and strategies in animal groups. The fundamental aspects of widely-used models, in common, are that they start with a hypothesis about individual decision-making strategies and then proceed the integration process as an aggregated outcome of collective decision. Such hypotheses about individual decision-making mainly include quorum response [7–9], state-dependent alignment [10–14], assertiveness/submission [15–17], and compromise strategies [2, 18, 19]. These agent-based models adopting various decision-making rules propelling individuals have been sufficiently realistic to reproduce numerous observed phenomena and hence are beneficial for better understanding the system complexity consisting of many self-propelled entities [20, 21].

More crucially, in order to deeply ascertain what decision-making strategy is based on for coordinated individuals, empirical study on collective movements and validation of those antecedent types of models are necessary. Recently, Ballerini et al. [22] adopted an image processing method in collecting the data of a huge flock of starlings and proposed a possibility of sharing information that each bird communicated with only a fixed number of topological neighbors, instead of individuals within a specific metric distance as assumed in the celebrated standard Vicsek model [23]. By using another newly emerged technique in collecting experimental data, i.e., GPS tracking device, Nagy et al. [24] revealed a hierarchical leadership structure in small pigeon flocks, which reflects the multilayer leader-follower relationship governing their decision-making process. Such a hierarchical leadership was found independent from the dominance rank in pigeon flocks [25]. In contrast, although the network in dogs [26] constructed from the leader-follower relations is also hierarchical, significant correlations were found between hierarchy and age, dominance rank, trainability, controllability, and aggression measures. Indeed, despotic decision-making is more obvious in the predatory behavior of wolf packs [27]. Previous studies on the mystery of despotism in African elephant herds provided a better sense that dominance rank relationships were transitive within families and highly asymmetrical within dyads, such that older, larger females consistently dominated decision-making of other individuals [28, 29]. An untested hypothesis thus naturally arises that animal groups with higher wisdom are more likely to adopt despotic decision-making strategy.

Hereby, to infer the decision-making strategy, we focus on free flights as shown in Figure 1(a) and carry out a detailed investigation on high spatial-temporal resolution GPS datasets. Therein, three flocks has ten individuals with 30 releases labeled as flocks A, B, and C from [25]. Each release lasts from two to seven minutes. Due to the limitation of the GPS device in z-axis and the average standard deviation of flight altitude in each release being sufficiently small, it suffices to only use the x- and y-axes data for investigation as the previous study [24, 25] did. Subsequently, we construct a structured spin model which shows how distinct individual preferences converge from one consensus homeostasis to another lowest-energy equilibrium. To verify the theoretical derivation, we study the dynamical evolution mechanism of systemic spins. Therein, we find successful transitions requiring a sufficient number of supporters. By further studying the origins, we reveal that decision-making sequences are triggered and influenced strongly by spatial locations and the leader-follower relationship. Thus, we can infer which individual is more likely to make the decision before the transition moment and who will complete the stop.

2. Results

2.1. Inferring the Decision-Making Strategy. Dynamical equilibrium is an axiomatic concept, which can be seen as no macroscopic change occurs in the system of internal thermodynamic equilibrium. Analogously, in bird flock flights with high polarization, e.g., homing flight of pigeon flocks [30], it is often encountered that the interagent spatial position keeps relatively constant, which can be considered as an equilibrium state. The degree of global ordering in a flock can be measured by the so-called order parameter $\phi = 1/N \left\| \sum_{i=1}^{N} \overrightarrow{v}_i(t)/\overrightarrow{v}_i(t) \right\|$, which is employed as a standard index of global order during the study of collective animal behavior [23]. Therein, $\overrightarrow{v}_i(t)$ denotes the velocity of bird i in the horizontal plane and N is the total number of individuals in the flock. Apparently, the index ϕ is zero for totally disordered and one for completely synchronized flocks. Furthermore, we introduce the concept of internal angular momentum from quantum mechanics [31], namely, the spin of an individual, $\overrightarrow{s}_i = \chi_i \cdot (\overrightarrow{v}_i \times \mathrm{d}\overrightarrow{v}_i/\mathrm{d}t)$ [32], with χ_i denoting the moment of inertia of individual i. In this study, we focus on the free flight of pigeon flocks, where individuals hover above their home loft with spontaneous changes of rotational directions. For simplicity, we consider every turning as a part of uniform circular motion without noise and normalize the spin in this study, i.e., $|\overrightarrow{s}_i| = |\overrightarrow{s}_j| = 1$ for any pair of individuals i and j. Thus, $\overrightarrow{s}_i$ is orthogonal to both the velocity and the centripetal acceleration vectors, and its positive and negative normalized values will denote the clockwise (C.W.) and the counterclockwise (C.C.W.) turning by simply using the right-hand screw rule. We then develop the alternative order parameter, namely, $S = 1/N \sum_{i=1}^{N} \overrightarrow{s}_i$, which characterizes the collective decision made by the entire flock. Note that S is an indicative vector with the value being equivalent to the summation of $\overrightarrow{s}_i$. The value of index S is zero for totally conflicted states with half supporters and half opponents, and $S = \pm 1$ for completely consensus in decision-making in C.W. and C.C.W. turnings, respectively.

As shown in Figures 1(b) and 1(c), we exhibit the probability distributions of the decision-making polarization S. With the reduction and increase of the index values, the entire flock gradually changes the rotational direction from counterclockwise to clockwise and vice versa, respectively. The symmetric probability distribution of S indicates that more individuals change their rotational direction in the medium period, whereas the number of directional switching pigeons reduces gradually in the earlier and later periods. The fitting curves indicate that spontaneous changes of decision on flight directions follow a Gaussian distribution. It provides an evidence how a heterogeneous flock of pigeons with varying stamina can achieve consensus in decision-making of the rotational direction in free flight. An explanation lies in previous numerical studies that every individual has a depletion time which follows a Gaussian distribution. When a pigeon feels tired or just wants to change the rotational direction, but most of the others not, it must compromise unwillingly but with increasingly greater intension to change. Thus, when a sufficient number of willing members have been accumulated to change their rotational direction, they drive the entire flock to switch [33, 34]. Moreover, the distributions of S shown in Figure 1 suggest that pigeon flocks in free flight

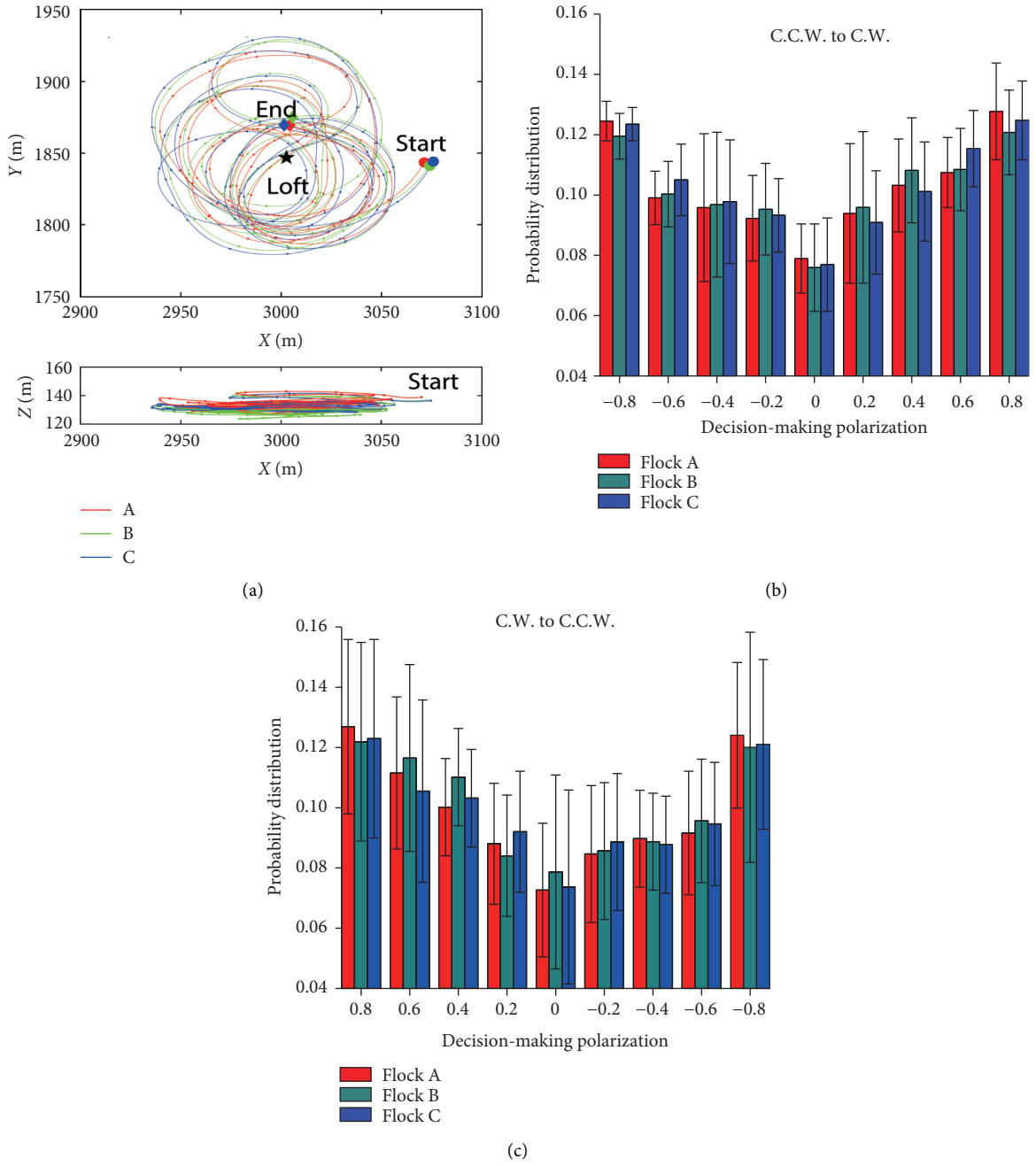


Figure 1: Trajectories and the distributions of the systematic spin. (a) Trajectories of ten pigeons in flock *A*. Each trajectory lies approximately on a plane. (b) Probability distributions of the systematic spin from C.C.W. turning to C.W. turning and (c) the reverse case. It is observed that the spontaneous changes of rotational directions follow a symmetric distribution.

employ a symmetric diamond-shaped structure of decision-making, instead of a pyramid-type leader-follower hierarchical structure [24]. It means that not only the leader initiates the change of rotational direction, there exists a symmetric rear individual completing the transient state to another equilibrium.

Theoretical validation: the distributions of S suggest a symmetric diamond-shaped structure of decision-making, which is in contrast to the previous pyramid-type [24], star-shaped [35], and V-shaped hierarchical structures [36] observed in bird flocks. Note that we normalize the phase transition time T defined as the completing time of a transition from one equilibrium to another, namely, the time from the initiator to the last one changing its direction of spin ($T = 3.25 \pm 1.67\,s$).

Theoretical descriptions of collective motion are widely based on alignment dynamics, namely, each individual tends to keep its direction of motion as close as possible to that of its neighbors. The Vicsek style dynamics of collective motion in a noiseless case is as follows:

$$\vec{v}_i(t+1) = \vec{v}_i(t) + \frac{1}{N_i}\sum_{j\in N_i}\vec{v}_j(t), \tag{1}$$

where the vector v_i is the velocity of individual i and the summation extends over all neighbors j of i. Analogously, we can follow the spirit of the alignment hypothesis and subsequently have the spin dynamics:

$$\vec{s}_i(t+1) = \vec{s}_i(t) + \sum_{j\in N_i} J_{ij}\vec{s}_j(t), \tag{2}$$

where J_{ij} is the interaction strength between two neighboring individuals i and j in the transition state of decision-making. We assume that, in a small pigeon flock, all individuals are mutual neighbors for each other. This dynamics of interaction rules is reminiscent of the classical Heisenberg model, which is formulated in a 3-dimensional lattice, $H = -\sum_{i,j} J_{ij}\vec{s}_i \cdot \vec{s}_j$ with $\vec{s}_i \in \mathbb{R}^3, |\vec{s}_i| = 1$ denoting the normalized spin and $J_{ij} = 1$ only if i and j are neighbors. A common simplification is to assume that all of the nearest neighbors have the same interaction strength, and then $J_{ij} = J$ for all individual pairs i and j. If the spin is one of the two canonical variables, the other would be velocity phase in the internal space of individual i in a uniform circular motion. However, the transient decision-making process definitely cannot be captured according to the simplest Heisenberg model with canonical variables spin and velocity phase.

In a continuous-time case of a flock with N individuals, the Hamiltonian and the update rule (2) can be equivalently derived to the following equations:

$$H = -J\sum_{i,j}\vec{s}_i \cdot \vec{s}_j, \frac{\vec{s}_i(t)}{\mathrm{d}t} = -\frac{\partial H}{\partial \vec{s}_i} = -J\sum_j \vec{s}_j. \tag{3}$$

Hence, each individual changes its decision following a social force $-(\partial H/\partial \vec{s}_i)$, produced by its neighbors. In the transition state of decision-making, it is clear that $S = 1$ evolves to $S = -1$ from consensus C.W. turning to a C.C.W. equilibrium, or the opposite for the reverse case. The total spin $S = \sum_i \vec{s}_i$ is constant for an equilibrium of C.W. or C.C.W turning due to the rotational symmetry, respectively. However, there is no conservation of S in the transition state since we know clearly the variant time-dependent evolution process of S, and we can obtain the following:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \sum_i \frac{\mathrm{d}\vec{s}_i}{\mathrm{d}t} = -\sum_i \frac{\partial H}{\partial \vec{s}_i} \sim -JN\sum_j \vec{s}_j. \tag{4}$$

Relation (4) contains the theoretical dynamics of S, which indicates that more supporters induce a faster transition of rotational direction change. Hamiltonian equation (3) in this study consists of discrete variables that represent intrinsic properties known as spins of the individuals allowing interaction with neighbors.

2.2. Matters Inducing Decision-Making. To explore the factors inducing the transient symmetric decision-making of pigeon flocks, we investigate the spatial locations of all the individuals. As shown in Figure 2, the sequence of individuals making decision in flock A is significantly influenced by their locations. The average location of the initiator in a transition from a C.C.W. turning to a C.W. turning always plays the leading role at the right front, whereas the followers are linearly distributed from the right front to the left rear (see flocks B, C, and D in Supplementary Materials). Interestingly, the individual who is the last to complete the turning locates symmetrically behind. It leaves the strong impression that the initiator at the cusp makes crucial decision (e.g., escaping natural enemies in the prey), others follow in order, and the protector locating symmetrically behind draws the completing stop.

The hierarchical leadership structure in small pigeon flocks has been revealed in the previous study [24], which suggests a multilayer leader-follower relationship governing their decision-making process. Furthermore, such a hierarchical leadership was found independent from the dominance rank in pigeon flocks [25]. Since spatial locations and leader-follower relationships are significantly correlated, we subsequently exhibit the distribution of the initiator occupation rate in Figure 3. Compared with the previous study [25] on the dominance rank and the hierarchical leadership structure, it is observed that the initiator occupation rate has no significant correlation with the dominance rank including feeding-queueing and pecking-order ranks. Moreover, it also suggests that age, weight, and sex which are the main factors inducing the dominance rank [25] have no correlation with decision-making sequence in pigeon flocks. However, to our surprise, the individual occupying the highest rate of the initiator position is just the one in the top layer of the hierarchical leadership structure constructed by the time delay in the velocity-velocity correlation [24]. Note that the hierarchical structure in Figure 3 indicates the average leader-follower relationship during the entire free flight. Thus, we may suggest that the leader in the top layer would be the initiator of decision-making and fly in front of the flock in most cases. Therefore, we have established the link among decision-making, spatial location, and hierarchical leadership, and we can infer who would be the initiator of decision-making by the most intuitive spatial distribution in pigeon flocks.

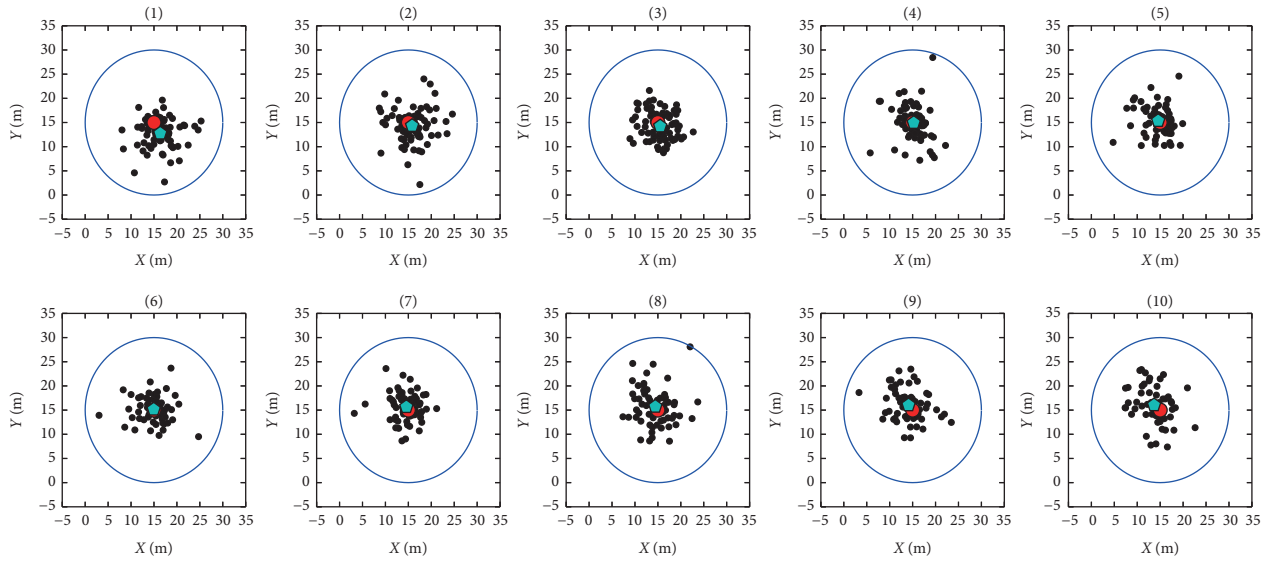


FIGURE 2: Spatial locations of individuals in flock A following the decision-making sequence from a C.C.W. turning to a C.W. turning. Numbers 1 to 10 indicate the switching sequence of the individuals. Here, the red point represents the flock center, the blue pentagram means the average center of the focal individual, and the black points denote the relative positions to the center of the individuals.

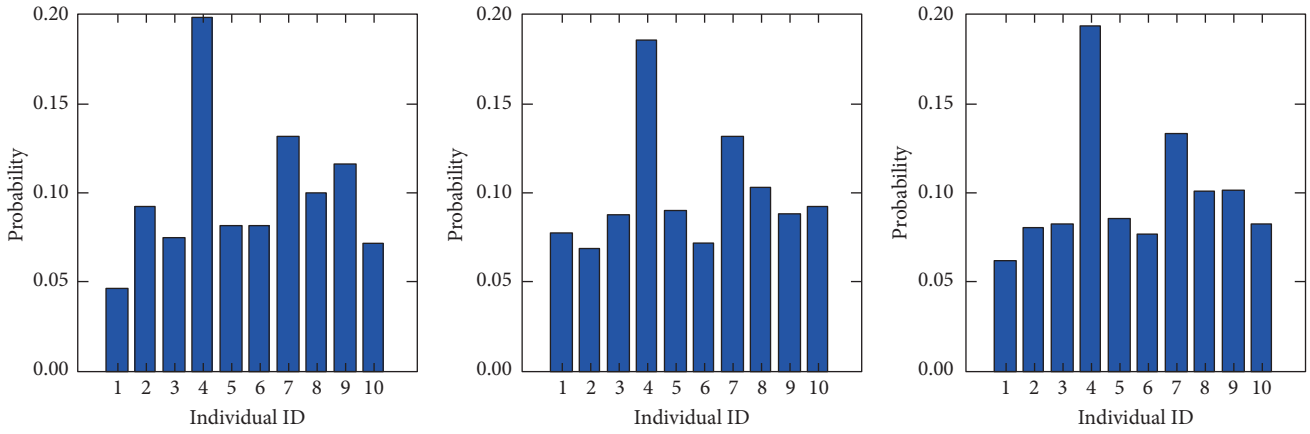


FIGURE 3: Probability distributions of the initiator occupation rate. (a) From C.W. to C.C.W. (b) From C.C.W. to C.W. (c) Average probability distribution of the initiator occupation rate.

3. Discussion and Conclusion

Besides focusing on the appealing behavioral characteristics of collective movements, e.g., consensus and other specific patterns, it is crucial to explore the mechanisms and dynamics of decision-making processes at the individual level. Classically, collective decision-making is viewed as the outcome of the leader-follower relationship among individuals, but how decisions are taken is still an under-investigated question. The hierarchical or centralized control model has been challenged by recent theoretical and experimental findings, which suggest that leadership can be more distributed. Moreover, self-organized processes can account for collective movements in many different species, even in those that are characterized by high cognitive complexity.

Hereby, to infer the decision-making strategy, we focus on free flights and carry out a detailed investigation on comprehensive datasets of four pigeon flocks with different flock numbers. We construct a minimally structured spin model in accordance with the maximum entropy method, which shows how distinct individual preferences converge from one consensus homeostasis to another lowest-energy equilibrium. To verify the theoretical derivation, we study the dynamical evolution mechanism of systemic spins denoting group decision, and we find that, in addition to

foreseeable compromise of a small quantity of initiators, successful transitions require a sufficient number of supporters. A few initiators trigger the phase transition from one equilibrium to another, where the symmetric transient state indicating the hierarchical network is completed by the intermediates and the rear individual. This observation of the diamond-shaped decision-making structure is different from typical pyramid-type, star-shaped, and V-shaped hierarchical structures observed in bird flocks.

To reveal the origins of the symmetric decision-making of pigeon flocks, we find that decision sequences are triggered and influenced strongly by spatial locations and the leader-follower relationship, rather than dominance rank. This would not be surprising since the hierarchical leadership patterns observed in bird flocks arise from an anonymous, self-organizing mechanism related to individual differences in flight speed. Leaders learned more effectively during free flights, and a possible explanation is that faster birds flying in front of the flock have no choice but to make decisions, whereas the slower followers are able to rely on social information. The enhanced learning by leaders would be expected to reinforce a particular direction of information transfer through the flock. However, the revealed symmetric decision-making strategy implies that other individuals except the initiator are not equivalent followers. The most special follower is the one flying in the rear, which leaves us the strong impression that it acts as a protector responding the initiator. This is due to the fact that the head and the tail would be the vital part when pigeon flocks are facing the predators. To sum up, we have established the link among decision-making, spatial locations, and hierarchical leadership structure, which allows us to predict which individual is more likely to make the decision before the transition moment and who will complete the transition.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the Fundamental Research Funds for the Central Universities (Grant no. 2242019K40111), the National Natural Science Foundation of China (Grant no. 61903079), "Zhishan" Scholars Programs of Southeast University, Scientific Research Foundation for the introduction of talents of NJIT (YKJ202015), project funded by Jiangsu Postdoctoral Science Foundation (2021K397C), and the Jiangsu Provincial Key Laboratory of Networked Collective Intelligence (Grant no. BM2017002).

References

[1] T. Vicsek and A. Zafeiris, "Collective motion," *Physics Reports*, vol. 517, no. 3-4, pp. 71–140, 2012.

[2] I. D. Couzin, J. Krause, N. R. Franks, and S. A. Levin, "Effective leadership and decision-making in animal groups on the move," *Nature*, vol. 433, no. 7025, pp. 513–516, 2005.

[3] I. D. Couzin, C. C. Ioannou, G. Demirel et al., "Uninformed individuals promote democratic consensus in animal groups," *Science*, vol. 334, no. 6062, pp. 1578–1580, 2011.

[4] C. Sueur and O. Petit, "Shared or unshared consensus decision in macaques?" *Behavioural Processes*, vol. 78, no. 1, pp. 84–92, 2008.

[5] J. Krause, D. Lusseau, and R. James, "Animal social networks: an introduction," *Behavioral Ecology and Sociobiology*, vol. 63, no. 7, pp. 967–973, 2009.

[6] A. Strandburg-Peshkin, D. R. Farine, I. D. Couzin, and M. C. Crofoot, "Shared decision-making drives collective movement in wild baboons," *Science*, vol. 348, no. 6241, pp. 1358–1361, 2015.

[7] D. J. T. Sumpter and S. C. Pratt, "Quorum responses and consensus decision making," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1518, pp. 743–753, 2009.

[8] J. L. Deneubourg and S. Goss, "Collective patterns and decision-making," *Ethology Ecology & Evolution*, vol. 1, no. 4, pp. 295–311, 1989.

[9] R. Beckers, J. L. Deneubourg, and S. Goss, "Modulation of trail laying in the antLasius Niger (Hymenoptera: formicidae) and its role in the collective selection of a food source," *Journal of Insect Behavior*, vol. 6, no. 6, pp. 751–759, 1993.

[10] S. A. Rands, G. Cowlishaw, R. A. Pettifor, J. M. Rowcliffe, and R. A. Johnstone, "Spontaneous emergence of leaders and followers in foraging pairs," *Nature*, vol. 423, no. 6938, pp. 432–434, 2003.

[11] L. Conradt, J. Krause, I. D. Couzin, and T. J. Roper, ""Leading according to need" in self-organizing groups," *The American Naturalist*, vol. 173, no. 3, pp. 304–312, 2009.

[12] K. Bhattacharya and T. Vicsek, "Collective decision making in cohesive flocks," *New Journal of Physics*, vol. 12, no. 9, Article ID 093019, 2010.

[13] D. Chen, X. Liu, and W. Yu, "Finite-time fuzzy adaptive consensus for heterogeneous nonlinear multi-agent systems," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 3057–3066, 2020.

[14] X. Liu, Y. W. Wang, D. Chen, and H. Chen, "Adaptive fuzzy fault-tolerant control for a class of unknown non-linear dynamical systems," *IET Control Theory & Applications*, vol. 10, no. 18, pp. 2357–2369, 2016.

[15] R. A. Johnstone and A. Manica, "Evolution of personality differences in leadership," *Proceedings of the National Academy of Sciences*, vol. 108, no. 20, pp. 8373–8378, 2011.

[16] L. Conradt and T. J. Roper, "Democracy in animals: the evolution of shared group decisions," *Proceedings of the Royal Society B: Biological Sciences*, vol. 274, no. 1623, pp. 2317–2326, 2007.

[17] L. Conradt and T. J. Roper, "Conflicts of interest and the evolution of decision sharing," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1518, pp. 807–819, 2009.

[18] E. A. Codling, J. W. Pitchford, and S. D. Simpson, "Group navigation and the "Many-wrongs principle" in models of animal movement," *Ecology*, vol. 88, no. 7, pp. 1864–1870, 2007.

[19] C. List, C. Elsholtz, and T. D. Seeley, "Independence and interdependence in collective decision making: an agent-based model of nest-site choice by honeybee swarms,"

Philosophical Transactions of the Royal Society B: Biological Sciences, vol. 364, no. 1518, pp. 755–762, 2009.

[20] D. Chen, Y. Wang, G. Wu, M. Kang, Y. Sun, and W. Yu, "Inferring causal relationship in coordinated flight of pigeon flocks," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 11, Article ID 113118, 2019.

[21] D. Chen, W. Li, X. Liu, W. Yu, and Y. Sun, "Effects of measurement noise on flocking dynamics of Cucker-Smale systems," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 10, pp. 2064–2068, 2020.

[22] M. Ballerini, N. Cabibbo, R. Candelier et al., "Interaction ruling animal collective behavior depends on topological rather than metric distance: evidence from a field study," *Proceedings of the National Academy of Sciences*, vol. 105, no. 4, pp. 1232–1237, 2008.

[23] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, "Novel type of phase transition in a system of self-driven particles," *Physical Review Letters*, vol. 75, no. 6, pp. 1226–1229, 1995.

[24] M. Nagy, Z. Ákos, D. Biro, and T. Vicsek, "Hierarchical group dynamics in pigeon flocks," *Nature*, vol. 464, no. 7290, pp. 890–893, 2010.

[25] M. Nagy, G. Vásárhelyi, B. Pettit, I. Roberts-Mariani, T. Vicsek, and D. Biro, "Context-dependent hierarchies in pigeons," *Proceedings of the National Academy of Sciences*, vol. 110, no. 32, pp. 13049–13054, 2013.

[26] Z. Ákos, R. Beck, M. Nagy, T. Vicsek, and E. Kubinyi, "Leadership and path characteristics during walks are linked to dominance order and individual traits in dogs," *PLoS Computational Biology*, vol. 10, no. 1, Article ID e1003446, 2014.

[27] L. D. Mech, "Wolf-pack buffer zones as prey reservoirs," *Science*, vol. 198, no. 4314, pp. 320-321, 1977.

[28] E. A. Archie, T. A. Morrison, C. A. H. Foley, C. J. Moss, and S. C. Alberts, "Dominance rank relationships among wild female African elephants, Loxodonta africana," *Animal Behaviour*, vol. 71, no. 1, pp. 117–127, 2006.

[29] G. Wittemyer and W. M. Getz, "Hierarchical dominance structure and social organization in African elephants, Loxodonta Africana," *Animal Behaviour*, vol. 73, no. 4, pp. 671–681, 2007.

[30] D. Chen, T. Vicsek, X. Liu, T. Zhou, and H.-T. Zhang, "Switching hierarchical leadership mechanism in homing flight of pigeon flocks," *EPL (Europhysics Letters)*, vol. 114, no. 6, p. 60008, 2016.

[31] A. Attanasi, A. Cavagna, L. Del Castello et al., "Information transfer and behavioural inertia in starling flocks," *Nature Physics*, vol. 10, no. 9, pp. 691–696, 2014.

[32] A. Cavagna, L. Del Castello, I. Giardina et al., "Flocking and turning: a new model for self-organized collective motion," *Journal of Statistical Physics*, vol. 158, no. 3, pp. 601–627, 2015.

[33] D. Chen, X. Liu, B. Xu, and H.-T. Zhang, "Intermittence and connectivity of interactions in pigeon flock flights," *Scientific Reports*, vol. 7, no. 1, Article ID 10452, 2017.

[34] D. Chen, B. Xu, T. Zhu, T. Zhou, and H. T. Zhang, "Anisotropic interaction rules in circular motions of pigeon flocks: an empirical study based on sparse Bayesian learning," *Physical Review E*, vol. 96, Article ID 022411, 2017.

[35] Z. Chen, H.-T. Zhang, X. Chen, D. Chen, and T. Zhou, "Two-level leader-follower organization in pigeon flocks," *EPL (Europhysics Letters)*, vol. 112, no. 2, Article ID 20008, 2015.

[36] J. R. Usherwood, M. Stavrou, J. C. Lowe, K. Roskilly, and A. M. Wilson, "Flying in a flock comes at a cost in pigeons," *Nature*, vol. 474, no. 7352, pp. 494–497, 2011.

Hindawi
Complexity
Volume 2021, Article ID 4113022, 15 pages
https://doi.org/10.1155/2021/4113022

Research Article

Collaborative Control of Multimotor Systems for Fixed-Time Optimisation Based on Virtual Main-Axis Speed Compensation Structure

Changfan Zhang[1], **Mingjie Xiao**[1], **Jing He**[1], **Zhitian Liu**[2], **Xingxing Yang**[1], **Qian Zhang**[1] and **Hongrun Chen**[1]

[1]*College of Electrical and Information Engineering, Hunan University of Technology, Zhuzhou 412000, China*
[2]*State Grid Fuyang City Chengjiao Electric Power Supply Company, Fuyang 236000, China*

Correspondence should be addressed to Jing He; hejing@263.net

Received 29 May 2021; Accepted 14 July 2021; Published 28 July 2021

Academic Editor: Xiao Ling Wang

In response to the high-speed and high-precision collaborative control requirements of the multimotor system for filling, a new type of virtual master-axis control structure is proposed and a multimotor fixed-time optimized collaborative control algorithm is designed. Firstly, coupling relationship between virtual and slave motors is effectively established by designing a velocity compensation module for the virtual motor. Secondly, the sliding mode observer (SMO) is used to reconstruct the composite disturbance composed of motor parameter perturbation and load disturbance. Finally, the variable gain terminal sliding mode controller (SMC) is designed to ensure that each slave motor can track the given value within a fixed time. The fast convergence of the system can be proved by the fixed-time convergence theorem and Lyapunov's stability theorem. The simulation results show that, compared with the traditional virtual main-axis control strategy, the proposed method is more effective for the tracking control of each slave motor in the initial stage.

1. Introduction

High-performance collaborative control of filling multimachine is one of the key technologies in the research and development of thick sauce and viscous food packaging equipment. It is also an important embodiment of high-speed and high-precision filling operation. At present, studies on filling multimachine collaborative control technologies usually focus on synchronous control of multimotor systems [1, 2]. In the multimotor synchronous control, the synchronous topology and control algorithm are two aspects worthy of attention. Therefore, the following is a summary of domestic and foreign research status from these two aspects.

Main multimotor synchronous control strategies for different synchronous topology structures are master-slave synchronous control [3], adjacent coupling synchronous control [4], relative coupling synchronous control [5], and virtual main-axis synchronous controls [6]. Yao et al. redefined adjacent cross-coupling error and designed auto disturbance rejection controller by extracting the Laplacian matrix [7]. Xu et al. proposed an optimal control strategy based on combined cross-coupling errors against various coupling errors and disturbance in multilayer and multiaxis control systems [8]. Huang et al. combined adjacent cross-coupling control strategy with adaptive algorithm to ease jitters when the speed is stable by considering phase and speed synchronisation control of multiactuators in the oscillation system [9]. Sun et al. combined fuzzy control technology with ring coupling synchronous structure to achieve the reliable operation of multimotor driving system [10]. Tao et al. proposed an improved adjacent coupling synchronous control strategy to simplify the structure of velocity synchronous control by introducing a new coupling

coefficient in the traditional adjacent coupling method [11]. Shi et al. put forward an improved multimotor relative coupling synchronous control, which also decoupled tracking and synchronisation errors of the system [12]. Zhang et al. mainly improved the relative coupling synchronization structure and then realized the smooth operation of the intelligent filling equipment for sticky food under complex working conditions [13]. Xu et al. investigated the adaptive synchronisation of coupled harmonic oscillator with switching topology and proposed an edge-based servo synchronous adaptive control protocol [14]. These control strategies promote the development of multimotor synchronous control technologies. However, the virtual main-axis control structure exerts a strong control effect on the application of multimotor synchronous technology with a unique feedback mechanism and characteristics of virtual motor as the navigator. Zhang and He explored a multimotor virtual main-axis control strategy based on the observer to reflect the dynamic relationship between virtual and slave motors accurately [15]. Lin and Cai estimated the composite disturbance of the motor using sliding mode observer to ensure that slave motors reach high synchronisation accuracy [16]. He and Chen focused on the total consensus control in multimotor cooperation to ensure that the total power of the system remains constant under power loss of a single motor [17].

The synchronous control algorithm mainly includes multiagent consensus [18, 19] and sliding mode variable structure control [20, 21] algorithms due to different design ideas of multimotor synchronous control. Xu and Li evaluated the multiagent consensus algorithm and proposed a semiglobal consensus control protocol based on low-gain state feedback with consideration for the intermittent semiglobal consensus problem of high-order multiagent system under input saturation [22]. Wang et al. analysed the structure of distributed observer of continuous linear time-invariant system and reduced the dependence of the system on topology structure information using adaptive rule [23]. Wang and Su investigated the collaborative control of linear time-invariant network system with uncertainty and constrained states of all agents by constructing an interval observer [24]. Liu and Su explored containment control of multiagent system based on data information of intermittent sampling position and obtained necessary and sufficient conditions for containment in the case of time delay and delay free [25]. Liu and Su further examined necessary and sufficient conditions, including control protocol under the condition of time delay, and effectively solved the problem of using the second-order multiagent system to control with only sampling position data [26]. Zheng et al. proposed an optimal operation model for microgrid based on the multiagent system to save the entire energy cost of locally observable convex function [27]. Fan and Wang assessed problems on following the consensus of a second-order multiagent system with input saturation and proved the stable consistency of the multiagent system using adaptive control technology [28]. Zhai et al. proposed a new pulse delay control method based on sampled data to realise the pulse synchronisation of multiagent systems [29]. Liu et al. examined the group controllability of the multiagent system, proposed a general definition of group controllability, and established a correlation criterion of group controllability [30]. In the research of sliding mode variable structure control algorithm, Shen et al. developed a trajectory tracking control scheme for the Mars orbiter and proposed a method of adaptive fixed-time control [31]. Shao et al. proposed a state observer with two polynomial feedback terms to identify unmeasurable speed states and unknown uncertain disturbances. The designed speedless fixed-time controller integrates the obtained observations [32]. Sun and Wang designed a sliding mode controller for angle tracking control of wearable exoskeleton based on model-free fixed-time control algorithm [33].

Although the multimachine cooperative control technology has been extensively investigated [34, 35], studies on multimachine high-performance collaborative control for thick sauce and viscous food filling are limited. In particular, the viscous characteristics of the filling material greatly hinder the speed and efficiency of the filling, resulting in two technical problems in the high-speed and high-precision operation of the filling multimotor system. (1) Virtual main-axis synchronous control is a common control strategy for filling multimotor production that can effectively improve the synchronisation accuracy of multimachine collaborative operation. The rotation speed of the slave filling conveying motor cannot be fed back to the virtual motor and the insufficient coupling correlation between them are limitations of the current feedback mechanism of the virtual main-axis control structure in terms of the multimotor collaborative control strategy. (2) As one of the effective controllers of the servo system, the sliding mode controller is widely used in various filling production occasions. Approaching speed of the existing sliding mode controller must be improved because its convergence speed demonstrates difficulty in meeting the requirements of high-speed operation of the servo system, especially when the sliding mode operation reaches the sliding surface from any initial position, in terms of the multimachine collaborative control algorithm.

Hence, a new multimotor collaborative control method for fixed-time optimisation based on the virtual main-axis structure is proposed in this study to solve these problems and meet the demand of the filling multimotor system for high-performance collaborative control. This study provides the following contributions:

(1) A speed compensator is added to the virtual motor to improve the existing virtual main-axis control structure, and the coupling relationship between virtual and slave motors in the speed is established given that the information on the velocity of slave motors and rotor position cannot be fed back to the virtual motor in the traditional virtual main-axis structure.

(2) A new type of variable gain terminal SMC is designed to realize the high-speed and high-precision synchronous control of the filling multimotor system. It mainly solves the problem of slow convergence speed when the system error reaches the terminal sliding surface from the initial state.

This paper is organized as follows: by taking advantage of the existing virtual main shaft synchronization control structure, a speed compensator is added to the virtual motor and the synchronization topology is improved in Section 2. The multimotor system model is established in Section 3. The sliding mode observer is used to overcome the parameter perturbation and load changes of the multimotor system in Section 4. The system controller is designed to ensure that the speed of each slave motor can quickly track the given value in Section 5. The experimental simulation is presented in Section 6. Finally, the conclusion is given in Section 7.

2. Improved Virtual Main Shaft Control Structure Design

A virtual main-axis control structure with speed compensator is designed for the multimotor synchronous technology in this section on the basis of fully absorbing the characteristics and principles of the existing virtual main-axis strategy. Specific details of each module of this control structure designed are shown in Figure 1.

Figure 1 shows the permanent magnet synchronous motor (PMSM) model, virtual motor, tracking error of permanent magnet synchronous motor, sliding mode function, controller module, virtual controller module, observer module, and virtual motor speed compensator as well as detailed formulas, logical relations, and related variables of each module. Compared with the diagrams of various modules shown in Figure 1, the golden area is the proposed virtual motor speed compensator designed in this study. The compensation signal generated by the virtual motor compensator can be directly fed back to the virtual motor and can also be adjusted by selecting a suitable virtual motor's moment of inertia. Although the compensation signal is known and available, $\omega_{\text{vir}} - \omega_j$ are bounded variables. Therefore, the output value of the virtual motor speed compensator is regarded as a variable with a known upper bound that can reduce the design difficulty and calculation of the virtual motor controller. The blue area denotes the designed variable-gain terminal SMC in this study. A dynamic switching gain term, including the sliding mode function, is designed to improve the convergence speed of the sliding mode movement because the tracking error of each slave motors is large in the initial state.

3. Establishment of the Multimotor System Model

The PMSM of the driving system for filling production is the research object of this study. The second-order kinetic model of motor j in the multimotor system can be expressed as follows [13]:

$$\begin{cases} \dot{\theta}_j = \omega_j, \\ \dot{\omega}_j = -a_j\omega_j + b_j u_j - p_j T_{Lj}, \end{cases} \quad j = 1, 2, \ldots, n, \tag{1}$$

where $a_j = R_{\Omega j}/J_j$; $b_j = 3/2 \cdot n_{pj}^2 \psi_{fj}/J_j$; $p_j = n_{pj}/J_j$; $u_j = i_{qj}$; θ_j and $w_j$are the position and electrical angular velocity of the motor rotor, respectively; T_{Lj} is the load torque on the shaft; $R_{\Omega j}$ is the rotational resistance coefficient; n_{Pj} is the number of pole pairs; J_j is the moment of inertia; Ψ_{fj} is the rotor flux; i_{qj} is the q-axis components of the stator current; and u_j is the controller to be designed.

The state equation of virtual motor based on formula (1) is expressed as follows:

$$\begin{cases} \dot{\theta}_{\text{vir}} = \omega_{\text{vir}}, \\ \dot{\omega}_{\text{vir}} = -a_{\text{vir}}\omega_{\text{vir}} + b_{\text{vir}}u_{\text{vir}} - p_{\text{vir}}T_{\text{Lvir}}, \end{cases} \tag{2}$$

where $a_{\text{vir}} = R_{\Omega\text{vir}}/J_{\text{vir}}$; $b_{\text{vir}} = 3/2 \cdot n_{p\text{vir}}^2 \psi_{f\text{vir}}/J_{\text{vir}}$; $p_{\text{vir}} = n_{p\text{vir}}/J_{\text{vir}}$; $u_{\text{vir}} = i_{q\text{vir}}$; θ_{vir} and ω_{vir} are the virtual motor rotor position and electrical angular velocity, respectively; $T_{L\text{vir}}$ is the virtual motor load torque; $R_{\Omega\text{vir}}$ is the virtual motor rotational resistance coefficient of motor; $n_{p\text{vir}}$ is the number of pole pairs of virtual motor; J_{vir} is the virtual motor moment of inertia; $\psi_{f\text{vir}}$ is the virtual motor rotor flux; $i_{q\text{vir}}$ is the virtual motor q-axis components of the stator current; and $u_{\text{vir}}$is the virtual controller to be designed.

Motor parameters will be subjected to uncertain perturbation during long-time operation of the motor, and the motor load torque will change uncertainly due to the change of filling materials in the filling process. Hence, formula (1) can be transformed into the following:

$$\begin{cases} \dot{\theta}_j = \omega_j, \\ \dot{\omega}_j = -\left(\overline{a}_j + \Delta a_j\right)\omega_j + \left(\overline{b}_j + \Delta b_j\right)u_j - p_j\Delta T_{Lj}, \end{cases} \quad j = 1, 2, \ldots, n, \tag{3}$$

where $\overline{a}_j, \overline{b}_j$ refer to nominal values of system parameters, $\Delta a_j, \Delta b_j$ refer to perturbation values of system parameters, and ΔT_{Lj} refers to the variation of load torque of the filling motor. These variables are considered unknown composite disturbances d_j. Hence, formula (3) can be modified as follows:

$$\begin{cases} \dot{\theta}_j = \omega_j, \\ \dot{\omega}_j = -\overline{a}_j\omega_j + \overline{b}_j u_j + d_j, \end{cases} \quad j = 1, 2, \ldots, n, \tag{4}$$

where $d_j = -\Delta a_j\omega_j + \Delta b_j u_j - p_j\Delta T_{Lj}$.

Define the vector form of each state variable as $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_j, \ldots, \theta_n]^T$; $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_j, \ldots, \omega_n]^T$; $\mathbf{u} = [u_1, \ldots, u_j, \ldots, u_n]^T$; and $\mathbf{d} = [d_1, \ldots, d_j, \ldots, d_n]^T$.

Formula (4) is modified in the following matrix form:

$$\begin{cases} \dot{\boldsymbol{\theta}} = \boldsymbol{\omega}, \\ \dot{\boldsymbol{\omega}} = -\mathbf{a}\boldsymbol{\omega} + \mathbf{b}\mathbf{u} + \mathbf{d}, \end{cases} \tag{5}$$

where $\mathbf{a} = \text{diag}\{\overline{a}_1, \ldots, \overline{a}_j, \ldots, \overline{a}_n\}$ and $\mathbf{b} = \text{diag}\{\overline{b}_1, \ldots, \overline{b}_j, \ldots, \overline{b}_n\}$ are coefficient matrices.

4. Sliding Mode Observer Design and Stability Analysis

The composite disturbance d_j of the motor system included in formula (4) will exert a strong influence on the synchronous performance. Therefore, the SMO will be introduced in this section to explore d_j.

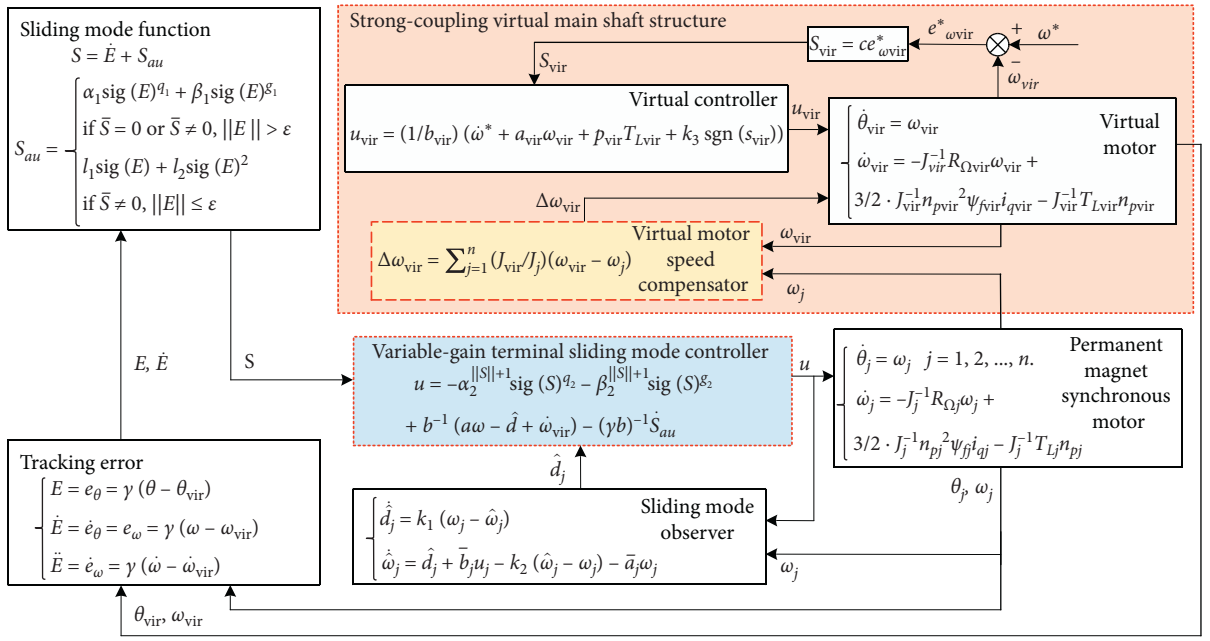


FIGURE 1: The schematic diagram of specific details of each module of this control structure.

The observer is designed for system (4) as follows [36]:

$$\begin{cases} \dot{\hat{\theta}}_j = \hat{\omega}_j + k_1 \mathrm{sgn}\left(\theta_j - \hat{\theta}_j\right), \\ \dot{\hat{\omega}}_j = -\bar{a}_j\omega_j + \bar{b}_j u_j + k_2 \mathrm{sgn}\left(\omega_j - \hat{\omega}_j\right), \\ \qquad\qquad j = 1, 2, \ldots, n, \end{cases} \tag{6}$$

where $\hat{\theta}_j$ and $\hat{\omega}_j$ refer to observed values of rotor position and speed of motor j, respectively, and k_1, k_2 refer to the parameters to be designed.

The observational error vector is defined as $\tilde{\mathbf{e}}_j = [\tilde{\theta}_j\ \tilde{\omega}_j]^T = [\theta_j - \hat{\theta}_j\ \omega_j - \hat{\omega}_j]^T$, and $\tilde{\theta}_j, \tilde{\omega}_j$ refer to observational errors of the rotor position and speed of motor j, respectively.

Hypothesis 1. A normal number $|d_j|_{\max}$ and $d_j \leq |d_j|_{\max}$ exist given that perturbation values of motor parameters $\Delta a_j, \Delta b_j$ and load torque variation ΔT_{Lj} are bounded in practice.

Theorem 1. *Observer (6) corresponding to motor j is designed for filling multimotor system (4). If $k_1 > |\tilde{\omega}_j|_{\max} + \eta_1, k_2 > |d_j|_{\max} + \eta_2$, $(j = 1, 2, \ldots, n)$ is satisfied, then the observational error can converge to 0 in a limited time $T_{\tilde{e}_j}$. The composite disturbance d_j can be expressed as follows:*

$$\lim_{x \longrightarrow T_{\tilde{e}_j}} \hat{d}_j = d_j = k_2 \mathrm{sgn}\left(k_1 \mathrm{sgn}\left(\theta_j - \hat{\theta}_j\right)\right). \tag{7}$$

Proof. The error equation of the observer system can be obtained when formula (6) is subtracted from formula (4) as follows:

$$\begin{cases} \dot{\tilde{\theta}}_j = \tilde{\omega}_j - k_1 \mathrm{sgn}\left(\tilde{\theta}_j\right), \\ \dot{\tilde{\omega}}_j = d_j - k_2 \mathrm{sgn}\left(\tilde{\omega}_j\right). \end{cases} \tag{8}$$

The selected sliding mode surface of the SMO is expressed as follows:

$$\mathbf{s}_{\tilde{e}_j} = \tilde{e}_j. \tag{9}$$

Take the derivative of the above formula and substitute formula (8) into it to get

$$\dot{s}_{\tilde{e}_j} = \dot{\tilde{e}}_j = \begin{bmatrix} \dot{\tilde{\theta}}_j & \dot{\tilde{\omega}}_j \end{bmatrix}^T = \left[\tilde{\omega}_j - k_1 \mathrm{sgn}\left(\tilde{\theta}_j\right)\ \ d_j - k_2 \mathrm{sgn}\left(\tilde{\omega}_j\right)\right]^T. \tag{10}$$

The Lyapunov function of the observer is expressed as follows:

$$V_{\tilde{\mathbf{e}}_j} = \frac{1}{2}\tilde{e}_j^T\tilde{e}_j. \tag{11}$$

The following equation can be derived from formula (11):

$$\begin{aligned} \dot{V}_{\tilde{e}_j} &= \tilde{e}_j^T\dot{\tilde{e}}_j = \tilde{\theta}_j\left[\tilde{\omega}_j - k_1 \mathrm{sgn}\left(\tilde{\theta}_j\right)\right] + \tilde{\omega}_j\left[d_j - k_2 \mathrm{sgn}\left(\tilde{\omega}_j\right)\right] \\ &= \tilde{\theta}_j\tilde{\omega}_j - k_1\left|\tilde{\theta}_j\right| + \tilde{\omega}_j d_j - k_2\left|\tilde{\omega}_j\right| \\ &\leq \left|\tilde{\theta}_j\right|\left(\left|\tilde{\omega}_j\right| - k_1\right) + \left|\tilde{\omega}_j\right|\left(\left|d_j\right| - k_2\right) \\ &\leq -\min\left\{k_1 - \left|\tilde{\omega}_j\right|, k_2 - \left|d_j\right|\right\}\left\|\tilde{e}_j\right\|. \end{aligned} \tag{12}$$

If the designed parameters meet $k_1 > |\tilde{\omega}_j|_{\max} + \eta_1, k_2 > |d_j|_{\max} + \eta_2$ and $\eta_1, \eta_2 > 0$, then the following equation can be obtained:

$$\dot{V}_{\tilde{e}_j} \leq -min\{\eta_1, \eta_2\}\left\|\tilde{e}_j\right\| = -\eta\left\|\tilde{e}_j\right\|, \tag{13}$$

where $\eta = \min\{\eta_1, \eta_2\}$.

Therefore, if $\tilde{\theta}_j \longrightarrow 0, \tilde{\omega}_j \longrightarrow 0$, then $\dot{V}_{\tilde{e}_j} \leq 0$ and $V_{\tilde{e}_j} \longrightarrow 0$. The observer system (6) at this moment demonstrates asymptotic stability. According to Hypothesis 1, the time when the observer system reaches the sliding surface is a limited value $T_{\tilde{e}_j}$. The following conclusions can be obtained from formulas (9) and (10) and the principle [37] of sliding mode equivalence when $t > T_{\tilde{e}_j}$:

$$\begin{cases} \mathbf{s}_{\tilde{e}_j} = 0 \\ \dot{s}_{\tilde{e}_j} = 0 \end{cases} \Rightarrow \begin{cases} \theta_j = \hat{\theta}_j, \\ \tilde{\omega}_j = k_1 \mathrm{sgn}\left(\tilde{\theta}_j\right) = k_1 \mathrm{sgn}\left(\theta_j - \hat{\theta}_j\right), \end{cases} \begin{cases} \omega_j = \hat{\omega}_j, \\ d_j = k_2 \mathrm{sgn}\left(\tilde{\omega}_j\right) = k_2 \mathrm{sgn}\left(\omega_j - \hat{\omega}_j\right). \end{cases} \tag{14}$$

The unknown composite disturbance d_j can then be restructured as follows:

$$\lim_{x \longrightarrow T_{\tilde{e}_j}} \hat{d}_j = d_j = k_2 \mathrm{sgn}\left(k_1 \mathrm{sgn}\left(\theta_j - \hat{\theta}_j\right)\right). \tag{15}$$

If $t > t_1$ when $\forall \delta > 0, \exists t_1 < T_{\tilde{e}_j}$, then

$$\left|\hat{d}_j - d_j\right| < \delta. \tag{16}$$

Hence, the stability of the SMO has been proven. □

5. System Controller Design and Stability Analysis

Definition 1 (see [38]). $\mathrm{diag}(\mathbf{x})$ refers to the diagonal matrix of Vector $\mathbf{x}$ for any vector $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T \in \mathbf{R}^n$ and expressed as follows:

$$\begin{cases} \mathbf{x}^a = \left[x_1^a, x_2^a, \ldots, x_n^a\right]^T, \\ |\mathbf{x}|^p = \left[|x_1|^a, |x_2|^a, \ldots, |x_n|^a\right]^T, \\ \mathrm{sgn}(\mathbf{x}) = \left[\mathrm{sgn}(x_1), \mathrm{sgn}(x_2), \ldots, \mathrm{sgn}(x_n)\right]^T, \\ \mathrm{sig}(\mathbf{x})^a = \left[\mathrm{sgn}(x_1)|x_1|^a, \mathrm{sgn}(x_2)|x_2|^a, \ldots, \mathrm{sgn}(x_n)|x_n|^a\right]^T, \end{cases} \tag{17}$$

where $\mathrm{sgn}(\cdot)$ represents the standard symbolic function and a refers to an arbitrary real number.

Lemma 1 (see [39]). *Stability theorem of fixed-time) For the system is expressed as follows:*

$$\dot{x}(t) = f(x(t)), f(0) = 0, t_0 = 0, x_0 \triangleq x(0), \tag{18}$$

where $x(t) \in R^n$ represents the state of the system, $f: U \longrightarrow R^n$ refers to a continuous function from the domain containing the original point U to n_dimensional space R^n, $0 \in \mathbf{R}^n$ represents the zero vector, and x_0 is defined as the initial state.

If a positive definite and continuous function $V(x)$ exists in the domain U, then

$$\dot{V}(x) \leq -\alpha V(x)^p - \beta V(x)^g, \quad x \in U \backslash \{0\}, \tag{19}$$

where α, β, γ, p, and g are normal numbers that meet $p < 1$ and $g > 1$, respectively, and the system is stable in the fixed time. Hence, the system state can converge to the equilibrium point in a fixed time independent of the initial value.

$$T \leq \frac{1}{\alpha(1-p)} + \frac{1}{\beta(g-1)}. \tag{20}$$

5.1. Design and Stability Analysis of the Virtual Sliding Mode Controller. The virtual controller u_{vir} can be used to control the velocity of the virtual motor in this study given that the upper bound of the virtual motor speed compensator signal is known. The tracking error $e^*_{\omega \mathrm{vir}}$ of the speed of the virtual motor is expressed as follows:

$$e^*_{\omega \mathrm{vir}} = \omega^* - \omega_{\mathrm{vir}}. \tag{21}$$

The following design of the virtual motor speed compensator module is added:

$$\Delta\omega_{\mathrm{vir}} = \sum_{j=1}^{n} \frac{J_{\mathrm{vir}}}{J_j}\left(\omega_{\mathrm{vir}} - \omega_j\right). \tag{22}$$

This signal can reduce the compensation by appropriately increasing the inertia J_j of the virtual motor to improve the convenience in parameter selection of the virtual controller. Meanwhile, $\Delta\omega_{\mathrm{vir}}$ is also a bounded variable, that is, $\Delta\omega_{\mathrm{vir}} \leq |\Delta\omega_{\mathrm{vir}}|_{\max}$, because ω_{vir} and ω_j are both bounded variables.

The second-order system model of the improved virtual main-axis structure can be obtained from formulas (2) and (22) as follows:

$$\begin{cases} \dot{\theta}_{\mathrm{vir}} = \omega_{\mathrm{vir}}, \\ \dot{\omega}_{\mathrm{vir}} = -a_{\mathrm{vir}}\omega_{\mathrm{vir}} + b_{\mathrm{vir}}u_{\mathrm{vir}} - p_{\mathrm{vir}}T_{L\mathrm{vir}} - \Delta\omega_{\mathrm{vir}}. \end{cases} \tag{23}$$

The following sliding mode surface of the virtual controller is selected:

$$s_{\mathrm{vir}} = ce^*_{\omega \mathrm{vir}}, \tag{24}$$

where c refers to the normal number to be designed.

The virtual controller is designed as follows:

$$u_{\mathrm{vir}} = \frac{1}{b_{\mathrm{vir}}}\left(\dot{\omega}^* + a_{\mathrm{vir}}\omega_{\mathrm{vir}} + p_{\mathrm{vir}}T_{L\mathrm{vir}} + k_3 \mathrm{sgn}\left(s_{\mathrm{vir}}\right)\right). \tag{25}$$

Theorem 2. *Virtual controller (2) is shown in formula (25). If the coefficient of the control can meet $k_3 > |\Delta\omega_{vir}|_{\max} + \eta_3$, then $e^*_{\omega vir}$ will gradually converge to 0.*

Proof. Define the Lyapunov function of the virtual controller as

$$V_{\text{vir}} = \frac{1}{2}s_{\text{vir}}^2. \tag{26}$$

The following equation can be derived from formula (26):

$$\dot{V}_{\text{vir}} = s_{\text{vir}}\dot{s}_{\text{vir}} = s_{\text{vir}}c\dot{e}^*_{\text{vir}} = s_{\text{vir}}c\left(\dot{\omega}^* + a_{\text{vir}}\omega_{\text{vir}} - b_{\text{vir}}u_{\text{vir}} + p_{\text{vir}}T_{L\text{vir}} + \Delta\omega_{\text{vir}}\right). \tag{27}$$

The following equation can be obtained by substituting formula (25) into formula (27):

$$\dot{V}_{\text{vir}} = s_{\text{vir}}c\left(-k_3\text{sgn}\left(s_{\text{vir}}\right) + \Delta\omega_{\text{vir}}\right) = -c\left(k_3\left|s_{\text{vir}}\right| - s_{\text{vir}}\Delta\omega_{\text{vir}}\right). \tag{28}$$

The following can be obtained when η_3 is any small normal number that only guarantees $k_3 > |\Delta\omega_{\text{vir}}|_{\max} + \eta_3$:

$$\dot{V}_{\text{vir}} < 0. \tag{29}$$

Formula (29) and the equivalent principle of sliding mode [37] showed that $e^*_{\omega\text{vir}}$ converges to 0 when $\omega^* = \omega_{\text{vir}}$.

Hence, the stability of the virtual motor controller has been proven. □

5.2. Design and Stability Analysis of the Filling Motor Controller. The tracking error of each slave motor is defined as follows:

$$\mathbf{e_\theta} = \boldsymbol{\gamma}\left(\boldsymbol{\theta} - \boldsymbol{\theta}_{\mathbf{vir}}\right) = \left[e_{\theta_1}, \ldots, e_{\theta_j}, \ldots, e_{\theta_n}\right]^T_{1\times n} = \left[\gamma_{1,\text{vir}}\left(\theta_1 - \theta_{\text{vir}}\right), \ldots, \gamma_{\text{j,vir}}\left(\theta_j - \theta_{\text{vir}}\right), \ldots, \gamma_{\text{n,vir}}\left(\theta_n - \theta_{\text{vir}}\right)\right]^T_{1\times n}, \tag{30}$$

$$\mathbf{e_\omega} = \dot{\mathbf{e}}_{\boldsymbol{\theta}} = \boldsymbol{\gamma}\left(\omega - \omega_{\mathbf{vir}}\right) = \left[e_{\omega_1}, \ldots, e_{\omega_j}, \ldots, e_{\omega_n}\right]^T_{1\times n} = \left[\gamma_{1,\text{vir}}\left(\omega_1 - \omega_{\text{vir}}\right), \ldots, \gamma_{j,\text{vir}}\left(\omega_j - \omega_{\text{vir}}\right), \ldots, \gamma_{n,\text{vir}}\left(\omega_n - \omega_{\text{vir}}\right)\right]^T_{1\times n}, \tag{31}$$

where $\gamma_{1,\text{vir}} = J_1/J_{\text{vir}}, \ldots, \gamma_{j,\text{vir}} = J_j/J_{\text{vir}}, \ldots, \gamma_{n,\text{vir}} = J_n/J_{\text{vir}}$; $\boldsymbol{\gamma} = \text{diag}\left\{\gamma_{1,\text{vir}}, \cdots, \gamma_{j,\text{vir}}, \cdots, \gamma_{n,\text{vir}}\right\}$; $\boldsymbol{\theta}_{\mathbf{vir}} = [\theta_{\text{vir}}, \ldots, \theta_{\text{vir}}, \ldots, \theta_{\text{vir}}]^T_{1\times n}$; and $\boldsymbol{\omega}_{\mathbf{vir}} = [\omega_{\text{vir}}, \ldots, \omega_{\text{vir}}, \ldots, \omega_{\text{vir}}]^T_{1\times n}$.

The simplification of this error equation obtained from formulas (30) and (31) can be expressed as follows:

$$\begin{cases} \mathbf{E} = \mathbf{e_\theta} = \boldsymbol{\gamma}\left(\boldsymbol{\theta} - \boldsymbol{\theta}_{\mathbf{vir}}\right), \\ \dot{\mathbf{E}} = \dot{\mathbf{e}}_{\boldsymbol{\theta}} = \mathbf{e}_\omega = \boldsymbol{\gamma}\left(\boldsymbol{\omega} - \boldsymbol{\omega}_{\mathbf{vir}}\right), \\ \ddot{\mathbf{E}} = \dot{\mathbf{e}}_\omega = \boldsymbol{\gamma}\left(\dot{\boldsymbol{\omega}} - \dot{\boldsymbol{\omega}}_{\mathbf{vir}}\right). \end{cases} \tag{32}$$

The selected terminal sliding surface is expressed as follows [38, 40]:

$$\mathbf{S} = \dot{E} + \mathbf{S}_{\text{au}}, \tag{33}$$

where

$$\mathbf{S_{au}} = \begin{cases} \alpha_1\text{sig}\left(\mathbf{E}\right)^{q_1} + \beta_1\text{sig}\left(\mathbf{E}\right)^{g_1}, & \text{if } \overline{\mathbf{S}} = \mathbf{0} \text{ or } \overline{\mathbf{S}} \neq \mathbf{0}, \|\mathbf{E}\| > \varepsilon, \\ l_1\text{sig}\left(\mathbf{E}\right) + l_2\text{sig}\left(\mathbf{E}\right)^2, & \text{if } \overline{\mathbf{S}} \neq \mathbf{0}, \|\mathbf{E}\| \leq \varepsilon, \end{cases} \tag{34}$$

where $\overline{\mathbf{S}} = \dot{\mathbf{E}} + \alpha_1\text{sig}\left(\mathbf{E}\right)^{q_1} + \beta_1\text{sig}\left(\mathbf{E}\right)^{g_1}$. α_1, β_1 refer to the coefficients to be designed that meet $\alpha_1 > 1, \beta_1 > 1$. q_1 and g_1 satisfy $q_1 \in (0.5, 1)$ and $g_1 > 1$, respectively. ε refers to the switching threshold. $\mathbf{S_{au}}$ and $\dot{S}_{\mathbf{au}}$ are continuous functions in $\|\mathbf{E}\| = \varepsilon$. $l_1 = 0.5\alpha_1\varepsilon^{q_1-1} + 0.5\beta_1\varepsilon^{g_1-1}$; $l_2 = 0.5\alpha_1\varepsilon^{q_1-2} + 0.5\beta_1\varepsilon^{g_1-2}$.

This switching rule (34) is inspired by [41, 42] to avoid the singularity.

The optimized variable gain terminal SMC is designed as follows:

$$\mathbf{u} = -\alpha_2^{\|\mathbf{S}\|+1}\text{sig}\left(\mathbf{S}\right)^{q_2} - \beta_2^{\|\mathbf{S}\|+1}\text{sig}\left(\mathbf{S}\right)^{g_2} + \mathbf{b}^{-1}\left(\mathbf{a}\boldsymbol{\omega} - \widehat{\mathbf{d}} + \dot{\boldsymbol{\omega}}_{\text{vir}}\right) - \left(\boldsymbol{\gamma}\mathbf{b}\right)^{-1}\dot{\mathbf{S}}_{\mathbf{au}}, \tag{35}$$

where α_2, β_2 refer to coefficients to be designed that meet $\alpha_2 > 1, \beta_2 > 1$, respectively, and q_2 and g_2 satisfy $q_2 \in (0.5, 1)$ and $g_2 > 1$, respectively.

Theorem 3. *The system controller is designed in formula (35). If the coefficient of the controller satisfies $\alpha_2 > 1, \beta_2 > 1$, $q_2 \in (0.5, 1)$, and $g_2 > 1$, then the system error (32) in formula*

(5) will converge to 0 within the fixed-time $T_2 + T_3$. The upper bound on this time of convergence is

$$\begin{cases} T_2 \le \dfrac{1}{\alpha_2^{\|S\|+1}\lambda_{\min}[(\mathbf{I}+\boldsymbol{\gamma})\mathbf{b}]\left(1-\theta_{q1}\right)(1-q_2)} + \dfrac{1}{2^{1-g_2}\beta_2^{\|S\|+1}\lambda_{\min}[(\mathbf{I}+\boldsymbol{\gamma})\mathbf{b}]\left(1-\theta_{g1}\right)(g_2-1)}, \\ T_3 \le \dfrac{1}{\alpha_1\left(1-\theta_{q_2}\right)(1-q_1)} + \dfrac{1}{2^{1-g_1}\beta_1\left(1-\theta_{g_2}\right)(g_1-1)}. \end{cases} \tag{36}$$

Proof. Define the Lyapunov function of the optimized variable gain terminal SMC as

$$V_{\mathbf{S}} = \mathbf{S}^{\mathbf{T}}\mathbf{S}. \tag{37}$$

Substituting formulas (32) and (33) into formula (37) gives the following:

$$\dot{V}_{\mathbf{S}} = 2\mathbf{S}^{\mathbf{T}}\dot{S} = 2\mathbf{S}^{\mathbf{T}}\left[\boldsymbol{\gamma}\left(\dot{\omega}-\dot{\omega}_{\mathbf{vir}}\right)+\dot{S}_{\mathrm{au}}\right]. \tag{38}$$

Substituting formulas (5) and (35) into formula (38) gives the following:

$$\begin{aligned} \dot{V}_S &= 2\mathbf{S}^T\left[-\alpha_2^{\|S\|+1}\gamma b\,\mathrm{sig}(\mathbf{S})^{q_2} - \beta_2^{\|S\|+1}\gamma b\,\mathrm{sig}(\mathbf{S})^{g_2}\right] \\ &\le -2\alpha_2^{\|S\|+1}\lambda_{\min}(\boldsymbol{\gamma}\mathbf{b})\left[\left(|s_1|^2\right)^{q_2+1/2} + \left(|s_2|^2\right)^{q_2+1/2} + \cdots + \left(|s_n|^2\right)^{q_2+1/2}\right] \\ &\quad -2\beta_2^{\|S\|+1}\lambda_{\min}(\boldsymbol{\gamma}\mathbf{b})\left[\left(|s_1|^2\right)^{g_2+1/2} + \left(|s_2|^2\right)^{g_2+1/2} + \cdots + \left(|s_n|^2\right)^{g_2+1/2}\right]. \end{aligned} \tag{39}$$

The following can be obtained on the basis of the conclusion in [43]:

$$\dot{V}_{\mathbf{S}} \le -2\alpha_2^{\|\mathbf{S}\|+1}\lambda_{\min}(\boldsymbol{\gamma}\mathbf{b})V_{\mathbf{S}}^{q_2+1/2} - 2^{2-g_2}\beta_2^{\|\mathbf{S}\|+1}\lambda_{\min}(\boldsymbol{\gamma}\mathbf{b})V_{\mathbf{S}}^{g_2+1/2}. \tag{40}$$

According to Lemma 1, formula (40) satisfies the fixed-time convergence condition, and its time upper bound is expressed as

$$T_2 \le \frac{1}{\alpha_2^{\|\mathbf{S}\|+1}\lambda_{\min}[(\mathbf{I}+\boldsymbol{\gamma})\mathbf{b}]\left(1-\theta_{q1}\right)(1-q_2)} + \frac{1}{2^{1-g_2}\beta_2^{\|\mathbf{S}\|+1}\lambda_{\min}[(\mathbf{I}+\boldsymbol{\gamma})\mathbf{b}]\left(1-\theta_{g1}\right)(g_2-1)}. \tag{41}$$

□

Remark 1. Formula (41) shows that nonlinear coefficients $\alpha_1^{\|\mathbf{S}\|+1}$ and $\beta_1^{\|\mathbf{S}\|+1}$ satisfy the conditions of fixed-time convergence in the convergence of $\mathbf{S}$.

Assuming that $\mathbf{S} = \varphi$ exists at the moment T_φ within T_2, formula (41) shows the following upper bound of the time when the sliding mode movement moves from an arbitrary initial position to the sliding surface:

$$T_{\boldsymbol{\varphi}} = \frac{1}{\alpha_2^{\|\boldsymbol{\varphi}\|+1}\lambda_{\min}[(\mathbf{I}+\boldsymbol{\gamma})\mathbf{b}]\left(1-\theta_{q1}\right)(1-q_2)} + \frac{1}{2^{1-g_2}\beta_2^{\|\boldsymbol{\varphi}\|+1}\lambda_{\min}[(\mathbf{I}+\boldsymbol{\gamma})\mathbf{b}]\left(1-\theta_{g1}\right)(g_2-1)}. \tag{42}$$

The comparison of formula (42) with the method proposed in [43] shows that

$$T_{\boldsymbol{\varphi}} < \frac{1}{\alpha_2 \lambda_{\min}[(\mathbf{I}+\boldsymbol{\gamma})\mathbf{b}]\left(1-\theta_{q1}\right)\left(1-q_2\right)} + \frac{1}{2^{1-g_2}\beta_2 \lambda_{\min}[(\mathbf{I}+\boldsymbol{\gamma})\mathbf{b}]\left(1-\theta_{g1}\right)\left(g_2-1\right)}. \tag{43}$$

Remark 2. The comprehensive analysis of formulas (41) and (42) showed that the large gain coefficient of the sig function in the controller accelerates the speed to the sliding surface when the sliding mode movement moves away from $\mathbf{S} = 0$. The gain coefficient of the sig function will continuously approach α_1 and β_1 when the SM movement is close to $\mathbf{S} = 0$. Therefore, the actual convergence time will be shortened. In the process of sliding mode movement from the initial state to the sliding surface, the optimized SMC has better convergence performance than the original terminal SMC.

The following can be obtained from the above conclusion and formula (33) when the sliding mode movement reaches the sliding surface:

$$\dot{E} = -\mathbf{S}_{\text{au}}. \tag{44}$$

In the process that the system moves to the equilibrium point on the sliding surface, based on sliding mode substitution method designed in formula (34), the classification should be discussed here.

When $\|\mathbf{E}\| > \varepsilon$, we can get

$$\dot{E} = -\alpha_1 \text{sig}(\mathbf{E})^{q_1} - \beta_1 \text{sig}(\mathbf{E})^{g_1}. \tag{45}$$

The Lyapunov function is defined, and the system error is analysed as follows:

$$V_{\mathbf{E}} = \mathbf{E}^{\mathbf{T}}\mathbf{E}. \tag{46}$$

The following can be obtained on the basis of formula (46):

$$\dot{V}_{\mathbf{E}} = 2\mathbf{E}^{\mathbf{T}}\dot{E}. \tag{47}$$

The following can be obtained after substituting formula (45) into formula (47):

$$\dot{V}_{\mathbf{E}} = 2\mathbf{E}^{\mathbf{T}}\left[-\alpha_1 \text{sig}(\mathbf{E})^{q_1} - \beta_1 \text{sig}(\mathbf{E})^{g_1}\right] \leq -2\alpha_1 V_{\mathbf{E}}^{(q_1+1)/2} - 2^{2-g_1}\beta_1 V_{\mathbf{E}}^{(g_1+1)/2}. \tag{48}$$

According to Lemma 1, formula (48) satisfies the fixed-time convergence condition, and the upper bound of time is expressed as follows:

$$T_3 \leq \frac{1}{\alpha_1\left(1-\theta_{q_2}\right)\left(1-q_1\right)} + \frac{1}{2^{1-g_1}\beta_1\left(1-\theta_{g_2}\right)\left(g_1-1\right)}, \tag{49}$$

where θ_{q2} and θ_{g2} refer to arbitrary normal numbers in the interval $(0, 1)$.

Formulas (41) and (49) demonstrate that the system error $\mathbf{E}$ will converge to 0 when the sliding mode moves for $t > T_2 + T_3$.

When $\|\mathbf{E}\| > \varepsilon$, we can get

$$\dot{E} = -l_1 \text{sig}(\mathbf{E}) - l_2 \text{sig}(\mathbf{E})^2. \tag{50}$$

Similarly, the system error is also analysed as follows:

$$\dot{V}_{\mathbf{E}} = 2\mathbf{E}^{\mathbf{T}}\left[-l_1 \text{sig}(\mathbf{E}) - l_2 \text{sig}(\mathbf{E})^2\right] \leq -2l_1 V_{\mathbf{E}}. \tag{51}$$

Hence, formula (51) shows that $V_{\mathbf{E}}$ will converge to 0 exponentially given that l_1 is a positive real number.

Remark 3. The comprehensive analysis on formulas (45)–(51) shows that the convergence time of the linear sliding surface of $\mathbf{S}_{\mathbf{au}} = l_1 \text{sig}(\mathbf{E}) + l_2 \text{sig}(\mathbf{E})^2$ instead of the terminal sliding mode can be ignored if the value of ε is sufficiently small. The convergence time is approximately equal to T_2 in this case. Therefore, system errors $\|\mathbf{E}\|$ and $\|\dot{E}\|$ can converge to 0 within $T_2 + T_3$; that is, the velocity of all motors in the multimotor system can achieve $\omega = \omega_{\mathbf{vir}}$ in a fixed time. Moreover, filling motors can track the given value in real time because the velocity of the virtual motor is $\omega_{\text{vir}} = \omega^*$. Hence, the stability of the filling motor controller has been proven.

The overall control block of the improved virtual main-axis system is shown in Figure 2 to understand the content of Figure 1 and design details of each module further.

Figure 2 shows that the most important characteristic of the control structure is the addition of the speed compensator to the virtual motor and the use of the virtual controller to provide feedback and adjust the tracking error and compensation signal of the virtual motor. Meanwhile, a variable-gain terminal SMC is designed to realise the high-speed and high-precision tracking of each filling motor speed to the given signal.

6. Simulation

The simulation of the multimotor system composed of three motors in this section is conducted using MATLAB/Simulink to verify the proposed technical solution on the basis of theoretical studies. The comparative experimental simulation is added in this section to reflect the advantages

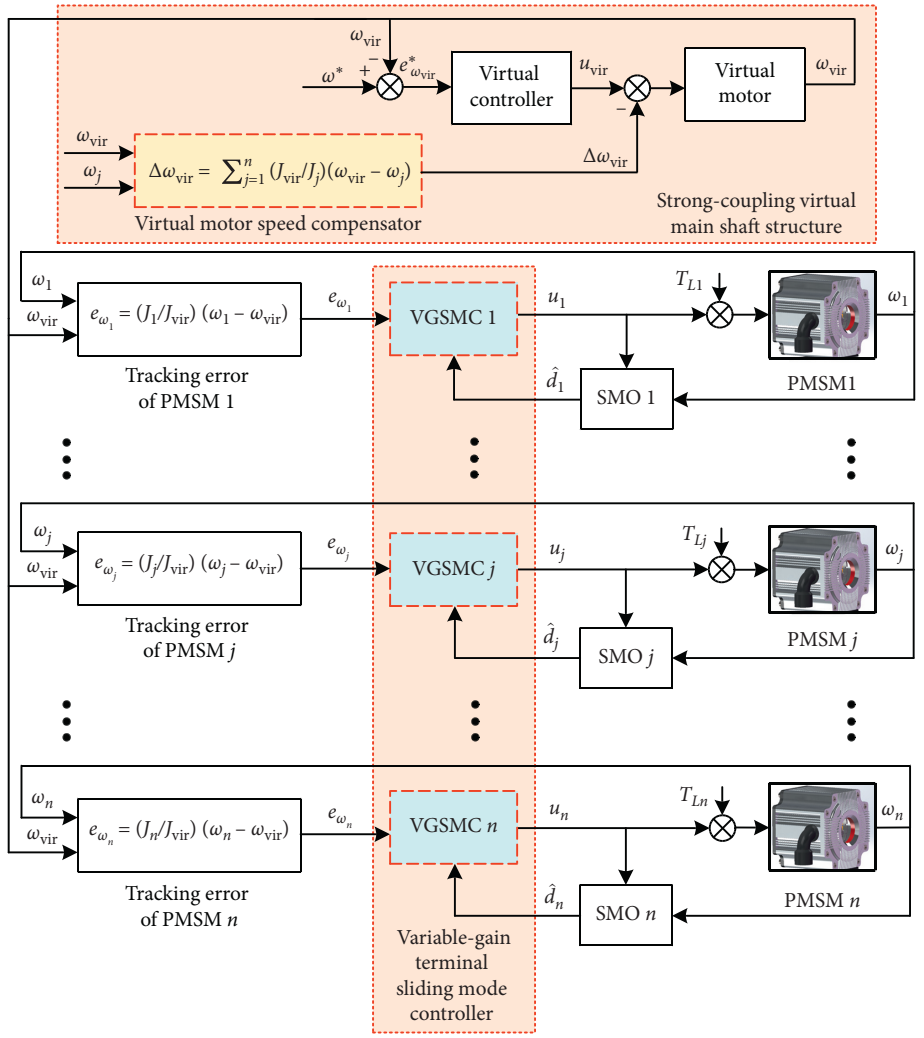


Figure 2: The control block diagram of the improved virtual main shaft structure.

of the improved synchronization topology and the optimised fixed-time controller. The terminal SMC proposed in [43] is used as the reference and modified according to the PMSM model in this study, but the control concept is consistent with the design idea of quotations. The given curve for multimotor system simulation is as follows:

$$\omega^* = \begin{cases} 500t, & 0 \le t < 0.6, \\ 300, & 0.6 \le t < 1.4, \\ -500t + 1000, & 1.4 \le t < 2. \end{cases} \quad (52)$$

Some typical motor parameters are shown in Table 1.

The simulation time in the entire experiment is 2 s. The experimental simulation of the improved virtual main-axis multimotor synchronous control structure is carried out for the optimised variable-gain terminal SMC based on fixed-time convergence and the traditional terminal SMC. The tracking error of each motor is compared with the simulation result of the synchronisation error under the enabling conditions of two controllers. Parameters of the observer are set to $k_1 = 4000$ and $k_2 = 10$. Virtual moment of inertia in the virtual motor is set to $J_{vir} = 0.0001$. The gain of controller

Table 1: Some typical motor parameters.

Parameter	PMSM 1	PMSM 2	PMSM 3
R_Ω (N · m · s)	4.831×10^{-5}	4.846×10^{-5}	4.827×10^{-5}
ψ_f (Wb)	0.175	0.173	0.178
n_p	2	2	2
J (kg · m^2)	8.5×10^{-4}	8.47×10^{-4}	8.51×10^{-4}

in the virtual controller is set to $k_3 = 200$ and $c = 50$. Coefficients of the sliding mode surface in the trial design of the variable-gain terminal SMC are set to $\alpha_1 = \beta_1 = 2.1$, $q_1 = 0.833$, and $g_1 = 2$. Coefficients of the controller are set to $\alpha_2 = \beta_2 = 2.1$, $q_2 = 0.833$, and $g_2 = 2$. The consistency between coefficients of the controller in the designed comparative experiment simulation in [43] and parameters of the observer ensures the validity and preciseness of the comparative experiment. The parameter perturbation value and motor torque variation of motors 1, 2, and 3 are $\Delta a_1 = \Delta b_1 = \Delta T_{L1} = \sin \pi t$, $\Delta a_2 = \Delta b_2 = \Delta T_{L2} = \sin 0.5\, \pi t$, and $\Delta a_3 = \Delta b_3 = \Delta T_{L3} = -\sin \pi t$, respectively. The following will show the simulation results of this experiment in turn.

As shown in Figure 3, motors of the filling multimotor system can successfully track the given value. The designed

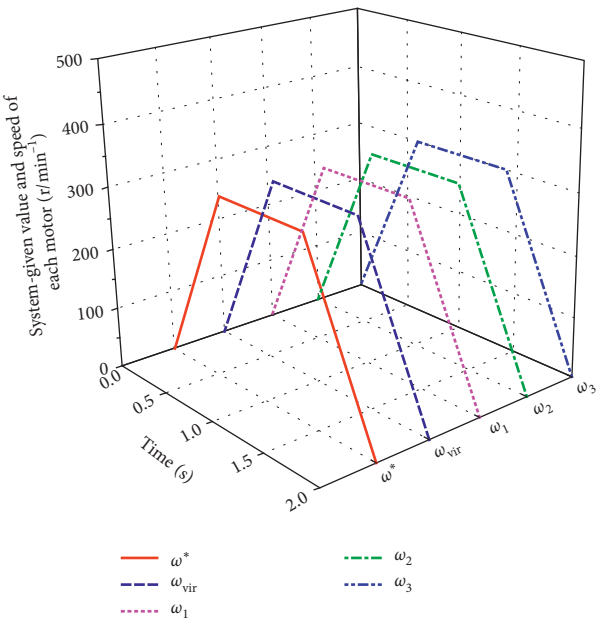


Figure 3: Schematic diagram of speed simulation results of multimotor system.

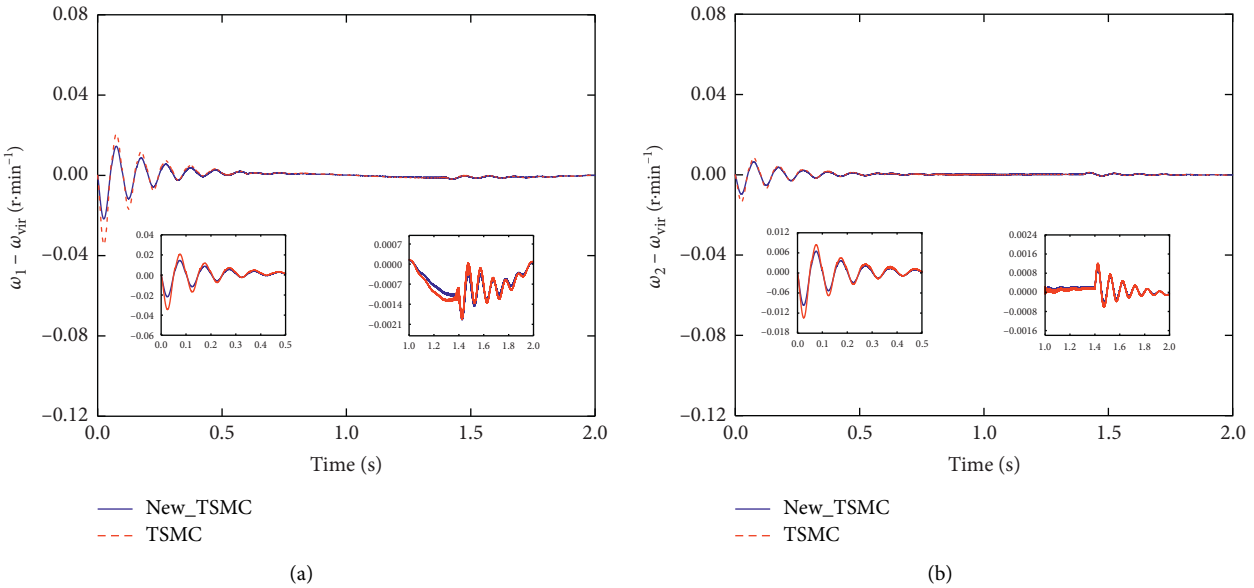
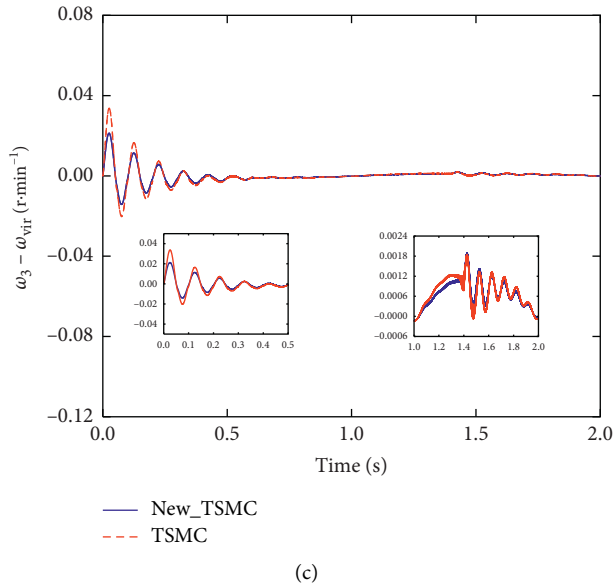


Figure 4: Continued.



(c)

Figure 4: Simulation comparison of tracking error between (a) motor 1 and virtual motor, (b) motor 2 and virtual motor, and (c) motor 3 and virtual motor.

three-stage given signal includes the following processes: (1) starting and accelerating, (2) running at a uniform velocity, and (3) slowing down to stop. The experimental results showed that virtual and slave motors can rapidly track the given value in the three stages and run stably without evident overshooting. The actual simulation effect of the virtual motor can verify that the selection of the virtual motor's moment of inertia and the design of the virtual controller are reasonable.

Simulation comparison charts in Figures 4(a)–4(c) that present the tracking error of the velocity of motors 1, 2, and 3, respectively, are further representations of the simulation results in Figure 3. New_TSMC and TSMC represent the simulation results of the tracking error between each slave motor and virtual motor under the enabling conditions of the improved variable-gain terminal SMC and controller in [43], respectively. The tracking error of motors 1, 2, and 3 refers to the error between the actual velocity of each motor and the velocity of the virtual motor. The analysis of the simulation results showed that the tracking performance of the improved variable-gain terminal SMC is significantly improved compared with that of the traditional terminal SMC. Both the amplitude of overshooting and the convergence time of all slave motors are reduced by approximately one-third, especially in the initial stage of the simulation.

Figure 5(a) represents the synchronization error between motor 1 and motor 2. Figure 5(b) represents the synchronization error between motor 1 and motor 3. Figure 5(c) represents the synchronization error between motor 2 and motor 3. New_TSMC and TSMC, respectively, represent the simulation results of the synchronisation error between adjacent motors under the enabling condition of the variable-gain terminal SMC and the controller enable condition proposed in [43]. The simulation results showed that the synchronisation performance of the improved controller significantly enhances compared with that of the traditional terminal SMC. Convergence amplitudes of all slave motors are reduced significantly, especially in the initial stage of the simulation.

Figure 6 shows that the tracking error between the real-time speed of the virtual motor and the given value is constantly controlled in a small range and vibrates in a small amplitude of 0.4–1.4 s during start-up, respectively. This phenomenon is mainly caused by the high change rate of the given value signal that causes fluctuation in the velocity of the virtual motor whilst the tracking error of the virtual motor can gradually converge within 0 to 0.001 s. Hence, the convergence is fast and the parameter selection for the virtual motor and the virtual controller is reasonable.

As shown in Figure 7, the velocity compensation signal constantly maintained in a small range is conducive to the design of the gain coefficient of the virtual SMC. This signal is mainly affected by the velocity of all slave motors. Therefore, the velocity compensator signal can only generate a small amplitude of vibration after selecting a small moment of inertia of the virtual motor. The velocity compensation signal gradually converges to zero when the velocity of each motor follows the given value. The result can be used to verify the effectiveness of the design of the velocity compensation signal.

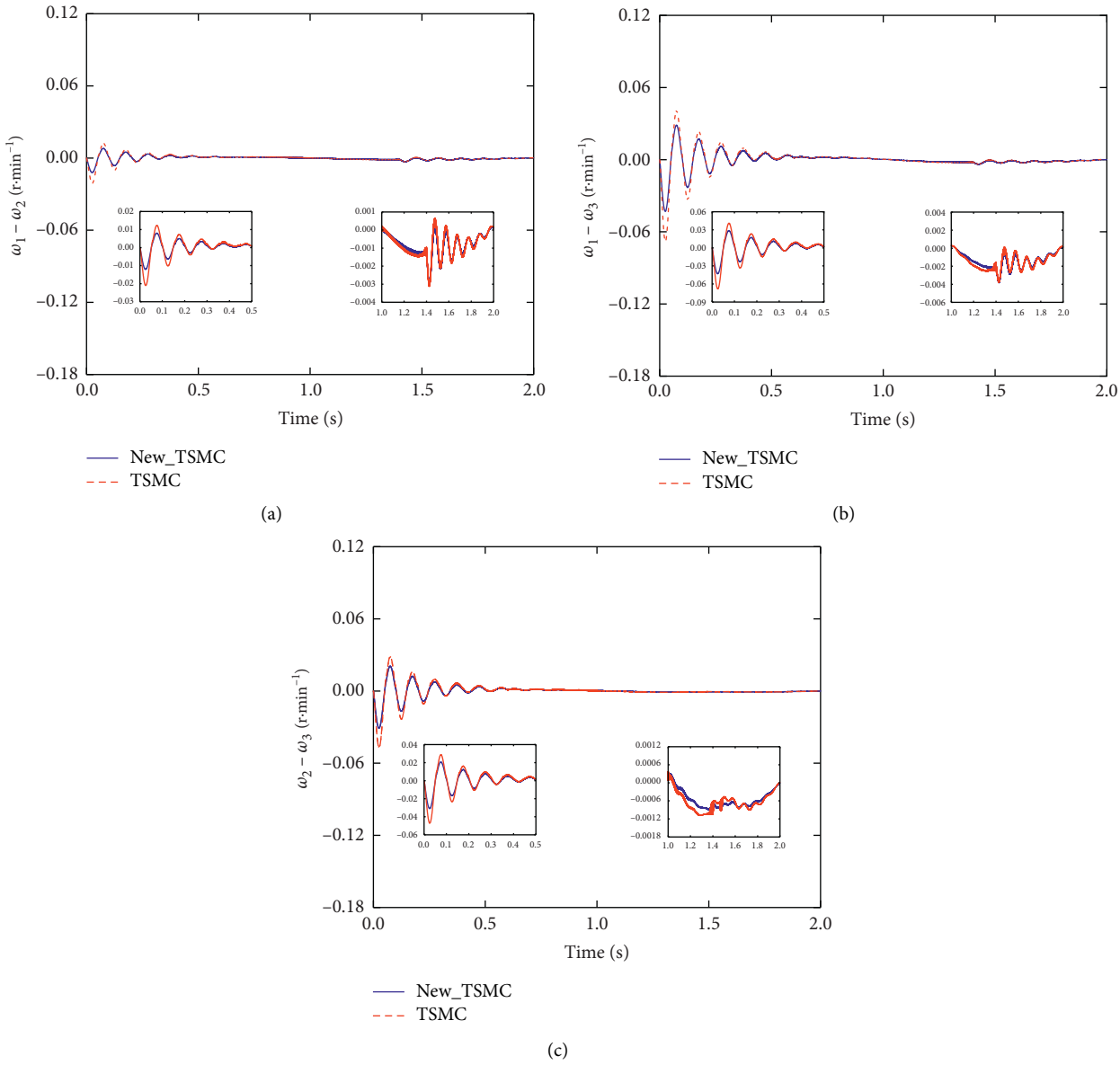


Figure 5: Comparison of simulation diagram of synchronization errors between (a) motors 1 and 2, (b) motors 1 and 3, and (c) motors 2 and 3.

Figures 8(a)–8(c) show the simulation results of disturbances and observed values of motors 1, 2, and 3, respectively. The effect of the estimation of the sliding mode observer of each motor that tracks the actual disturbance is better for different unknown composite disturbances of the three slave motors. The observed value of each motor will evidently jitter in the initial stage of the simulation but can be controlled to follow the actual value of the disturbance within 0.2 s, which exerts a minimal impact on the entire system. This result is consistent with the expectation of the simulation experiment.

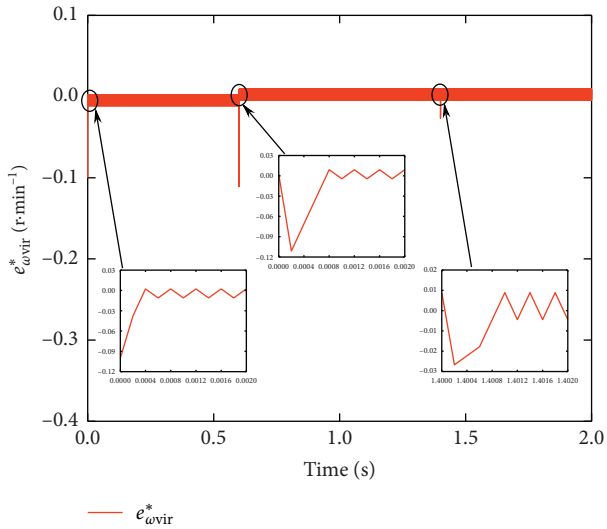


Figure 6: Simulation results of virtual motor speed tracking error.

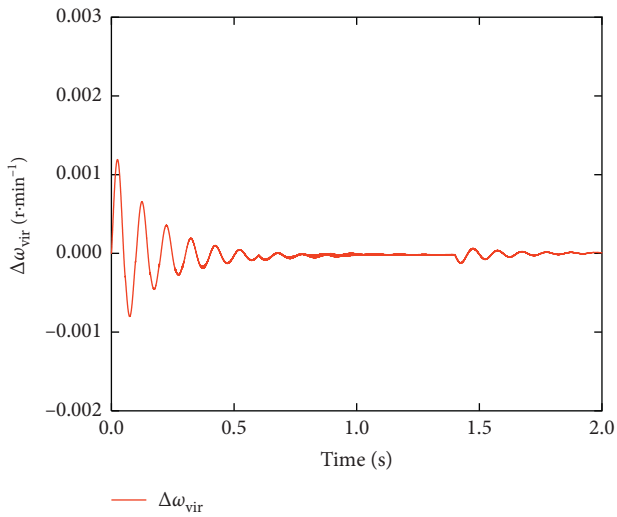
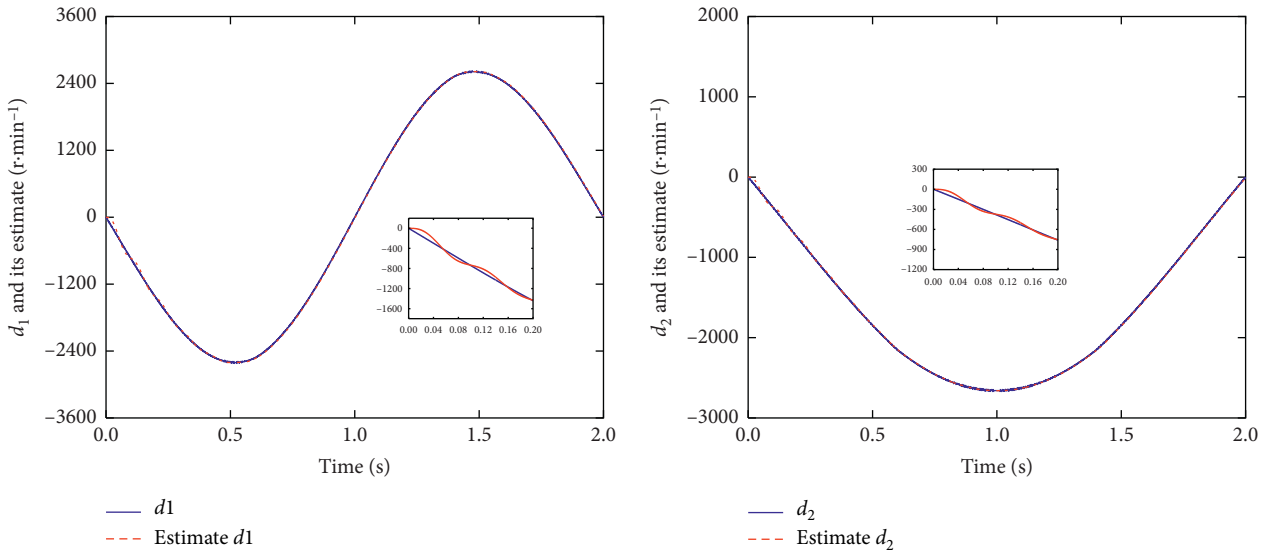
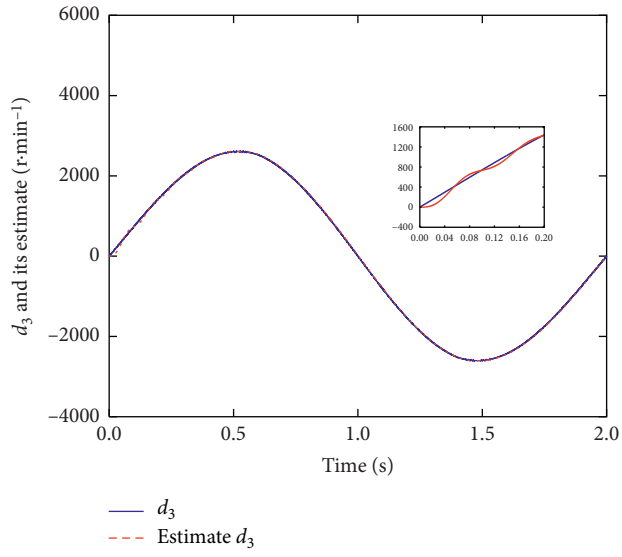


Figure 7: Speed compensator signal simulation results.



(a) (b)

Figure 8: Continued.



(c)

Figure 8: The disturbance of (a) motor 1, (b) motor 2, and (c) motor 3 and the simulation results of their observations.

7. Conclusions

A novel synchronous control method for the filling multimotor system is put forward in this study. On the one hand, the velocity compensator added to the virtual motor to establish the coupling relationship between virtual and slave motors in the output velocity improves the correlation between systems on the basis of the existing virtual main-axis control structure. On the other hand, the variable-gain terminal sliding mode controller is designed on the basis of the tracking error between each slave motor and the virtual motor to ensure that the multimotor system can be stable within a fixed time. The proposed method can effectively weaken the adverse effects of the system parameter perturbation and motor load variation and improve the convergence robustness performance of the system. Notably, the convergence of each slave motor in the initial stage of the start-up phase significantly improves. The simulation results verify that the method has a relatively prominent effect on the high-performance cooperative control of filling multimotor system.

Data Availability

The model data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The study was supported by the National Key R&D Program of China (no. 2018YFD0400705).

References

[1] V. Jerković Štil, T. Varga, T. Benšić, and M. Barukčić, "A survey of fuzzy algorithms used in multimotor systems control," *Electronics*, vol. 9, no. 11, p. 1788, 2020.

[2] Z. Gao, S. K. Nguang, and D. Kong, "Advances in modelling, monitoring, and control for complex industrial systems," *Complexity*, vol. 2019, Article ID 2975083, 3 pages, 2019.

[3] H. R. Karimi, M. Zapateiro, and N. Luo, "Adaptive synchronization of master-slave systems with mixed neutral and discrete time-delays and nonlinear perturbations," *Asian Journal of Control*, vol. 14, no. 1, pp. 251–257, 2012.

[4] Y.-J. Liu, L. Liang, T.-T. Chu, and M.-Y. Wu, "N-PD cross-coupling synchronization control based on adjacent coupling error analysis," *Journal of Central South University*, vol. 25, no. 5, pp. 1154–1164, 2018.

[5] M. Wang, X. Ren, and Q. Chen, "Robust tracking and distributed synchronization control of a multi-motor servo-mechanism with H-infinity performance," *ISA Transactions*, vol. 72, pp. 147–160, 2018.

[6] R. Savelsberg, J. Andert, S. Klein, and S. Pischinger, "Virtual shaft: robust coupling by bidirectional and distributed prediction of coupling values," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 234, no. 10-11, pp. 2419–2428, 2020.

[7] S. Yao, G. Gao, Z. Gao, and S. Li, "Active disturbance rejection synchronization control for parallel electro-coating conveyor," *ISA Transactions*, vol. 101, pp. 327–334, 2020.

[8] J. Xu, H. Lu, and X. Liu, "Synchronization control strategy in multi-layer and multi-axis systems based on the combine cross coupling error," *Advances in Mechanical Engineering*, vol. 9, no. 6, 2017.

[9] Z. Huang, Y. Li, G. Song, X. Zhang, and Z. Zhang, "Speed and phase adjacent cross-coupling synchronous control of multi-exciters in vibration system considering material influence," *IEEE Access*, vol. 7, pp. 63204–63216, 2019.

[10] C. Sun, G. Gong, H. Yang, and F. Wang, "Fuzzy sliding mode control for synchronization of multiple induction motors drive," *Transactions of the Institute of Measurement and Control*, vol. 41, no. 11, pp. 3223–3234, 2019.

[11] L. Tao, Q. Chen, Y. Nan, F. Dong, Y. Jin, and Y. Zhou, "Speed tracking and synchronization of a multimotor system based on fuzzy ADRC and enhanced adjacent coupling scheme," *Complexity*, vol. 2018, Article ID 5632939, 16 pages, 2018.

[12] T. Shi, H. Liu, Q. Geng, and C. Xia, "Improved relative coupling control structure for multi-motor speed synchronous driving system," *IET Electric Power Applications*, vol. 10, no. 6, pp. 451–457, 2016.

[13] C. Zhang, M. Xiao, and J. He, "Multimotor improved relative coupling cooperative control based on sliding mode controller," *Mathematical Problems in Engineering*, vol. 2020, Article ID 5638462, 10 pages, 2020.

[14] C. Xu, Y. Zhao, B. Qin, and H. Zhang, "Adaptive synchronization of coupled harmonic oscillators under switching topology," *Journal of the Franklin Institute*, vol. 356, no. 2, pp. 1067–1087, 2018.

[15] C. Zhang, J. He, L. Jia, C. Xu, and Y. Xiao, "Virtual line-shafting control for permanent magnet synchronous motor systems using sliding-mode observer," *IET Control Theory & Applications*, vol. 9, no. 3, pp. 456–464, 2015.

[16] S. Lin, Y. Cai, B. Yang, and W. Zhang, "Electrical line-shafting control for motor speed synchronisation using sliding mode controller and disturbance observer," *IET Control Theory & Applications*, vol. 11, no. 2, pp. 205–212, 2017.

[17] J. He, X. Chen, S. Mao, C. Zhang, and J. Liu, "Virtual line shafting-based total-amount coordinated control of multi-motor traction power," *Journal of Advanced Transportation*, vol. 2020, Article ID 4735397, 9 pages, 2020.

[18] C. Xu, H. Xu, H. Su, and C. Liu, "Disturbance-observer based consensus of linear multi-agent systems with exogenous disturbance under intermittent communication," *Neurocomputing*, vol. 404, pp. 26–33, 2020.

[19] X. Wang and H. Su, "Completely model-free RL-based consensus of continuous-time multi-agent systems," *Applied Mathematics and Computation*, vol. 382, 2020.

[20] S. Hou, Y. Chu, and J. Fei, "Modified fuzzy neural network control using sliding mode technique for power quality improvement system with experimental verification," *IET Control Theory & Applications*, vol. 14, no. 19, pp. 3029–3037, 2020.

[21] J. He, L. Mi, S. Mao, C. Zhang, and H. Chu, "Fault-tolerant control of a nonlinear system actuator fault based on sliding mode control," *Journal of Control Science and Engineering*, vol. 2017, Article ID 8595960, 13 pages, 2017.

[22] C. Xu, B. Li, and L. Yang, "Semi-global containment of discrete-time high-order multi-agent systems with input saturation via intermittent control," *IET Control Theory & Applications*, vol. 14, no. 16, pp. 2303–2309, 2020.

[23] X. Wang, G.-P. Jiang, H. Su, and Z. Zeng, "Consensus-based distributed reduced-order observer design for LTI systems," *IEEE Transactions on Cybernetics*, pp. 1–11, 2020.

[24] X. Wang, X. Wang, H. Su, and L. James, "Coordination control for uncertain networked systems using interval observers," *IEEE Transactions on Cybernetics*, vol. 50, no. 9, pp. 4008–4019, 2019.

[25] Y. Liu and H. Su, "Containment control of second-order multi-agent systems via intermittent sampled position data communication," *Applied Mathematics and Computation*, vol. 362, 2019.

[26] Y. Liu and H. Su, "Some necessary and sufficient conditions for containment of second-order multi-agent systems with sampled position data," *Neurocomputing*, vol. 378, pp. 228–237, 2020.

[27] Y. Zheng, Y. Song, D. J. Hill, and Y. Zhang, "Multiagent system based microgrid energy management via asynchronous consensus ADMM," *IEEE Transactions on Energy Conversion*, vol. 33, no. 2, pp. 886–888, 2018.

[28] M.-C. Fan and M. Wang, "Second-order consensus for a class of uncertain multi-agent systems subject to input saturation," *Transactions of the Institute of Measurement and Control*, vol. 41, no. 7, pp. 1957–1964, 2019.

[29] Y. Zhai, Z. Liu, Z. Guan, and Z. Gao, "Resilient delayed impulsive control for consensus of multiagent networks subject to malicious agents," *IEEE Transactions on Cybernetics*, pp. 1–10, 2020.

[30] B. Liu, H. Su, F. Jiang, Y. Gao, L. Liu, and J. Qian, "Group controllability of continuous-time multi-agent systems," *IET Control Theory & Applications*, vol. 12, no. 11, pp. 1665–1671, 2018.

[31] G. Shen, Y. Xia, J. Zhang, and B. Cui, "Adaptive fixed-time trajectory tracking control for mars entry vehicle," *Nonlinear Dynamics*, vol. 102, no. 4, pp. 2687–2698, 2020.

[32] X. Shao, B. Tian, and W. Yang, "Fixed-time trajectory following for quadrotors via output feedback," *ISA Transactions*, vol. 110, pp. 213–224, 2021.

[33] J. Sun, J. Wang, P. Yang, and S. Guo, "Model-free prescribed performance fixed-time control for wearable exoskeletons," *Applied Mathematical Modelling*, vol. 90, pp. 61–77, 2021.

[34] Y. Zheng, Y. Song, and D. J. Hill, "A general coordinated voltage regulation method in distribution networks with soft open points," *International Journal of Electrical Power and Energy Systems*, vol. 116, 2020.

[35] R. Rahimilarki, Z. Gao, A. Zhang, and R. Binns, "Robust neural network fault estimation approach for nonlinear dynamic systems with applications to wind turbine systems," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 12, pp. 6302–6312, 2019.

[36] C. Zhang, Q. Zhang, J. He, J. Liu, X. Yang, and S. Mao, "Consistent total traction torque-oriented coordinated control of multimotors with input saturation for heavy-haul locomotives," *Journal of Advanced Transportation*, vol. 2020, Article ID 1390764, 11 pages, 2020.

[37] J. He, C. Zhang, S. Mao, H. Wu, and K. Zhao, "Demagnetization fault detection in permanent magnet synchronous motors based on sliding observer," *Journal of Nonlinear Science Application*, vol. 9, no. 5, pp. 2039–2048, 2016.

[38] B. Jiang, Q. Hu, and M. I. Friswell, "Fixed-time attitude control for rigid spacecraft with actuator saturation and faults," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 5, pp. 1892–1898, 2016.

[39] A. Polyakov, "Nonlinear feedback design for fixed-time stabilization of linear control systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 8, pp. 2106–2110, 2012.

[40] Z. Zhu, Y. Xia, and M. Fu, "Attitude stabilization of rigid spacecraft with finite-time convergence," *International Journal of Robust and Nonlinear Control*, vol. 21, no. 6, pp. 686–702, 2011.

[41] K. Lu and Y. Xia, "Adaptive attitude tracking control for rigid spacecraft with finite-time convergence," *Automatica*, vol. 49, no. 12, pp. 3591–3599, 2013.

[42] L. Liangyong Wang, T. Lianfei Zhai, and L. Zhai, "Neural-network-based terminal sliding-mode control of robotic manipulators including actuator dynamics," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 9, pp. 3296–3304, 2009.

[43] B. Jiang, Q. Hu, and M. I. Friswell, "Fixed-time rendezvous control of spacecraft with a tumbling target under loss of actuator effectiveness," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 4, pp. 1576–1586, 2016.

Hindawi
Complexity
Volume 2021, Article ID 4760249, 14 pages
https://doi.org/10.1155/2021/4760249

Research Article

Observer-Based Leader-Following Consensus of General Linear Multiagent Systems Based on Novel Event Trigger Mechanism with Input Time Delay under Directed Graphs

Hong Zhang, Changshun Chen, and Feng Wei

School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430074, China

Correspondence should be addressed to Hong Zhang; sunracer@hust.edu.cn

Received 27 April 2021; Accepted 19 June 2021; Published 2 July 2021

Academic Editor: Xiao Ling Wang

This paper considers the tracking and containment consensus for the general linear systems with input time delays under directed communication networks. The distributed observer-based algorithm on the basis of event-triggering mechanism will be designed by using only neighboring agents information. In this way, we can save network resource effectively. The event-based protocol with input time delays will be proposed for the leader-follower systems. Appropriate feedback gain matrices and trigger parameters can be designed by using Lyapunov stability theory. Based on the designed control algorithm, if the feedback gain matrices and the event trigger are designed appropriately, the leader-follower general linear system can eventually reach tracking and containment consensus. Then, two simulation results are provided to demonstrate the practicability of the theoretical analysis.

1. Introduction

In recent years, the multiagent systems have been paid considerable attentions in various scientific communities, such as spacecraft formation flying, complex network control, and collaborative monitoring [1, 2]. Consensus is one of the important issues that is worth studying in the field of multiagent systems' cooperative control. For the consensus issues of multiagent systems, the main work is to design a suitable information exchange method so that all agents can agree on the state of interest. The research studies on multiagent systems' consensus issues mainly focus on communication topology, agent dynamic systems, and control method, which emphasize the specific analysis of the whole system. And some related results have been obtained. Consensus protocols are proposed in [3–6] for a set of single, double integrators, general linear, and Lipschitz nonlinear systems with undirected and directed network topologies.

The earlier related research studies mainly focus on the leaderless consensus issues, which means the final state of the agents will be not set in advance. However, there might be one or even more leaders in the agent network to realize that the agents reach the state we expected in some practical applications. The leader-follower tracking control issue is an extremely important part of the multiagent systems' consensus issues, which needs the states of all followers follow that of the leader. And in this way, consensus can be achieved more quickly and efficiently. By only using the neighboring information, the tracking control algorithm for the leader-follower single-integrator system with a single leader is designed in [7, 8]. Considering the actual situation, consensus-tracking control strategies are put forward in [9–12] for the leader-follower systems with time delays, noises, external disturbances, and so forth. Under the designed algorithms, control goals can be reached; that is to say, the leader-follower systems can achieve consensus. Considering the situation of multiple leaders, as a special branch of multiagent systems' control methods, containment has been widely used. The main goal of containment control is to ensure that the states of all followers can eventually converge to the convex hull spanned by that of the leaders. Up to now, there are many related outstanding achievements. The distributed control strategy is investigated to solve the containment consensus issues for not only

discrete-time but also continuous-time general linear systems in [13]. The leader-follower systems can achieve containment consensus by this means. In [14], the authors study the distributed containment control issues and design the corresponding control algorithm for the general linear system which can make the leader-follower system reach containment consensus. They designed distributed output-feedback controllers based on the relative information of neighboring agents for the leader-follower systems. The authors in [15] proposed a stochastic sampling control algorithm to solve the issue of formation-containment for the leader-follower linear system. Under this control algorithm, the update frequency of the controller and energy can be reduced. The containment control protocol for heterogeneous linear systems on the basis of neighboring agents' relative information is designed in [16], which will guarantee that the novel system can reach containment consensus. The authors in [17] further extended the containment control method of general linear systems to directed random networks. The fully distributed control algorithm by using the formation-containment control method for heterogeneous linear systems is proposed in [18]. The formation-containment consensus issue is studied by an observer with directed graph. However, these mentioned references do not focus on the research on the update frequency of the controller.

Furthermore, the updates frequency of the controllers can be significantly reduced by using the event-based control method, which may save communication resources of the communication networks. The main idea of the event-triggered control method is that it depends on the predefined trigger condition to update the controller. In order to solve the problem of event-triggered consensus, we need to design the distributed event-based control protocols, which include the event-based control strategies and the triggering functions. In the past decade, the event-based consensus problem has been studied a lot. In [19], a new distributed sampled-data control protocol is designed for fractional-order multiagent systems under directed graph. Event-triggered and self-triggered control protocols are designed for single-integrator agents systems with undirected communication topologies in [20–22]. The event-triggered problems with general linear systems are studied in [23–26]. In [27], the work is extended to the leader-follower systems with event-triggered consensus issues. The distributed control algorithm by using the event-triggered method is designed to let the leader-follower system reach consensus. An adaptive distributed observer-based control strategy by using the event-triggered method is designed in [28] to achieve the predicted control target. The authors in [29] proposed distributed adaptive event-triggered algorithms for both the leaderless and leader-follower linear systems on the basis of the local sampled state information. In [30], the authors designed an event-triggered consensus algorithm for the second-order hybrid systems with continuous-time and discrete-time individuals. In practice, there are many constraints that affect the analysis of the multiagent systems' consistency issues, such as time delays, noises, and external disturbance. Furthermore, time delays are considered into the multiagent systems' consensus issues. In [31], the event-based control law of the leader-follower linear system with input time delay is designed. By algebraic Riccati equation-based method and low-gain output-feedback mechanism, the observer-based protocol is designed to solve the issue of edge-consensus in [32]. The distributed observer-based tracking and containment control algorithms by using the event-triggered control method are, respectively, proposed for general linear systems with time delays in [33, 34]. In [35], a novel distributed tracking control algorithm with input time delays is designed for leader-follower linear systems. A new event trigger is designed to let the multiagent networks reach tracking consensus as well as save network energy. As far as we know, the containment consensus issues with input time delays by using the distributed event-triggered control approach have not been adequately studied, which inspires the current work.

In this paper, we will consider the tracking and containment consensus issues for the leader-follower general linear systems with time delays under directed graph. And the observer-based event-triggered controller will be designed to satisfy control requirements. The main contributions of this work can be summarized as follows. First of all, two fully distributed observer-based control protocols by using the event-triggered method are proposed for both tracking control and containment control problems of the leader-follower linear systems with time delays. Then, appropriate feedback gain matrix and trigger parameters can be obtained by Lyapunov stability theory and matrix analysis. At last, under the proposed protocol, it is shown that the leader-follower linear systems with time delays can eventually achieve tracking and containment consensus.

The remainder of this paper is organized as follows. In Section 2, we give some basic notations and graph theory to be used as well as state the problem to be solved. By using some assumptions and lemmas, the main results are given in Section 3, which contains tracking control and containment control individuals. Section 4 shows the simulation examples to confirm our theoretical analysis. At last, conclusions are drawn in Section 5.

Notations:$R^{n\times n}$ represents the set of $n\times n$ order real matrices. A^T means the transposed matrix of the real matrix A. I_N denotes the identity matrix of order N. $A\otimes B$ denotes the Kronecker product of matrices A and B. λ_m means the maximum eigenvalue of a real matrix. $\mathrm{conv}\{x_1, x_2, \ldots, x_N\} = \left\{\sum_{i=1}^{N} \alpha_i x_i | \alpha_i \in R, \alpha_i \geq 0, \sum_{i=1}^{N} \alpha_i = 1\right\}$ represents the convex hull of the set $x_1, x_2, \ldots, x_N$. $\mathrm{dis}\{x, U\} = \inf_{y\in U}\|x - y\|$ represents the Euclidean distance between the point $x \in R$ and the set $U\subseteq R$.

2. Preliminaries and Problem Formulation

2.1. Graph Theory. In this paper, the communication of N agents with each other can be regarded as a network topology, which is represented as directed graph G. The directed graph G can be denoted by $G = (\nu, \varepsilon)$, where $\nu = 1, 2, \ldots, N$ represents the set of agents and $\varepsilon \in \nu \times \nu$ represents the set of edges. The weighted adjacency matrix $A = [a_{ij}] \in R^{n\times n}$ is described as $a_{ii} = 0, a_{ij} = 1$ if $(j, i) \in \varepsilon$, and $a_{ij} = 0$ while others. The Laplacian matrix of

$GL = [L_{ij}] \in R^{n\times n}$ is described as $L_{ii} = \sum_{j\neq i} a_{ij}$ and $L_{ij} = -a_{ij}$ if $i \neq j$.

2.2. Problem Formulation. Consider a leader-follower general linear system with a single leader and N followers. For a leader-follower system, the follower can receive information and the leader cannot receive information. The state update of the agents depends on the communication network. The communication network among agents can be regarded as a directed graph $G = (\nu, \varepsilon)$. Each node represents an agent, and each edge represents the communication between neighboring agents. The existence of an edge between the two agents indicates that there is information exchange between them.

With loss of generality, we suppose that 0 represents the leader and others represent the followers. The expression of the followers' dynamics can be described as follows:

$$\begin{aligned} \dot{x}_i(t) &= Ax_i(t) + Bu_i(t), \\ y_i(t) &= Cx_i(t), \quad t \geq 0, i \in \nu, \end{aligned} \tag{1}$$

where $x_i(t) \in R^n, u_i(t) \in R^r$, and $y_i(t) \in R^m$, respectively, represent the state, control input, and output of the followers. The expression of the leader' dynamic can be described as follows:

$$\dot{x}_0(t) = Ax_0(t), \quad t \geq 0, \tag{2}$$

where $x_0(t) \in R^n$ means the state of the leader.

Definition 1. If the states of each followers satisfy

$$\lim_{x\longrightarrow\infty} \left\| x_i(t) - x_0(t) \right\| = 0, \quad i = 1, 2, \ldots, N, \tag{3}$$

then we think the leader-follower multiagent systems reach tracking consensus.

Definition 2. If the states of each followers satisfy

$$\lim_{t\longrightarrow\infty} \operatorname{dis}\left\{ x_i(t), \operatorname{conv}\left\{ x_j(t) | j \in R \right\} \right\} = 0, \quad i \in F, \tag{4}$$

in other words, the multiagent system control goal has been achieved, then we think the leader-follower systems reach containment consensus.

In practice, the full state information of the agent maybe unavailable due to the physical limitation, and we can consider the following state observer to solve this problem:

$$\begin{aligned} \dot{\overline{x}}_i(t) &= A\overline{x}_i(t) + Bu_i(t) + F\left(\overline{y}_i(t) - y_i(t)\right), \\ \overline{y}_i(t) &= C\overline{x}_i(t), \quad t \geq 0, i \in \nu, \end{aligned} \tag{5}$$

where $\overline{x}_i(t) \in R^n$ and $\overline{y}_i(t) \in R^m$ are the state and output of the observers and $F \in R^{n\times m}$ is the feedback gain matrix which is derived in the following section. If the leader-follower multiagent system wants to reach consensus, some assumptions are indispensable.

Assumption 1. For the agent system we used, the system matrix pairs need to satisfy that (A, C) is observable and (A, B) is controllable.

Assumption 2. There is at least one leader with a directional path to each follower in the multiagent system network.

Generally speaking, the communication among agents is through the network, which will cause transmission delays and should not be ignored. In addition, the update of relative state is continuous, but it is unnecessary in actual situation; in order to save network energy, we can update the control input by some events, that is, event-triggered control method. Therefore, we can consider the distributed tracking control algorithm with input time delays as follows:

$$u_i(t) = -Kq_i\left(t_k^i - \tau\right), \quad t \in \left[t_k^i, t_{k+1}^i\right], \tag{6}$$

where $\tau \geq 0$ is the input time delay of the network, $q_i(t) = \sum_{j=1}^{N} a_{ij}\left(\overline{x}_i(t) - \overline{x}_j(t)\right) + h\left(\overline{x}_i(t) - x_0(t)\right)$, h represents the adjacency matrix between the leader and the followers, and $K \in R^{r\times n}$ is the feedback gain matrix to be determined. Let us combine (1) and (5), and the expression of control system can be written as follows:

$$\dot{\overline{x}}_i(t) = A\overline{x}_i(t) + Bu_i(t) + \mathrm{FC}\left(\overline{x}_i(t) - x_i(t)\right), \tag{7}$$

for $t \geq 0, i \in \nu$. Further, combining with (6), we can obtain

$$\dot{\overline{x}}_i(t) = A\overline{x}_i(t) - \mathrm{BK}q_i\left(t_k^i - \tau\right) + \mathrm{FC}\left(\overline{x}_i(t) - x_i(t)\right), \tag{8}$$

for $t \geq 0, i \in \nu$. Then, let $z_i(t) = \overline{x}_i(t) - x_i(t)$ be the state error between the observer and the agent. Let $\varsigma_i(t) = \overline{x}_i(t) - x_0(t)$ be the state error between the leader and follower. Let $e_i(t) = q_i(t_k^i) - q_i(t)$ be the state error between the triggering time and the real time, then the following expression can be obtained:

$$\begin{aligned} \dot{\varsigma}_i(t) &= A\varsigma_i(t) - \mathrm{BK}e_i(t-\tau) - \mathrm{BK}q_i(t-\tau) + \mathrm{FC}z_i(t) \\ &= A\varsigma_i(t) - \mathrm{BK}e_i(t-\tau) - \mathrm{BK}\sum_{j=1}^{N} a_{ij}\left(\overline{x}_i(t-\tau) - \overline{x}_j(t-\tau)\right) - \mathrm{BK}h\varsigma_i(t-\tau) + \mathrm{FC}z_i(t) \\ &= A\varsigma_i(t) - \mathrm{BK}e_i(t-\tau) - \mathrm{BK}\sum_{j=1}^{N} a_{ij}\left(\varsigma_i(t-\tau) - \varsigma_j(t-\tau)\right) - \mathrm{BK}h\varsigma_i(t-\tau) + \mathrm{FC}z_i(t). \end{aligned} \tag{9}$$

Let $\eta = L + h$, $\varsigma(t) = [\varsigma_1(t), \varsigma_2(t), \ldots, \varsigma_N(t)]^T$, and $e(t) = [e_1(t), e_2(t), \ldots, e_N(t)]^T$, then we can use the Kronecker product of matrix to obtain the following dynamic:

$$\begin{aligned}\dot{\varsigma}(t) &= (I_N \otimes A)\varsigma(t) - (I_N \otimes \mathrm{BK})e(t-\tau) - (\eta \otimes \mathrm{BK})\varsigma(t-\tau) + (I_N \otimes \mathrm{FC})z(t),\\ \dot{z}(t) &= [I_N \otimes (A + \mathrm{FC})]z(t).\end{aligned} \tag{10}$$

Lemma 1. *From the previous assumption, we can obtain $\eta = L + h$ is a positive definite matrix.*

Lemma 2. *For given orthogonal matrix U and $\eta > 0$, we can obtain that $U^T \eta U = \theta = diag(\nu_1, \nu_2, \ldots, \nu_N)$.*

Lemma 3. *For given matrix $S = \begin{bmatrix} S_{11} & S_{12} \\ S_{12}^T & S_{22} \end{bmatrix}$, where S_{11}, S_{12}, and S_{22} are compatible dimension matrices, the following descriptions are equivalent [35]:*

(1) $S < 0$

(2) $S_{22} < 0, S_{11} - S_{12}S_{22}^{-1}S_{12}^T < 0$

Lemma 4. *From the previous assumption, let $(\alpha_1^T, \ldots, \alpha_N^T)^T = (-L_1^{-1}L_2 \otimes I_M)(\beta_1^T, \ldots, \beta_M^T)^T$, where $\alpha_i \in R^n$ and $\beta_i \in R^n$. Then, we can obtain $\alpha_i \in conv\{\beta_1, \ldots, \beta_M\}$ [34].*

3. Main Result

3.1. Tracking Control. In this section, we will discuss the distributed tracking consensus issue for the leader-follower linear system with input time delay by using the event-triggered method. And sufficient event-triggering function and distributed control algorithm are given to achieve our control goals; that is to say, the tracking consensus can be reached for the leader-follower general linear system under the proposed protocol.

At first, we put forward the following event-triggered function for each follower:

$$t_{k+1}^i = \inf\{t:\ t > t_k^i, f_i(t) < 0\}, \tag{11}$$

where

$$f_i(t) = a\|e_i(t)\|^2 - b\|q_i(t)\|^2 - c\|q_i(t_k^i - \tau)\|^2, \tag{12}$$

a, b, and c are derived in the following section and t_κ^i denotes the k_{th} triggering time of the k_{th} follower, which is defined by $f_i(t) = 0$.

Theorem 1. *Consider the leader-follower linear system (1), the observer (5), the control algorithm (6), and the triggering conditions (11). The leader-follower linear system can reach tracking consensus by distributed event-triggered control protocol if there exist $P > 0$, $Q > 0$, $D > 0$, $R > 0$, $H > 0$, $W > 0$, F, and K satisfying:*

$$M = \begin{bmatrix} M_{11} & M_{12} & M_{13} & M_{14} & M_{15} \\ * & M_{22} & 0 & 0 & M_{25} \\ * & * & M_{33} & 0 & M_{35} \\ * & * & * & M_{44} & M_{45} \\ * & * & * & * & M_{55} \end{bmatrix} < 0, \tag{13}$$

where

$$\begin{aligned}
M_{11} &= \theta \otimes (\mathrm{PA} + A^T P) - \frac{1}{\tau}(\theta \otimes W) + (\theta \otimes R) + b(\theta^2 \otimes H),\\
M_{12} &= -2\theta \otimes \mathrm{PBK},\\
M_{13} &= -2\theta^2 \otimes \mathrm{PBK} + \frac{1}{\tau}(\theta \otimes W),\\
M_{14} &= 2\theta \otimes \mathrm{PFC},\\
M_{15} &= U^T \otimes A,\\
M_{22} &= -I_N \otimes H,\\
M_{25} &= -U^T \eta \otimes (\mathrm{BK})^T,\\
M_{33} &= -\theta \otimes R + b(\theta^2 \otimes D) - \frac{1}{\tau}(\theta \otimes W),\\
M_{35} &= -U^T \otimes (\mathrm{BK})^T,\\
M_{44} &= 2I_N \otimes (\mathrm{QA} + \mathrm{QFC}),\\
M_{45} &= -U^T \otimes (\mathrm{FC})^T,\\
M_{55} &= -\frac{1}{\tau}(\eta^{-1} \otimes W^{-1}).
\end{aligned} \tag{14}$$

That is, all followers can finally track the leader as well as we can avoid the Zeno behavior through the tracking control protocol we designed.

Proof. Assuming that all proposed assumptions are satisfied, then consider the following Lyapunov–Krasovskii functional $V(t) = V_1(t) + V_2(t) + V_3(t)$ with

$$\begin{aligned} V_1(t) &= \varsigma^T(t)(\eta \otimes P)\varsigma(t) + z^T(t)(I_N \otimes Q)z(t), \\ V_2(t) &= \int_{t-\tau}^{t} \left(e^T(s)(I_N \otimes H)e(s) + \varsigma^T(s)(\eta \otimes R)\varsigma(s)\right)\mathrm{d}s, \\ V_3(t) &= \int_0^{\tau}\int_{t+\theta}^{t} \dot{\varsigma}^T(s)(\eta \otimes W)\dot{\varsigma}(s)\mathrm{d}s\mathrm{d}\theta, \end{aligned} \tag{15}$$

where $P > 0$, $Q > 0$, $H > 0$, $R > 0$, and $W > 0$, so we can get that $V(t)$ is positive definite. Then, the time derivation of $V(t)$ can be given as follows:

$$\begin{aligned} \dot{V}_1(t) &= 2\varsigma^T(t)(\eta \otimes P)\dot{\varsigma}(t) + 2z^T(t)(I_N \otimes Q)\dot{z}(t), \\ \dot{V}_2(t) &= e^T(t)(I_N \otimes H)e(t) - e^T(t-\tau)(I_N \otimes H)e(t-\tau) + \varsigma^T(t)(\eta \otimes R)\varsigma(t) - \varsigma^T(t-\tau)(\eta \otimes R)\varsigma(t-\tau), \\ \dot{V}_3(t) &= \tau\dot{\varsigma}^T(t)(\eta \otimes W)\dot{\varsigma}(t) - \int_{t-\tau}^{t} \dot{\varsigma}^T(s)(\eta \otimes W)\dot{\varsigma}(s)\mathrm{d}s. \end{aligned} \tag{16}$$

Let $\delta(t) = (U^T \otimes I_N)\varsigma(t)$, $\hat{e}(t) = (U^T \otimes I_N)e(t)$, and $\hat{z}(t) = (U^T \otimes I_N)z(t)$, we can get

$$\begin{aligned} \dot{V}_1(t) &= 2\delta^T(t)\left[(\theta \otimes \mathrm{PA})\delta(t) - (\theta \otimes \mathrm{PBK})\hat{e}(t-\tau) - \left(\theta^2 \otimes \mathrm{PBK}\right)\delta(t-\tau) + (\theta \otimes \mathrm{PFC})\hat{z}(t)\right] \\ &\quad + 2\hat{z}^T(t)\left[I_N \otimes (\mathrm{QA} + \mathrm{QFC})\right]\hat{z}(t), \\ \dot{V}_2(t) &= \hat{e}^T(t)(I_N \otimes H)\hat{e}(t) - \hat{e}^T(t-\tau)(I_N \otimes H)\hat{e}(t-\tau) \\ &\quad + \delta^T(t)(\theta \otimes R)\delta(t) - \delta^T(t-\tau)(\theta \otimes R)\delta(t-\tau), \\ \dot{V}_3(t) &= \tau\dot{\varsigma}^T(t)(\eta \otimes W)\dot{\varsigma}(t) - \int_{t-\tau}^{t} \dot{\delta}^T(s)(\eta \otimes W)\dot{\delta}(s)\mathrm{d}s \\ &\le \tau\dot{\varsigma}^T(t)(\eta \otimes W)\dot{\varsigma}(t) - \frac{1}{\tau}(\delta(t) - \delta(t-\tau))^T(\theta \otimes W)(\delta(t) - \delta(t-\tau)) \\ &= \tau\left[\delta^T(t)\left(\theta \otimes A^T\mathrm{WA}\right)\delta(t) - \delta^T(t)\left(\theta \otimes A^T\mathrm{WBK}\right)\hat{e}(t-\tau)\right. \\ &\quad - \delta^T(t)\left(\theta^2 \otimes A^T\mathrm{WBK}\right)\delta(t) + \delta^T(t)\left(\theta \otimes A^T\mathrm{WFC}\right)\hat{z}(t) \\ &\quad - \hat{e}^T(t-\tau)\left(\theta \otimes (\mathrm{BK})^T\mathrm{WA}\right)\delta(t) + \hat{e}^T(t-\tau)\left(\theta \otimes (\mathrm{BK})^T\mathrm{WBK}\right)\hat{e}(t-\tau) \\ &\quad + \hat{e}^T(t-\tau)\left(\theta^2 \otimes (\mathrm{BK})^T\mathrm{WBK}\right)\delta(t-\tau) - \hat{e}^T(t-\tau)\left(\theta \otimes (\mathrm{BK})^T\mathrm{WFC}\right)\hat{z}(t) \\ &\quad - \delta^T(t-\tau)\left(\theta^2 \otimes (\mathrm{BK})^T\mathrm{WA}\right)\delta(t) + \delta^T(t-\tau)\left(\theta^2 \otimes (\mathrm{BK})^T\mathrm{WBK}\right)\hat{e}(t-\tau) \\ &\quad + \delta^T(t-\tau)\left(\theta^3 \otimes (\mathrm{BK})^T\mathrm{WBK}\right)\delta(t-\tau) - \delta^T(t-\tau)\left(\theta^2 \otimes (\mathrm{BK})^T\mathrm{WFC}\right)\hat{z}(t) \\ &\quad + \hat{z}^T(t)\left(\theta \otimes (\mathrm{FC})^T\mathrm{WA}\right)\delta(t) - \hat{z}^T(t)\left(\theta \otimes (\mathrm{FC})^T\mathrm{WBK}\right)\hat{e}(t-\tau) \\ &\quad \left. -\hat{z}^T(t)\left(\theta^2 \otimes (\mathrm{FC})^T\mathrm{WBK}\right)\delta(t-\tau) + \hat{z}^T(t)\left(\theta \otimes (\mathrm{FC})^T\mathrm{WFC}\right)\hat{z}(t)\right] \\ &\quad - \frac{1}{\tau}(\delta(t) - \delta(t-\tau))^T(\theta \otimes W)(\delta(t) - \delta(t-\tau)). \end{aligned} \tag{17}$$

Then, according to the proposed event-triggering function (12), we can easily get

$$\begin{aligned}\widehat{e}^T(t)(I_N\otimes H)\widehat{e}(t) &\le \lambda_m(H)\sum_{i=1}^{N}\left\|e_i(t)\right\|^2 \le b\sum_{i=1}^{N}\left\|q_i(t)\right\|^2 + c\sum_{i=1}^{N}\left\|q_i(t-\tau)\right\|^2\\ &= b\varsigma^T(t)\left(\eta^2\otimes I_N\right)\varsigma(t) + c\varsigma^T(t-\tau)\left(\eta^2\otimes I_N\right)\varsigma(t-\tau)\\ &\le k_1\delta^T(t)\left(\theta^2\otimes H\right)\delta(t) + k_2\delta^T(t-\tau)\left(\theta^2\otimes D\right)\delta(t-\tau),\end{aligned} \tag{18}$$

where $a=\lambda_m(H)$, $b=k_1\lambda_m(H)$, and $c=k_2\lambda_m(D)$. Then, the following expression can be summarized:

$$\dot{V}(t)\le \varepsilon^T(t)\widehat{M}\varepsilon(t), \tag{19}$$

where

$$\varepsilon(t)=\left[\delta^T(t),\widehat{e}^T(t-\tau),\delta^T(t-\tau),\widehat{z}(t)\right],$$

$$\widehat{M}=\begin{bmatrix}\widehat{M}_{11} & \widehat{M}_{12} & \widehat{M}_{13} & \widehat{M}_{14}\\ * & \widehat{M}_{22} & \widehat{M}_{23} & \widehat{M}_{24}\\ * & * & \widehat{M}_{33} & \widehat{M}_{34}\\ * & * & * & \widehat{M}_{44}\end{bmatrix},$$

$$\begin{aligned}\widehat{M}_{11} &= \theta\otimes\left(\mathrm{PA}+A^T P\right)+\tau\left(\theta\otimes A^T\mathrm{WA}\right)-\frac{1}{\tau}(\theta\otimes W)+(\theta\otimes R)+k_1\left(\theta^2\otimes H\right),\\ \widehat{M}_{12} &= -2\theta^2\otimes\mathrm{PBK}-\tau\left(\theta\otimes A^T\mathrm{WBK}\right),\\ \widehat{M}_{13} &= -2\theta^2\otimes\mathrm{PBK}-\tau\left(\theta^2\otimes A^T\mathrm{WBK}\right)+\frac{1}{\tau}(\theta\otimes W),\\ \widehat{M}_{14} &= 2\theta\otimes\mathrm{PFC}+\tau\left(\theta\otimes A^T\mathrm{WFC}\right),\\ \widehat{M}_{22} &= -(I_N\otimes H)+\tau\left(\theta\otimes(\mathrm{BK})^T\mathrm{WBK}\right),\\ \widehat{M}_{23} &= \tau\left(\theta^2\otimes(\mathrm{BK})^T\mathrm{WBK}\right),\\ \widehat{M}_{24} &= -\tau\left(\theta\otimes(\mathrm{BK})^T\mathrm{WFC}\right),\\ \widehat{M}_{33} &= -\theta\otimes R+\tau\left(\theta^3\otimes(\mathrm{BK})^T\mathrm{WBK}\right)+k_2\left(\theta^2\otimes D\right)-\frac{1}{\tau}(\theta\otimes W),\\ \widehat{M}_{34} &= -\tau\left(\theta^2\otimes(\mathrm{BK})^T\mathrm{WFC}\right),\\ \widehat{M}_{44} &= \tau\left(\theta\otimes(\mathrm{FC})^T\mathrm{WFC}\right).\end{aligned} \tag{20}$$

By (19), we can get that only if $\widehat{M}<0$ and $\varepsilon(t)\neq 0$, $\dot{V}(t)<0$ is established. In other words, the control error system (10) can be proven to be progressively stable. Then, according to Lemma 3, $\widehat{M}<0$ can be equivalent to

$$M=\begin{bmatrix}M_{11} & M_{12} & M_{13} & M_{14} & M_{15}\\ * & M_{22} & 0 & 0 & M_{25}\\ * & * & M_{33} & 0 & M_{35}\\ * & * & * & M_{44} & M_{45}\\ * & * & * & * & M_{55}\end{bmatrix}<0, \tag{21}$$

where

$$
\begin{aligned}
M_{11} &= \theta \otimes \left(\mathrm{PA} + A^T P\right) - \frac{1}{\tau}(\theta \otimes W) + (\theta \otimes R) + b\left(\theta^2 \otimes H\right),\\
M_{12} &= -2\theta \otimes \mathrm{PBK},\\
M_{13} &= -2\theta^2 \otimes \mathrm{PBK} + \frac{1}{\tau}(\theta \otimes W),\\
M_{14} &= 2\theta \otimes \mathrm{PFC},\\
M_{15} &= U^T \otimes A,\\
M_{22} &= -I_N \otimes H,\\
M_{25} &= -U^T \eta \otimes (\mathrm{BK})^T,\\
M_{33} &= -\theta \otimes R + b\left(\theta^2 \otimes D\right) - \frac{1}{\tau}(\theta \otimes W),\\
M_{35} &= -U^T \otimes (\mathrm{BK})^T,\\
M_{44} &= 2I_N \otimes (\mathrm{QA} + \mathrm{QFC}),\\
M_{45} &= -U^T \otimes (\mathrm{FC})^T,\\
M_{55} &= -\frac{1}{\tau}\left(\eta^{-1} \otimes W^{-1}\right).
\end{aligned}
\tag{22}
$$

So, we can obtain that $\lim_{t\longrightarrow\infty}\varsigma(t) = 0$ by using Lyapunov stability theory; that is to say, all followers of the multiagent system can finally track the leader. Zeno behavior is an important problem for the event-triggered control method. We should ensure strictly positive triggering interval to avoid it. The proof is as follows. The time derivative of $\|e_i(t)\|$ for $t \in (t_k^i, t_{k+1}^i)$ can be obtained as follows:

$$
\begin{aligned}
\frac{d\left\|e_i(t)\right\|}{dt} &\le \left\|\dot{e}_i\ (t)\right\| = \left\|-\dot{q}_i\ (t)\right\|\\
&= \left\|-\sum_{j=1}^{N} a_{ij}\left(\dot{\bar{x}}_i(t) - \dot{\bar{x}}_j(t)\right) + h\left(\dot{\bar{x}}_i(t) - \dot{x}_0\ (t)\right)\right\|\\
&= \left\|-Aq_i(t) - \left(\sum_{j=1}^{N} a_{ij}B\left(u_i(t) - u_j(t)\right) + \mathrm{FC}\left(z_i(t) - z_j(t)\right) + hBu_i(t) + hFCz_i(t)\right)\right\|\\
&\le \|A\|\left\|e_i(t)\right\| + \omega,
\end{aligned}
\tag{23}
$$

where $\omega = \max_{t\in[t_k^i, t_{k+1}^i))}\|Aq_i(t_k^i) + \sum_{j=1}^{N} a_{ij}B(u_i(t) - u_j(t)) + \mathrm{FC}(z_i(t) - z_j\ (t)) + hBu_i(t) + hFCz_i(t)\|$; then, consider a nonnegative function that satisfies the following conditions:

$$
\begin{aligned}
\dot{\varphi} &= \|A\|\varphi + \omega,\\
\varphi(0) &= \left\|e_i\left(t_k^i\right)\right\| = 0,
\end{aligned}
\tag{24}
$$

and hence we can obtain $\|e_i(t\| \le \varphi(t - t_k^i)$, where $\varphi(t)$ is the analytical solution to (24), which is given by $\varphi(t) = (\omega/\|A\|)(e^{\|A\|t} - 1)$. In order to ensure the event trigger function $f_i(t) \le 0$, we give the following sufficient condition:

$$
a\left\|e_i(t)\right\|^2 \le c\left\|q_i\left(t_k^i - \tau\right)\right\|^2,
\tag{25}
$$

so we can easily get $\|e_i(t)\| \le \sqrt{(c/a)}\|q_i(t_k^i - \tau)\|$; let $\rho_k^i = \sqrt{(c/a)}\|q_i(t_k^i - \tau)\|$. It follows

$$\left\|e_i\left(t_{k+1}^i\right)\right\| = \rho_k^i \le \left(\frac{\omega}{\|A\|}\right)\left(e^{\|A\|\left(t_{k+1}^i - t_k^i\right)} - 1\right). \tag{26}$$

Then, we can obtain

$$\tau_k^i = t_{k+1}^i - t_k^i \ge \frac{1}{\|A\|}\ln\left(\frac{\|A\|\rho_k^i}{\omega} + 1\right). \tag{27}$$

We can know that the trigger interval always exists and is strictly positive at any time. Only when $t \longrightarrow \infty$, we can obtain $t_k^i \longrightarrow \infty$ and $k \longrightarrow \infty$. Therefore, all the agents will not exist Zeno behavior at any finite time. So, the proof is thus completed. □

3.2. Containment Control. In this section, we will discuss the distributed containment consensus problem for the leader-follower linear system with input time delay by using the event-triggered method. And sufficient event-triggering function and distributed control algorithm are given to achieve our control goals; that is to say, the containment consensus can be reached for the general linear system under the proposed protocol. The general linear system contains M leaders and N followers. Without loss of generality, we assume that the followers are labeled as $F = \{1, 2, \ldots, N\}$ and the leaders are labeled as $R = \{N+1, N+2, \ldots, N+M\}$, and the communication among agents can also be defined as $G = (\nu, \varepsilon)$. Under the previous assumptions, we can obtain that the Laplacian matrix of G can be written as $L = \begin{bmatrix} L_1 & L_2 \\ 0 & 0 \end{bmatrix}$, where L_1 is also a positive definite matrix. The dynamics of the followers can be expressed as (1), and the expression of the leaders' dynamics can be described as follows:

$$\dot{x}_i(t) = Ax_i(t), \quad t \ge 0, i \in R, \tag{28}$$

where $x_i(t) \in R^n$ means the state of the leaders. For convenience, let ${}^*f = ({}^*1^T, \ldots, {}^*N^T)^T$ represents the information of the followers and ${}^*r = ({}^*N+1^T, \ldots, {}^*N+M^T)^T$ represents the information of the leaders. We can consider the distributed event-triggered containment control protocol with input time delays as follows:

$$u_i(t) = -Kq_i\left(t_k^i - \tau\right), \quad t \in \left[t_k^i, t_{k+1}^i\right], \tag{29}$$

where $q_i(t) = \sum_{j\in F} a_{ij}(\overline{x}_i(t) - \overline{x}_j(t)) + \sum_{j\in R} a_{ij}(\overline{x}_i(t) - x_j(t))$. Then, the expression of control system can be written as follows:

$$\begin{aligned}
\dot{x}_r(t) &= (I_M \otimes A)x_r(t), \\
\dot{\overline{x}}_f(t) &= (I_N \otimes A)\overline{x}_f(t) - (I_N \otimes \mathrm{BK})e_f(t-\tau) - (L_1 \otimes \mathrm{BK})\overline{x}_f(t-\tau) - (L_2 \otimes \mathrm{BK})x_r(t-\tau) + (I_N \otimes \mathrm{FC})z_f(t), \\
\dot{z}_f(t) &= [I_N \otimes (A + \mathrm{FC})]z_f(t).
\end{aligned} \tag{30}$$

Let $\varsigma_f(t) = \overline{x}_f(t) + (L_1^{-1}L_2 \otimes I_M)x_r(t)$. Then, we have the following expression:

$$\begin{aligned}
\dot{\varsigma}_f(t) &= (I_N \otimes A)\varsigma_f(t) - (I_N \otimes \mathrm{BK})e_f(t-\tau) - (L_1 \otimes \mathrm{BK})\varsigma_f(t) + (I_N \otimes \mathrm{FC})z_f(t), \\
\dot{z}_f(t) &= [I_N \otimes (A + \mathrm{FC})]z_f(t).
\end{aligned} \tag{31}$$

Theorem 2. *Consider the leader-follower linear system (1), the observer (5), the control algorithm (29), and the triggering conditions (11). The leader-follower linear system can reach containment consensus by distributed event-triggered control protocol if there exists $P > 0$, $Q > 0$, $D > 0$, $R > 0$, $H > 0$, and $W > 0$ and F and K satisfied:*

$$\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} & \Phi_{13} & \Phi_{14} & \Phi_{15} \\ * & \Phi_{22} & 0 & 0 & \Phi_{25} \\ * & * & \Phi_{33} & 0 & \Phi_{35} \\ * & * & * & \Phi_{44} & \Phi_{45} \\ * & * & * & * & \Phi_{55} \end{bmatrix} < 0, \tag{32}$$

where

$$\begin{aligned}
\Phi_{11} &= \theta\otimes\left(\mathrm{PA}+A^T P\right)-\frac{1}{\tau}(\theta\otimes W)+(\theta\otimes R)+b\left(\theta^2\otimes H\right),\\
\Phi_{12} &= -2\theta\otimes\mathrm{PBK},\\
\Phi_{13} &= -2\theta^2\otimes\mathrm{PBK}+\frac{1}{\tau}(\theta\otimes W),\\
\Phi_{14} &= 2\theta\otimes\mathrm{PFC},\\
\Phi_{15} &= U^T\otimes A,\\
\Phi_{22} &= -I_N\otimes H,\\
\Phi_{25} &= -U^T L_1\otimes(\mathrm{BK})^T,\\
\Phi_{33} &= -\theta\otimes R+b\left(\theta^2\otimes D\right)-\frac{1}{\tau}(\theta\otimes W),\\
\Phi_{35} &= -U^T\otimes(\mathrm{BK})^T,\\
\Phi_{44} &= 2I_N\otimes(\mathrm{QA}+\mathrm{QFC}),\\
\Phi_{45} &= -U^T\otimes(\mathrm{FC})^T,\\
\Phi_{55} &= -\frac{1}{\tau}\left(L_1^{-1}\otimes W^{-1}\right).
\end{aligned} \tag{33}$$

Then, we can get that the leader-follower linear system can eventually achieve containment consensus.

Proof. Assuming that all proposed assumptions are satisfied, then consider the following Lyapunov–Krasovskii functional $V(t)=V_1(t)+V_2(t)+V_3(t)$ with

$$\begin{aligned}
V_1(t) &= \varsigma_f^T(t)\left(L_1\otimes P\right)\varsigma_f(t)+z_f^T(t)\left(I_N\otimes Q\right)z_f(t),\\
V_2(t) &= \int_{t-\tau}^{t}\left(e_f^T(s)\left(I_N\otimes H\right)e_f(s)+\varsigma_f^T(s)\left(L_1\otimes R\right)\varsigma_f(s)\right)\mathrm{d}s,\\
V_3(t) &= \int_0^{\tau}\int_{t+\theta}^{t}\dot{\varsigma}_f^T(s)\left(L_1\otimes W\right)\dot{\varsigma}_f(s)\mathrm{d}s\mathrm{d}\theta,
\end{aligned} \tag{34}$$

where $P>0$, $Q>0$, $H>0$, $R>0$, and $W>0$, so we can get that $V(t)$ is positive definite. Then, the time derivation of $V(t)$ can be given as follows:

$$\begin{aligned}
\dot{V}_1(t) &= 2\varsigma_f^T(t)\left(L_1\otimes P\right)\dot{\varsigma}_f(t)+2z_f^T(t)\left(I_N\otimes Q\right)\dot{z}_f(t),\\
\dot{V}_2(t) &= e_f^T(t)\left(I_N\otimes H\right)e_f(t)-e_f^T(t-\tau)\left(I_N\otimes H\right)e_f(t-\tau)\\
&\quad+\varsigma_f^T(t)\left(L_1\otimes R\right)\varsigma_f(t)-\varsigma_f^T(t-\tau)\left(L_1\otimes R\right)\varsigma_f(t-\tau),\\
\dot{V}_3(t) &= \tau\dot{\varsigma}_f^T(t)\left(L_1\otimes W\right)\dot{\varsigma}_f(t)-\int_{t-\tau}^{t}\dot{\varsigma}_f^T(s)\left(L_1\otimes W\right)\dot{\varsigma}_f(s)\mathrm{d}s.
\end{aligned} \tag{35}$$

Let $\delta_f(t)=(U^T\otimes I_N)\varsigma_f(t)$, $\hat{e}_f(t)=(U^T\otimes I_N)e_f(t)$, and $\hat{z}_f(t)=(U^T\otimes I_N)z_f(t)$, we can get

$$\begin{aligned}
\dot{V}_1(t) &= 2\delta_f^T(t)\left[(\theta\otimes\mathrm{PA})\delta_f(t)-(\theta\otimes\mathrm{PBK})\hat{e}_f(t-\tau)-\left(\theta^2\otimes\mathrm{PBK}\right)\delta_f(t-\tau)+(\theta\otimes\mathrm{PFC})\hat{z}_f(t)\right]\\
&\quad+2\hat{z}_f^T(t)\left[I_N\otimes(\mathrm{QA}+\mathrm{QFC})\right]\hat{z}_f(t),\\
\dot{V}_2(t) &= \hat{e}_f^T(t)\left(I_N\otimes H\right)\hat{e}_f(t)-\hat{e}_f^T(t-\tau)\left(I_N\otimes H\right)\hat{e}_f(t-\tau)\\
&\quad+\delta_f^T(t)(\theta\otimes R)\delta_f(t)-\delta_f^T(t-\tau)(\theta\otimes R)\delta_f(t-\tau),\\
\dot{V}_3(t) &= \tau\dot{\varsigma}_f^T(t)\left(L_1\otimes W\right)\dot{\varsigma}_f(t)-\int_{t-\tau}^{t}\dot{\delta}_f^T(s)\left(L_1\otimes W\right)\dot{\delta}_f(s)\mathrm{d}s\\
&\le\tau\dot{\varsigma}_f^T(t)\left(L_1\otimes W\right)\dot{\varsigma}_f(t)-\frac{1}{\tau}\left(\delta_f(t)-\delta_f(t-\tau)\right)^T(\theta\otimes W)\left(\delta_f(t)-\delta_f(t-\tau)\right)\\
&=\tau\left[\delta_f^T(t)\left(\theta\otimes A^T\mathrm{WA}\right)\delta_f(t)-\delta_f^T(t)\left(\theta\otimes A^T\mathrm{WBK}\right)\hat{e}_f(t-\tau)\right.\\
&\quad-\delta_f^T(t)\left(\theta^2\otimes A^T\mathrm{WBK}\right)\delta_f(t)+\delta_f^T(t)\left(\theta\otimes A^T\mathrm{WFC}\right)\hat{z}_f(t)\\
&\quad-\hat{e}_f^T(t-\tau)\left(\theta\otimes(\mathrm{BK})^T\mathrm{WA}\right)\delta_f(t)+\hat{e}_f^T(t-\tau)\left(\theta\otimes(\mathrm{BK})^T\mathrm{WBK}\right)\hat{e}_f(t-\tau)\\
&\quad+\hat{e}_f^T(t-\tau)\left(\theta^2\otimes(\mathrm{BK})^T\mathrm{WBK}\right)\delta_f(t-\tau)-\hat{e}_f^T(t-\tau)\left(\theta\otimes(\mathrm{BK})^T\mathrm{WFC}\right)\hat{z}_f(t)\\
&\quad-\delta_f^T(t-\tau)\left(\theta^2\otimes(\mathrm{BK})^T\mathrm{WA}\right)\delta_f(t)+\delta_f^T(t-\tau)\left(\theta^2\otimes(\mathrm{BK})^T\mathrm{WBK}\right)\hat{e}_f(t-\tau)\\
&\quad+\delta_f^T(t-\tau)\left(\theta^3\otimes(\mathrm{BK})^T\mathrm{WBK}\right)\delta_f(t-\tau)-\delta_f^T(t-\tau)\left(\theta^2\otimes(\mathrm{BK})^T\mathrm{WFC}\right)\hat{z}_f(t)\\
&\quad+\hat{z}_f^T(t)\left(\theta\otimes(\mathrm{FC})^T\mathrm{WA}\right)\delta_f(t)-\hat{z}_f^T(t)\left(\theta\otimes(\mathrm{FC})^T\mathrm{WBK}\right)\hat{e}_f(t-\tau)\\
&\quad\left.-\hat{z}_f^T(t)\left(\theta^2\otimes(\mathrm{FC})^T\mathrm{WBK}\right)\delta_f(t-\tau)+\hat{z}_f^T(t)\left(\theta\otimes(\mathrm{FC})^T\mathrm{WFC}\right)\hat{z}_f(t)\right]\\
&\quad-\frac{1}{\tau}\left(\delta_f(t)-\delta_f(t-\tau)\right)^T(\theta\otimes W)\left(\delta_f(t)-\delta_f(t-\tau)\right).
\end{aligned} \tag{36}$$

Then, according to the proposed event-triggering function (12), we can easily obtain

$$\begin{aligned}
&\widehat{e}^T(t)(I_N \otimes H)\widehat{e}(t) \le \lambda_m(H)\sum_{i=1}^{N}\left\|e_i(t)\right\|^2 \\
&\le b\sum_{i=1}^{N}\left\|q_i(t)\right\|^2 + c\sum_{i=1}^{N}\left\|q_i(t-\tau)\right\|^2 \\
&= b\varsigma_f^T(t)\left(L_1^2 \otimes I_N\right)\varsigma_f(t) + c\varsigma_f^T(t-\tau)\left(L_1^2 \otimes I_N\right)\varsigma_f(t-\tau) \\
&\le k_1\delta_f^T(t)\left(\theta^2 \otimes H\right)\delta_f(t) + k_2\delta_f^T(t-\tau)\left(\theta^2 \otimes D\right)\delta_f(t-\tau).
\end{aligned} \tag{37}$$

Then, the following expression can be summarized:

$$\dot{V}(t) \le \varphi^T(t)\widehat{\Phi}\varphi(t), \tag{38}$$

where

$$\varphi(t) = \left[\delta_f^T(t), \widehat{e}_f^T(t-\tau), \delta_f^T(t-\tau), \widehat{z}_f(t)\right],$$

$$\widehat{\Phi} = \begin{bmatrix} \widehat{\Phi}_{11} & \widehat{\Phi}_{12} & \widehat{\Phi}_{13} & \widehat{\Phi}_{14} \\ * & \widehat{\Phi}_{22} & \widehat{\Phi}_{23} & \widehat{\Phi}_{24} \\ * & * & \widehat{\Phi}_{33} & \widehat{\Phi}_{34} \\ * & * & * & \widehat{\Phi}_{44} \end{bmatrix},$$

$$\begin{aligned}
\widehat{\Phi}_{11} &= \theta \otimes \left(\mathrm{PA} + A^T P\right) + \tau\left(\theta \otimes A^T \mathrm{WA}\right) - \frac{1}{\tau}(\theta \otimes W) + (\theta \otimes R) + k_1\left(\theta^2 \otimes H\right), \\
\widehat{\Phi}_{12} &= -2\theta^2 \otimes \mathrm{PBK} - \tau\left(\theta \otimes A^T \mathrm{WBK}\right), \\
\widehat{\Phi}_{13} &= -2\theta^2 \otimes \mathrm{PBK} - \tau\left(\theta^2 \otimes A^T \mathrm{WBK}\right) + \frac{1}{\tau}(\theta \otimes W), \\
\widehat{\Phi}_{14} &= 2\theta \otimes \mathrm{PFC} + \tau\left(\theta \otimes A^T \mathrm{WFC}\right), \\
\widehat{\Phi}_{22} &= -(I_N \otimes H) + \tau\left(\theta \otimes (\mathrm{BK})^T \mathrm{WBK}\right), \\
\widehat{\Phi}_{23} &= \tau\left(\theta^2 \otimes (\mathrm{BK})^T \mathrm{WBK}\right), \\
\widehat{\Phi}_{24} &= -\tau\left(\theta \otimes (\mathrm{BK})^T \mathrm{WFC}\right), \\
\widehat{\Phi}_{33} &= -\theta \otimes R + \tau\left(\theta^3 \otimes (\mathrm{BK})^T \mathrm{WBK}\right) + k_2\left(\theta^2 \otimes D\right) - \frac{1}{\tau}(\theta \otimes W), \\
\widehat{\Phi}_{34} &= -\tau\left(\theta^2 \otimes (\mathrm{BK})^T \mathrm{WFC}\right), \\
\widehat{\Phi}_{44} &= \tau\left(\theta \otimes (\mathrm{FC})^T \mathrm{WFC}\right).
\end{aligned} \tag{39}$$

By (38), we can get that only if $\widehat{\Phi} < 0$ and $\varphi(t) \neq 0$, $\dot{V}(t) < 0$ is established. In other words, the control error system (31) can be proven to be progressively stable. Then, according to Lemma 3, $\widehat{\Phi} < 0$ can be equivalent to

$$\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} & \Phi_{13} & \Phi_{14} & \Phi_{15} \\ * & \Phi_{22} & 0 & 0 & \Phi_{25} \\ * & * & \Phi_{33} & 0 & \Phi_{35} \\ * & * & * & \Phi_{44} & \Phi_{45} \\ * & * & * & * & \Phi_{55} \end{bmatrix} < 0, \tag{40}$$

where

$$\Phi_{11} = \theta \otimes \left(PA + A^T P\right) - \frac{1}{\tau}(\theta \otimes W) + (\theta \otimes R) + b\left(\theta^2 \otimes H\right),$$

$$\Phi_{12} = -2\theta \otimes PBK,$$

$$\Phi_{13} = -2\theta^2 \otimes PBK + \frac{1}{\tau}(\theta \otimes W),$$

$$\Phi_{14} = 2\theta \otimes PFC,$$

$$\Phi_{15} = U^T \otimes A,$$

$$\Phi_{22} = -I_N \otimes H,$$

$$\Phi_{25} = -U^T L_1 \otimes (BK)^T,$$

$$\Phi_{33} = -\theta \otimes R + b\left(\theta^2 \otimes D\right) - \frac{1}{\tau}(\theta \otimes W),$$

$$\Phi_{35} = -U^T \otimes (BK)^T,$$

$$\Phi_{44} = 2I_N \otimes (QA + QFC),$$

$$\Phi_{45} = -U^T \otimes (FC)^T,$$

$$\Phi_{55} = -\frac{1}{\tau}\left(L_1^{-1} \otimes W^{-1}\right). \tag{41}$$

Then, we have $\lim_{t\longrightarrow\infty}\varsigma_f(t) = 0$ by using Lyapunov stability theory, that is, $\lim_{t\longrightarrow\infty}\|\overline{x}_f(t) + (L_1^{-1}L_2 \otimes I_M)x_r(t)\| = 0$ and $\lim_{t\longrightarrow\infty}z_f(t) = 0$. Thus, we can get that $\lim_{t\longrightarrow\infty}\|x_f(t) + (L_1^{-1}L_2 \otimes I_M)x_r(t)\| = 0$. Then, according to Lemma 4, distributed control goals can be achieved; that is, the leader-follower system can finally reach the containment consensus. The proof of Zeno behavior is similar to that in Theorem 1. And the proof is thus completed. □

4. Simulation Example

First, we discuss the situation of the leader-follower linear system with a single leader, and the MAS with identical agent dynamics is as follows:

$$x_i(t) = \begin{bmatrix} x_{i1}(t) \\ x_{i2}(t) \end{bmatrix},$$
$$A = \begin{bmatrix} 0 & 0.4 \\ 0.2 & -0.6 \end{bmatrix},$$
$$B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{42}$$
$$C = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

Then, we consider that the communication connection weights are $a_{12} = a_{13} = a_{24} = a_{34} = 1$, $h_1 = 1$, and others are zero. Let $x_0(t) = \begin{bmatrix} -2 & 3 \end{bmatrix}^T$, $x_1(t) = \begin{bmatrix} -10 & 1 \end{bmatrix}^T$, $x_2(t) = \begin{bmatrix} 8 & -2 \end{bmatrix}^T$, $x_4(t) = \begin{bmatrix} -12 & 4 \end{bmatrix}^T$, and $x_4(t) = \begin{bmatrix} 13 & 4 \end{bmatrix}^T$. We should note that when the event trigger method is used to deal with the delay problem, it is necessary to ensure that the delay is bigger than the event trigger time interval; otherwise, the time delay problem will become meaningless; therefore, we will choose the largest possible time delay bigger than the trigger time interval. So, we choose the appropriate input time delay $\tau = 0.1s$. By solving (13) in Theorem 1, we can get $K = \begin{bmatrix} 1.6538 & 1.7246 \end{bmatrix}$ and $F = \begin{bmatrix} 1.3429 & 0.8573 \end{bmatrix}^T$. The results are shown in Figures 1–3.

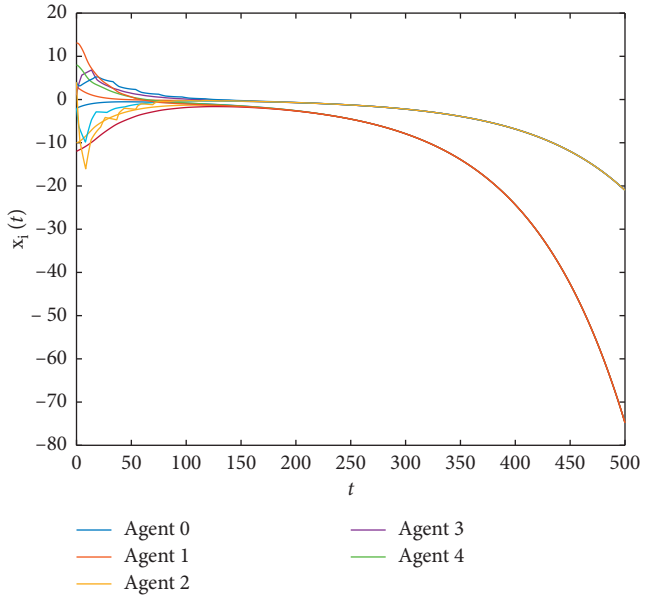


Figure 1: The states for tracking control.

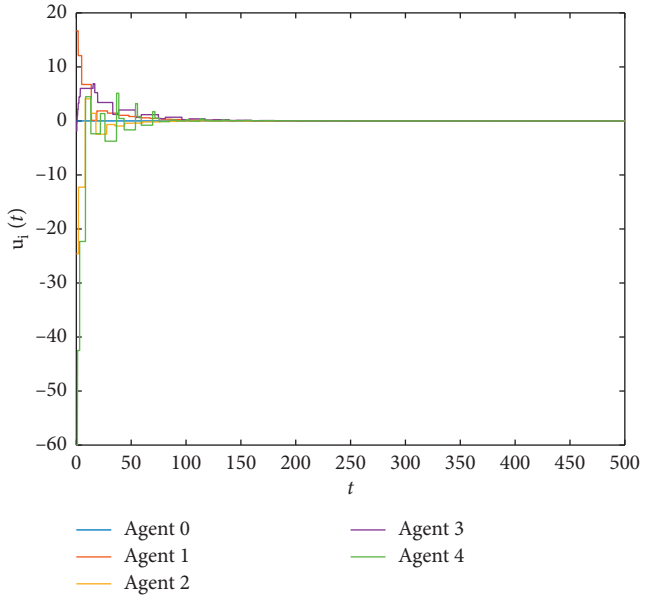


Figure 2: The control inputs for tracking control.

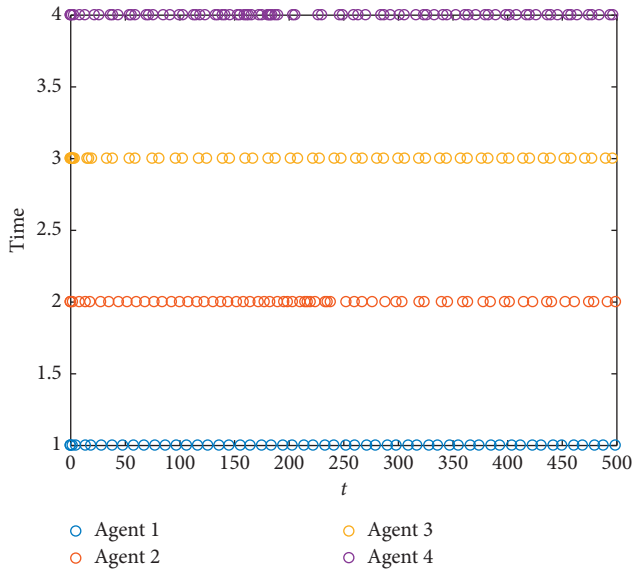


Figure 3: The event-triggered time for tracking control.

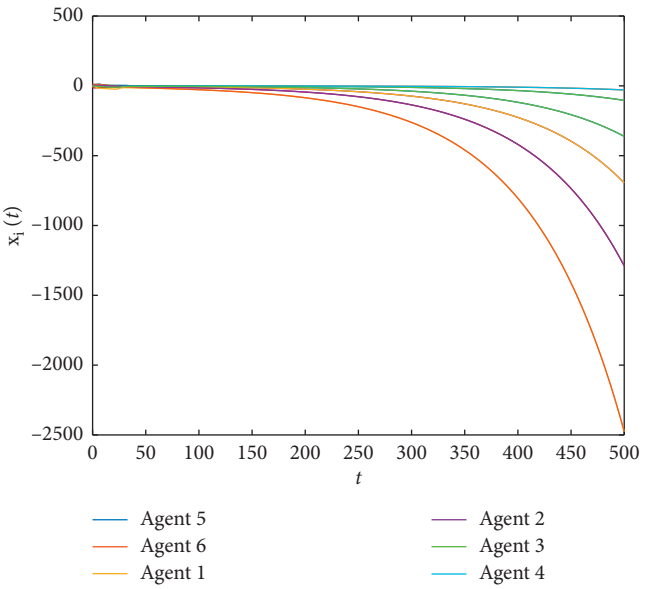


Figure 4: The states for containment control.

Then, we discuss the situation of the leader-follower linear system with four followers and two leaders, the followers are marked as $\{1, 2, 3, 4\}$, and the leaders are marked as $\{5, 6\}$. The MAS with identical agent dynamics is as follows:

$$x_i(t) = \begin{bmatrix} x_{i1}(t) \\ x_{i2}(t) \end{bmatrix},$$
$$A = \begin{bmatrix} 0 & 0.4 \\ 0.2 & -0.6 \end{bmatrix},$$
$$B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad (43)$$
$$C = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

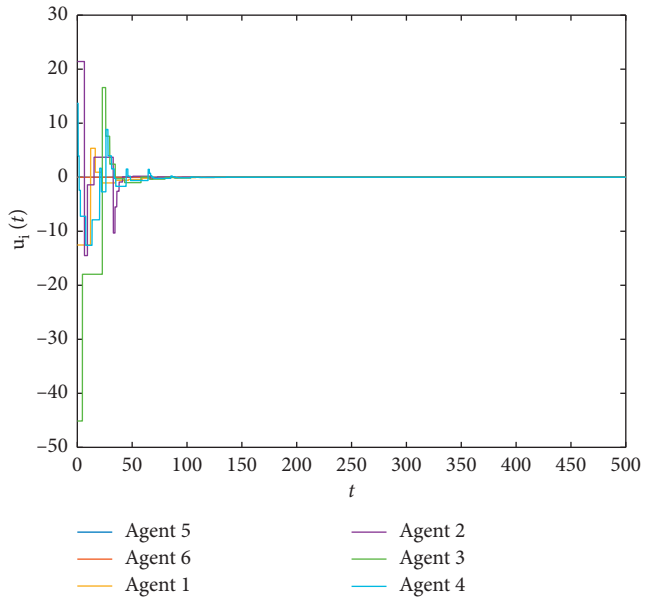


Figure 5: The control inputs for containment control.

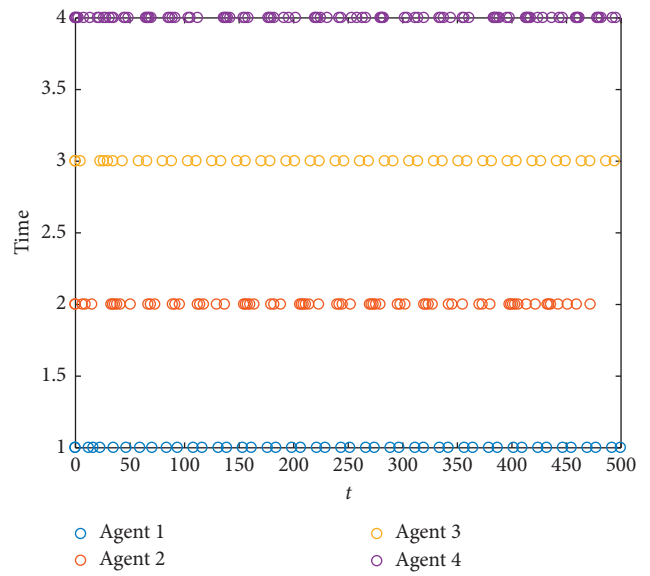


Figure 6: The event-triggered time for containment control.

Let $x_1(t) = \begin{bmatrix} 9 & -3 \end{bmatrix}^T$, $x_2(t) = \begin{bmatrix} -15 & 5 \end{bmatrix}^T$, $x_3(t) = \begin{bmatrix} 12 & 3 \end{bmatrix}^T$, $x_4(t) = \begin{bmatrix} 4 & -4 \end{bmatrix}^T$, $x_5(t) = \begin{bmatrix} -1 & 1 \end{bmatrix}^T$, and $x_6(t) = \begin{bmatrix} -11 & 1 \end{bmatrix}^T$. As with the above conditions, we choose the appropriate input time delay $\tau = 0.1s$. By solving (32) in Theorem 2, we can get $K = \begin{bmatrix} 1.8356 & 1.4528 \end{bmatrix}$ and $F = \begin{bmatrix} 1.3429 & 1.2176 \end{bmatrix}^T$. The results are shown in Figures 4–6.

The following information can be obtained through the above simulation results:

(1) For the leader-follower linear system with a single leader, by designing appropriate feedback gain matrix and event trigger parameters, the followers'

states can be same as the leader and the followers' control inputs tend to zero eventually

(2) For the leader-follower linear system with multiple leaders, by designing appropriate feedback gain matrix and event trigger parameters, the system can reach containment consensus and the followers' control inputs also tend to zero eventually

5. Conclusions

This paper considered the observer-based tracking and containment control on the basis of event-triggering mechanism for the leader-follower linear systems with input time delays. A fully distributed delay-dependent containment controller and event trigger have been designed for the general linear systems under directed graphs. Under the designed algorithms, the leader-follower linear system can achieve tracking consensus and containment. Two simulation results showed the effect of the designed control protocol. In the future, we will study the containment problems of the leader-follower systems with switching networks by using the event-triggered method.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Key Natural Science Foundation of Hubei (grant 2019CFb423).

References

[1] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[2] W. Ren, R. W. Beard, and E. M. Atkins, "Information consensus in multivehicle cooperative control," *IEEE Control Systems Magazine*, vol. 27, no. 2, pp. 71–82, 2007.

[3] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[4] W. Ren and R. W. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Transactions on Automatic Control*, vol. 50, no. 5, pp. 655–661, 2005.

[5] W. Ren, "On consensus algorithms for double-integrator dynamics," *IEEE Transactions on Automatic Control*, vol. 53, no. 6, pp. 1503–1509, 2008.

[6] Z. Li, W. Ren, X. Liu, and M. Fu, "Consensus of multi-agent systems with general linear and lipschitz nonlinear dynamics using distributed adaptive protocols," *IEEE Transactions on Automatic Control*, vol. 58, no. 7, pp. 1786–1791, 2013.

[7] Y. Hong, J. Hu, and L. Gao, "Tracking control for multi-agent consensus with an active leader and variable topology," *Automatica*, vol. 42, no. 7, pp. 1177–1182, 2006.

[8] Y. Hong, G. Chen, and L. Bushnell, "Distributed observers design for leader-following control of multi-agent networks," *Automatica*, vol. 44, no. 3, pp. 846–850, 2008.

[9] J. Hu and Y. Hong, "Leader-following coordination of multi-agent systems with coupling time delays," *Physica A: Statistical Mechanics and Its Applications*, vol. 374, no. 2, pp. 853–863, 2007.

[10] J. Hu and G. Feng, "Distributed tracking control of leader-follower multi-agent systems under noisy measurement," *Automatica*, vol. 46, no. 8, pp. 1382–1387, 2010.

[11] X. Wang, Y. Hong, J. Huang, and Z. Jiang, "A distributed control approach to a robust output regulation problem for multi-agent linear systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 12, pp. 2891–2895, 2010.

[12] Y. Fei, S. Bo, O. Lin, and Z. Dong, "Disturbance observer-based control for consensus tracking of multi-agent systems with input delays from a frequency domain perspective," *Systems And Control Letters*, vol. 114, pp. 66–75, 2018.

[13] Z. Li, W. Ren, X. Liu, and M. Fu, "Distributed containment control of multi-agent systems with general linear dynamics in the presence of multiple leaders," *International Journal of Robust and Nonlinear Control*, vol. 23, no. 5, pp. 534–547, 2013.

[14] M. Qian and M. Ying, "Distributed containment control of linear multi-agent systems," *Neurocomputing*, vol. 133, pp. 399–403, 2014.

[15] S. Sheng, Z. Xin, and Z. Gang, "Formation-containment control of multi-robot systems under a stochastic sampling mechanism," *Science China Technological Sciences*, vol. 63, no. 6, pp. 1025–1034, 2020.

[16] H. Hamed, B. Mohammad, and B. Mahdi, "Containment control of heterogeneous linear multi-agent systems," *Automatica*, vol. 54, pp. 210–216, 2015.

[17] K. Zhen, S. John, and D. Warren, "Leader–follower containment control over directed random graphs," *Automatica*, vol. 66, pp. 56–62, 2016.

[18] W. Jiang, G. Wen, Z. Peng, T. Huang, and A. Rahmani, "Fully distributed formation-containment control of heterogeneous linear multiagent systems," *IEEE Transactions on Automatic Control*, vol. 64, no. 9, pp. 3889–3896, 2019.

[19] S. Sheng, Y. Yan, C. Xia, and H. Bo, "Necessary and sufficient conditions for consensus in fractional-order multiagent systems via sampled data over directed graph," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 4, pp. 2501–2511, 2021.

[20] D. V. Dimarogonas, E. Frazzoli, and K. H. Johansson, "Distributed event-triggered control for multi-agent systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 5, pp. 1291–1297, 2012.

[21] G. Eloy, C. Can, Y. Han, and A. Panos, "Decentralised event-triggered cooperative control with limited communication," *International Journal of Control*, vol. 86, no. 9, pp. 1479–1488, 2013.

[22] M. Yu and C. Wen, "Event based agreement protocols for multi-agent networks," *Automatica*, vol. 49, no. 7, pp. 2125–2132, 2013.

[23] Y. Peng, R. Wei, L. Dong, and C. Sheng, "Decentralized event-triggered consensus for linear multi-agent systems under general directed graphs," *Automatica*, vol. 69, pp. 242–249, 2016.

[24] H. Zhang, G. Feng, H. Yan, and Q. Chen, "Observer-based output feedback event-triggered control for consensus of multi-agent systems," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 9, pp. 4885–4894, 2014.

[25] W. Zhu, Z. Jiang, and G. Feng, "Event-based consensus of multi-agent systems with general linear models," *Automatica*, vol. 50, no. 2, pp. 552–558, 2014.
[26] E. Garcia, Y. Cao, and D. W. Casbeer, "Decentralized event-triggered consensus with general linear dynamics," *Automatica*, vol. 50, no. 10, pp. 2633–2640, 2014.
[27] Y. Cheng and V. Ugrinovskii, "Event-triggered leader-following tracking control for multivariable multi-agent systems," *Automatica*, vol. 70, pp. 204–210, 2016.
[28] X. Li and D. Qi, "Distributed event-triggered observer-based tracking control of leader–follower multi-agent systems," *Neurocomputing*, vol. 273, pp. 650–658, 2018.
[29] B. Cheng and Z. Li, "Fully distributed event-triggered protocols for linear multiagent networks," *IEEE Transactions on Automatic Control*, vol. 64, no. 4, pp. 1655–1662, 2019.
[30] S. Sheng, W. Xin, and Z. Gang, "Consensus of second-order hybrid multiagent systems by event-triggered strategy," *IEEE Transactions on Cybernetics*, vol. 50, no. 11, pp. 4648–4657, 2020.
[31] W. Zhu and Z. Jiang, "Event-based leader-following consensus of multi-agent systems with input time delay," *IEEE Transactions on Automatic Control*, vol. 60, no. 5, pp. 1362–1367, 2015.
[32] S. Sheng, S. Ping, and Z. Gang, "Semiglobal observer-based non-negative edge consensus of networked systems with actuator saturation," *IEEE Transactions on Cybernetics*, vol. 50, no. 6, pp. 2827–2836, 2020.
[33] W. Liu, C. Yang, Y. Sun, and J. Qin, "Observer-based event-triggered tracking control of leader-follower systems with time delay," *Journal of Systems Science and Complexity*, vol. 29, no. 4, pp. 865–880, 2016.
[34] W. Liu, C. Yang, Y. Sun, and J. Qin, "Observer-based event-triggered containment control of multi-agent systems with time delay," *International Journal of Systems Science*, vol. 48, no. 6, pp. 1217–1225, 2017.
[35] W. Chun, G. Qiang, X. Peng, and Z. Guang, "Delay-dependent distributed event-triggered tracking control for multi-agent systems with input time delay," *Neurocomputing*, vol. 333, pp. 200–210, 2019.

Hindawi
Complexity
Volume 2021, Article ID 9995030, 14 pages
https://doi.org/10.1155/2021/9995030

Research Article

A Novel Fast Image Stitching Method Based on the Combination of SURF and Cell

Qing An,[1] **Xijiang Chen**,[1,2] **and Shusen Wu**[3]

[1]*Artificial Intelligence School, Wuchang University of Technology, Wuhan 430223, China*
[2]*School of Safety Science and Emergency Management, Wuhan University of Technology, Wuhan 430070, Hubei, China*
[3]*State Key Laboratory of Materials Processing and Die & Mould Technology, School of Materials Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China*

Correspondence should be addressed to Xijiang Chen; cxj_0421@163.com and Shusen Wu; ssw636@hust.edu.cn

Received 2 April 2021; Accepted 10 June 2021; Published 18 June 2021

Academic Editor: Dan Selişteanu

The traditional image stitching method has some shortcomings such as double shadow, chromatic aberration, and stitching. In view of this, this paper proposes a power function-weighted image stitching method that combines SURF optimization and improved cell acceleration. First, the method uses the cosine similarity to preliminarily judge the similarity of the feature points and then uses the two-way consistency mutual selection to filter the feature point pairs again. Simultaneously, some incorrect matching points in the reverse matching are eliminated. Finally, the method uses the MSAC algorithm to perform fine matching. Then, the power function-weighted fusion algorithm is used to calculate the weight of the center point. The power function weight of the accelerated cell is used to perform the final image fusion. The experimental results show that the matching accuracy rate of the proposed method is about 11 percentage points higher than the traditional SURF algorithm, and the time is reduced by about 1.6 s. In the image fusion stage, this paper first selects images with different brightness, angles, resolutions, and scales to verify the effectiveness of the proposed method. The results show that the proposed method effectively solves the ghosting and stitching seams. Comparing with the traditional fusion algorithm, the time consumption is reduced by at least 2 s, the mean square error is reduced by about 1.32%~1.48%, and the information entropy is improved by about 0.98%~1.70%. The proposed method has better performance in matching accuracy and fusion effect and has better stitching quality.

1. Introduction

Image stitching technology includes image registration and fusion. The target of image stitching is to form a wide viewing angle and seamless panoramic image from multiple partially overlapping images. Simultaneously, it needs to ensure that the image is not distorted [1]. These images can be taken from different times, heights, viewing angles, and sensors. Image stitching technology is widely used in computer graphics, computer vision, virtual reality technology, medical image analysis, and other fields [2].

Image registration is the core of image stitching, which includes the extraction and matching of feature points. There are many ways to extract feature points, such as Harris, SIFT, SURF, FAST, and ORB. The SURF algorithm can not only adapt to image rotation and zooming but also maintain stability under the conditions of viewing angle changes, scale changes, and illumination changes. Therefore, the proposed method uses the SURF algorithm [3] to extract feature points. The SURF algorithm is an improvement algorithm of the SIFT (scale-invariant feature transform), and it is not affected by the image rotation, scale scaling, and brightness changes. Simultaneously, it has a faster calculation speed [4]. In 2012, Pang et al. proposed a fully radiation-invariant SURF algorithm, which combines the scale-invariant feature transform (SIFT), accelerated robust feature (SURF), and affine SIFT (ASIFT). Simultaneously, it uses the invariance of ASIFT's affine and advantage of high efficiency of SURF [5]. The improved SURF algorithm was proposed by using the positive and

negative characteristics of the Hessian matrix trace, which saved the time of feature point detection and matching [6]. Additionally, another improved SURF algorithm was proposed by reducing the SURF feature point descriptor from 64 to 36 dimensions. However, the algorithm is not universal and is only suitable for images with large rotation angles and low brightness [7]. In view of this, an improved SURF algorithm based on feature points was proposed on the basis of SURF. The algorithm uses Hessian matrix trajectory to extract the feature points of the image, which improves the speed and accuracy of the algorithm. However, it is complicated and relatively time-consuming [8]. In order to improve the efficiency of registration, a fast image stitching method based on the phase correlation and the improved SURF algorithm was proposed. The algorithm can not only reduce computational complexity but also reduce mismatches. However, it is also easy to eliminate some correct matching points, which is easy to cause stitching in the later image fusion process [9]. In 2021, the performance of the traditional Hitmap version is evaluated in similar GPU clusters. Fernandez-Fabeiro et al. proposed to extend this implementation by adding new functions to the latest version of Hitmap to support arbitrary load distribution of multinode heterogeneous GPU clusters to achieve images [10]. Pasha Hosseinbor et al. proposed 2D point-set registration algorithm for unlabeled data based on linear least squares and subsequently utilized for minutia-based fingerprint matching. The matcher considers all possible minutia pairings and iteratively aligns the two sets until the number of minutia pairs does not exceed the maximum number of allowable one-to-one pairings. The first alignment establishes a region of overlap between the two minutia sets, which is then (iteratively) refined by each successive alignment. After each alignment, minutia pairs that exhibit weak correspondence are discarded [11].

Image fusion plays an important role in image stitching. Because there might be changes in brightness, angle, and scale between images to be stitched, direct fusion will cause problems such as ghosting and stitching gap. A Markov Random Field (MRF) fusion model for image fusion was proposed. The salient structure of the input image is fused in the gradient domain, and then the final result is reconstructed by solving the Poisson equation. For the fusion images, the Poisson equation makes the gradient of the fusion image be close to the fusion gradient [12]. An improved method based on the relative orientation and small area fusion was proposed, which can select the appropriate stitching surface according to the half-angle correction method. Then, it selects a strip area in the middle part of the image overlap area as the transition area to conduct gradual fusion [13]. In 2017, an adaptive weighted image fusion method was proposed. The method designs an adaptive weighted image fusion strategy by combining the fuzzy theory idea based on curvelet transformation [14]. Additionally, the S-shaped nonlinear fusion strategy was proposed to conduct the image fusion stitching, which effectively solved the problems of stitching gap and ghosting [15]. Jung and Hong propose a method for quantifying the parallax level of the input images and clustering them accordingly. This facilitates a quantitative assessment of the various stitching methods for each parallax level. The parallax levels of the images are grouped based on the magnitude and variation in the planar parallax, as estimated with the proposed metric using matching errors and patch similarity [16]. This article has introduced a robust and reliable image stitching methodology (l, r-Stitch Unit), which considers multiple nonhomogeneous image sequences as input to generate a reliable panoramically stitched wide view as the final output. Besides, it has also introduced a novel convolutional-encoder-decoder deep-neural-network (l, r-PanoED-network) with a unique split-encoding-network methodology, to stitch noncoherent input left, right stereo image pairs [17]. Xue et al. proposed a high-quality fisheye image stitching algorithm (LLBI-AW). In the correction part, the traditional bilinear interpolation algorithm leads to blurred edges, and an interpolation algorithm based on latitude and longitude (LLBI) is proposed to solve this problem. In the fusion part, a weighted fusion algorithm based on arc function is proposed for the ghost problem in the overlapped area of the image after stitching. We establish an arc function to determine the weight of the left and right images and then perform the fusion calculation [18]. Vishwakarma and Bhuyan proposed an automatic classification algorithm and an image stitching algorithm based on Harris features of local differences. In the image mosaic process, the automatic classification results are distorted due to image noise and pseudoperiod. By using the structural similarity index to solve the above problems and replacing the standard Sobel edge detector with a local difference operation, the high time complexity problem of the conventional angle detector can be reduced [19]. Although the above methods have been relatively improved, some image stitching quality problems are still unsolved, and the stitching efficiency also needs to be improved.

Aiming at the problems of low matching accuracy and a large amount of calculation in the image registration process, this paper combines improved SURF algorithm and SURF optimization to improve the matching accuracy of image, and the main contributions of this paper are described as follows:

(1) First, the cosine similarity is used to detect the preliminary screening of the sign points. If the cosine similarity is greater than K value, the sign points are examined by two-way consistency detection, and the mismatch points are filtered out again.

(2) Second, the MSAC algorithm is proposed for the fine matching to determine the final matching pair.

(3) Third, aiming at problems such as ghosts and stitching seams that are likely to occur in the image fusion process, a cell-accelerated power function-weighted fusion algorithm is proposed, which uses the power function-weighted fusion algorithm to obtain the weight of the cell center point in the cell acceleration. Finally, an image stitching technology

with good stitching quality and high efficiency is finally realized according to the proposed method.

2. Optimize the SURF Algorithm

2.1. Extraction of Feature Points. The SURF algorithm uses the Hessian matrix $H(x,\sigma)$ to detect feature points, and each pixel can find a Hessian matrix. Thus, the Hessian matrix $H(x,\sigma)$ with coordinate x and scale σ can be expressed as

$$H(x,\sigma) = \begin{pmatrix} l(x,\sigma) & l_{xy}(x,\sigma) \\ l_{xy}(x,\sigma) & l_{yy}(x,\sigma) \end{pmatrix}, \quad (1)$$

where x is the image pixel coordinate and σ is the scale of the image

The SURF algorithm uses a box filter instead of the Gaussian second-order partial derivative. A multiscale spatial function is constructed by changing the size of the box filter. Then, the multiscale spatial function is convolved with the original image in different directions.

The convolution sum of l_{xy}, l_{xy}, l_{yy} in equation (1) can be expressed as D_{xx}, D_{xy}, D_{yy}, so the Hessian matrix can be simplified as

$$\mathrm{Det}(\mathrm{Hessian}) = D_{xx}D_{yy} - (\omega D_{xx})^2, \quad (2)$$

where ω is the weight coefficient and it is generally 0.9.

First, the Hessian matrix is used to find the local maximum pixel point, and the point is marked as the interest point. The corresponding two-dimensional array subscript is the position of the interest point in the image. Then, the Harr wavelet is used to determine the main direction of the interest point. Finally, a 64-dimensional SURF characteristic descriptor is constructed.

2.2. Improved Feature Point Matching. For the traditional SURF algorithm, there are two steps to determine the matching degree between the feature points. First, the Euclidean distance between the two feature points is calculated, and then the signs of the Hessian matrix traces are calculated, which are used to determine whether the feature points match or not. Although it has a great advantage over the SIFT algorithm, there are still a large number of mismatched points affecting the stitching effect, and it takes a long time and is less efficient. In view of the above problems, this paper proposes a matching method that combines cosine similarity, two-way consistency, and MSAC algorithm. Cosine similarity and two-way consistency are used for rough matching, and then the MSAC algorithm is used for fine matching.

2.2.1. Cosine Similarity. Cosine similarity means the degree of similarity between feature points. According to the coordinates of feature points, the cosine between the vectors is determined. The cosine value means the similarity between the two feature points. The smaller the cosine value, the larger the angle between the two feature points, the smaller the matching similarity, and vice versa.

First, the Euclidean distance is used to select the preliminary feature point, and then the cosine phase velocity function is used to further determine the fine feature point. If the cosine value of the two vectors is greater than the threshold K, it will be retained, and vice versa. The cosine similarity of vector a and b is obtained by

$$S(a,b) = \frac{(a,b)}{\|a\| \cdot \|b\|} = \frac{\sum_{i=1}^{n} a_i b_i}{\left(\sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2\right)^{1/2}}. \quad (3)$$

The cosine similarity is used to test four types of images. The first type of image generates scaling and rotation, the second has the variation of brightness and contrast, the third has the variation of blur, and the fourth has the variation of viewing angle. According to Table 1, we select the average as the final threshold $K = 0.975$.

2.2.2. Two-Way Consistency. After matching with cosine similarity, there will still be some mismatches. In order to improve the accuracy of matching, this paper uses a two-way consistency constraint to eliminate mismatched point pairs. First, the matching pair is saved after the one-way matching. Second, the relationship between the matching graph and the original graph will be exchanged. Then, we perform reverse matching and select the common matching pair as the final matching result. The single-direction matching point pair is deleted. The relative position relationship calculation of the matching point pairs can effectively improve the accuracy of the matching result. The two-way consistency constraint matching is shown in Figure 1. Assume the homonymy points of point P_i are P_i', and conduct the reverse registration. The homonymy point of P_i' is P_i, which meets the two-way consistency constraint. Otherwise, the point P_i is regarded as a mismatched point and should be eliminated.

2.2.3. MSAC Algorithm. The MSAC algorithm is an optimization of the RANSAC algorithm. By modifying the RANSAC cost function, the issue of RANSAC being sensitive to the threshold was resolved. The paper uses the MSAC algorithm to conduct the image registration. The algorithm is used to eliminate the mismatched point pairs in the rough matching. The holography matrix H between two adjacent images is calculated according to the correct matching point pairs. Finally, the matrix to each point in the image to be registered is mapped to the reference coordinate system, and the bilinear interpolation method is used to achieve precise image matching.

3. Improved Image Fusion

Traditional image fusion algorithms perform direct fusion on the registered images, which will produce very obvious stitches and color differences, and severely cause excessive uneven color or artifacts. In addition, some improved methods have been proposed, but these algorithms are cumbersome and time-consuming. In view of this, this paper

Table 1: K value test result.

	Rotate and zoom	Brightness	Blurry	Perspective
Optimal threshold	0.97	0.96	0.98	0.99

proposes a cell-accelerated power function-weighted fusion algorithm.

3.1. Gradually Weighted Fusion Algorithm. The gradual weighted fusion algorithm is an improvement of the traditional fusion algorithm. According to the traditional single weighting coefficient, the weight is determined by the distance between the pixel points of the overlapping area and the boundary of the overlapping area. Thereby, a slow and smooth transition is achieved and the stitching gap of the stitched image is eliminated [20].

$$I(x,y) = \begin{cases} I_1(x,y), & (x,y) \in I_1, \\ \omega_1 I_1(x,y) + \omega_2 I_2(x,y), & (x,y) \in I_1 \cap I_2, \\ I_2(x,y), & (x,y) \in I_2, \end{cases} \tag{4}$$

where I_1 is the reference image; I_2 is the original image; $\omega_1 = (x_2 - x/x_2 - x_1), \omega_1 + \omega_2 = 1\,(0 < \omega_1, \omega_2 < 1)$.

Figure 2 shows that the gradual weighted fusion algorithm is easy to implement, and the stitching speed is fast. The weights ω_1 and ω_2 vary linearly from 0 to 1. This algorithm is better than the direct fusion algorithm. However, the overlapping regions have shadow after registration because of the same intersection of the two weighting functions and the inflection point of the fusion function. Moreover, the effect of the proposed method is not good when the brightness of the two stitched images differs greatly. It can be seen that the gradual and fade-out weighting method is no longer suitable for most image fusion situations.

3.2. Cell-Accelerated Power Function-Weighted Fusion Algorithm. In order to solve the problems of stitching seams and ghosts in image stitching due to brightness differences and noise generated in image acquisition, this paper proposes a cell-accelerated power function-weighted fusion algorithm, which not only solves the ghosting problem but also saves time and improves stitching efficiency.

According to the Weber–Fechner law [21], the human eye perception response shows a power function nonlinear change law as the background brightness stimulus becomes larger. This law is more appropriate to the trajectory of plant growth. Therefore, the paper uses the Logistic and Cubic function commonly used in plant growth models to derive the power function-weighted fusion model, as shown in equations (5) and (6) [15]:

$$\frac{dx}{dt} = r\left(1 - \frac{x}{x_{\max}}\right)x, \tag{5}$$

$$x(t) = -2\frac{\Delta x}{T^3}t^3 + 3\frac{\Delta x}{T^2}t^2 + x_{\min}, \tag{6}$$

where t is the time, T is the plant growth cycle, and $t \in [0, T]$, $x_{\min}, x_{\max}$ is the initial value and the maximum value, respectively. When $t = T$, the plant growth is close to the maximum value.

According to equations (5) and (6), the weighting function of the power function model is obtained:

$$x(t) = -\frac{r\Delta x}{3T^3}t^3 + \frac{r\Delta x}{2T^2}t^2 + x_{\min}. \tag{7}$$

According to equation (7), appropriate r is selected that is more suitable for the weighting function of the power function model.

$$\begin{aligned} x_{\min} &= 0, \\ r &= 6, \\ \Delta x &= x_{\max} = 1, \\ T &= 1, \\ x(t) &= 6\left[-\frac{1}{3}t^3 + \frac{1}{2}t^2\right], \quad t = [0,1]. \end{aligned} \tag{8}$$

Finally, combined with the change trend of the power function weighting function ω_1 in this paper from 1 to 0, equation (8) is adjusted to obtain equation (9). The final improved weighted fusion process is obtained, as shown in Figure 3(a).

$$\begin{aligned} \omega_1 = &-2\left(\frac{x - x_1}{x_2 - x_1}\right)^3 + 3\left(\frac{x - x_1}{x_2 - x_1}\right)^2 \\ &- 2\left(\frac{x - x_1}{x_2 - x_1}\right) + 1, \quad x_1 \le x \le x_2. \end{aligned} \tag{9}$$

In order to clearly express the changing trend of ω_1, the derivation of equation (9) is taken.

$$\omega_1' = -\frac{6}{q^3}x^2 + \left(\frac{12x_1}{q^3} + \frac{6}{q^2}\right)x - \left(\frac{6x_1^2}{q^3} + \frac{6x_1}{q^2} + \frac{2}{q}\right). \tag{10}$$

From the parabolic properties, we can get that when $x = 0$, $y < 0$, and the coefficient before x^2 is less than 0, the one-variable quadratic equation discriminant $\Delta = -(12/q^4) < 0$ shows that there is no intersection with the x-axis, so you can draw a rough image of ω_1', as shown in Figure 3 on the right.

In combination with Figure 3, it can be seen that, in the process of approximating from the edge of the coincident

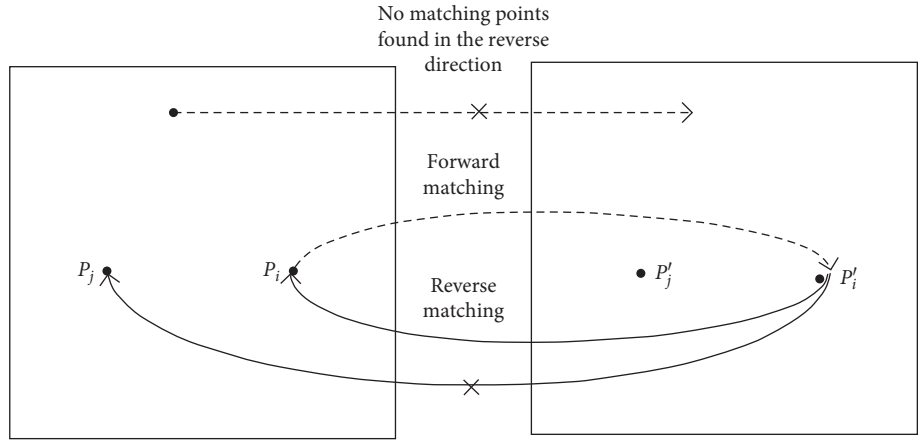


Figure 1: Bidirectional consistency constraint matching.

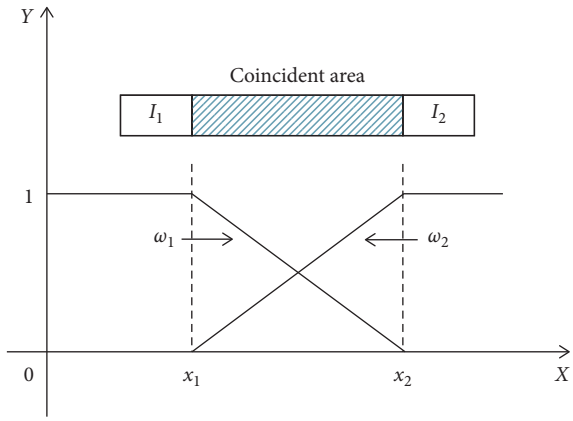


Figure 2: Gradual weighted fusion.

area to the middle position, the derivative ω_1' of the power function weighting function first increases and then decreases. This makes the power function weighting function ω_1 follow the first fast, then slow, and change. The fast speed means that, during the refusion process of the left and right images, the proportion of pixels will also change relatively, thus alleviating the problem of differences in the fusion process of two images with large chromatic aberration and also solving the problem of stitching seams.

Since each pixel in the overlapping area needs to calculate a weighting coefficient and the weighting coefficient does not change linearly, this will cause the calculation to be more complicated and time-consuming. In response to this situation, a cell-based integration acceleration strategy is proposed. The specific steps are as follows:

(1) Divide the overlapping area in the fusion into as much as possible $n \times n (n \geq 1)$ small squares of equal size, as shown in Figure 4, which are defined as cells.

(2) Take the center point of the cell or the point $P(x, y)$ close to the center, calculate the value of this point by equation (9), and use the weighting coefficient of point P as the weighting coefficient of each pixel in the entire cell.

(3) Points that are close to the edge of the overlapping area and do not participate in the cell need to calculate the weighting coefficient separately. The appropriate value of n can be selected according to the size and shape of the overlapping area between the images to be stitched.

As a result, the number of calculation weighting coefficients will be reduced exponentially, and the fusion efficiency will be significantly improved.

4. Experimental Results and Analysis

The experiment in this paper uses Matlab 2014 for programming and image stitching under a Windows 10 system with an Intel i5 processor and a memory of 12G, which verifies the feasibility and superiority of the calculation in this paper. The first is the comparison between the traditional SURF algorithm and the proposed method in the image registration stage, and the second is the comparison between the image fusion based on this algorithm and the other two algorithms in the image fusion stage.

4.1. Registration Analysis of Feature Points. In order to verify the universality of the proposed method, outdoor parking lots and indoor hall images are used to conduct the image stitching, as shown in Figures 5 and 6. The proposed method extracts and matches the feature points of adjacent images, compared with the traditional SURF algorithm.

According to the proposed method and the traditional SURF algorithm, the feature points of the outdoor parking lot image are extracted and the feature points are matched, as shown in Figure 5. The traditional SURF algorithm will produce a small number of mismatched point pairs, as shown in the red circle and purple square in Figure 5(a). The red circle marks the ground in the two images to be stitched. The reason for the mismatch is that the Euclidean distance between the ground feature points is extremely small, which results in high similarity between the feature points and mismatches; the purple square in Figure 5 shows the matching results of the back of two white cars, the same matrix trace symbols are generated after Hessian matrix

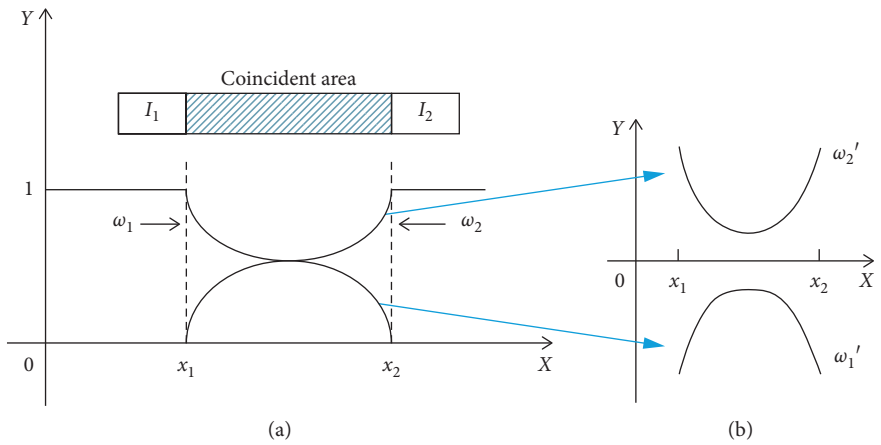


Figure 3: Power function-weighted fusion process.

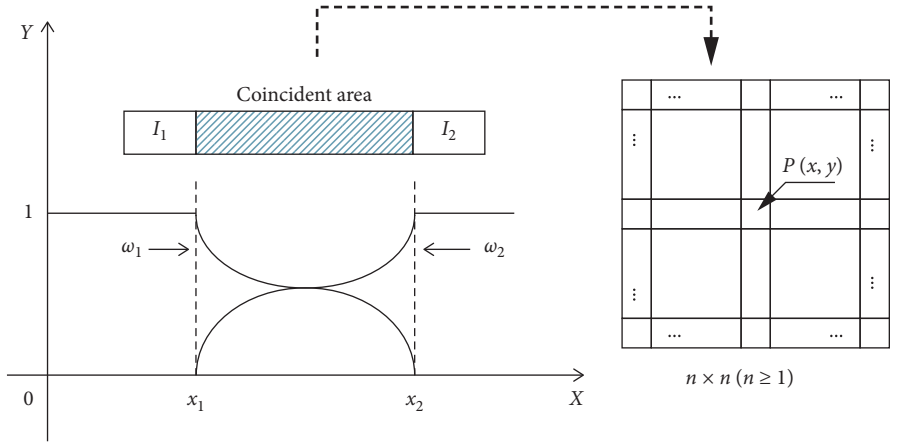


Figure 4: Cell acceleration power function weighting coefficient.

Figure 5: Image feature point matching of outdoor parking lot. (a) Traditional SURF algorithm matching. (b) Algorithm matching in this paper.

calculation, which leads to mismatches. In Figure 5(b), there is no mismatch phenomenon in the matching algorithm used in this paper, and more stable interest points are generated as matching points, which improves the efficiency of the algorithm.

Similarly, according to the proposed method and the traditional SURF algorithm, the feature points of the indoor hall image are extracted and the feature points are matched. The result is shown in Figure 6. The red circle in Figure 6(a) shows the mismatch between the left pillar of the image to be stitched and the right wall of the image to be stitched. The left oblique pillar and the wall of the image to be stitched are mismatched, as shown in the purple box of Figure 6(a). In Figure 6(b), the proposed algorithm eliminates mismatched point pairs.

In order to ensure the accuracy of the conclusion, this paper analyzes the traditional algorithm and the algorithm of this paper to further verify the superiority of the proposed method. In this paper, the correct matching rate [22] is used as the evaluation criterion:

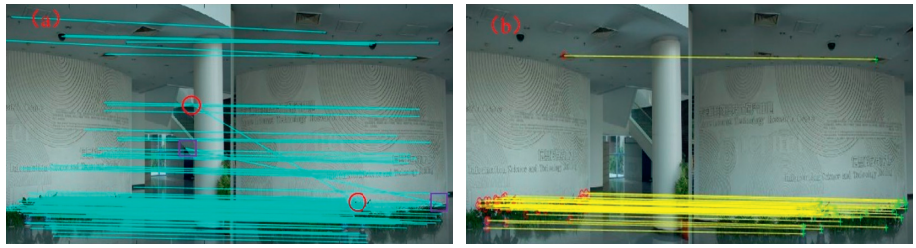


Figure 6: Image feature point matching in indoor lobby. (a) Traditional SURF algorithm matching. (b) Algorithm matching in this paper.

$$1 - \text{precision} = \frac{\text{falmatches}}{\text{correctmatches} + \text{falmatches}}, \tag{11}$$

where precision is the correct rate; falmatches is the wrong match; correctmatches is the correct match.

Use equation (11) to calculate the matching precision of outdoor parking lots and indoor halls, as shown in Table 2.

From Table 2, it can be seen that the matching precision of outdoor parking images by the proposed method is greater than 80%. However, the matching precision of indoor hall image is 76.8%, and it is lower than outdoor parking. The reason for this matching result is that the feature of outdoor parking is more obvious than that of indoor hall image. The same phenomenon appears in the proposed method. But, the matching precision of the traditional SURF algorithm is lower than the proposed method, and it is about 11% lower than that of the proposed method. It illustrates that the proposed method reduces the wrong match and improves the correct match. Additionally, the time of image match by the proposed method is smaller than the traditional SURF algorithm, and the registration effectiveness of the proposed method is superior to the traditional SURF algorithm. Therefore, the proposed method has obvious advantages in both speed and accuracy of image registration.

4.2. Analysis of Image Fusion. In order to verify the effectiveness of the proposed method in terms of image fusion, this paper uses the improved SURF matching algorithm based on the selection of the progressive gradual-out weighted fusion algorithm, and the algorithm used in [23] is compared with the proposed method for multiple experiments. In order to verify the universality of the proposed method, images with different brightness, different angles, different resolutions, and different scales are selected for fusion.

4.2.1. Effect Analysis of Fusion Results

(1) Different Brightness. For the cameras deployed in different positions of the parking lot, the adjacent cameras are used to collect parking lot images at different times, as shown in Figure 7. Figure 7(a) is a parking lot image taken on a cloudy day, and Figure 7(b) is a parking lot image taken on a sunny day. It can be seen that the brightness of the two images is obviously different.

Two images with different brightness are extracted and matched using the optimized SURF algorithm. First, the Hessian matrix is used to detect the feature points to form a 64-dimensional feature descriptor. Then, the cosine similarity is used to preliminarily judge the similarity between the feature points. The two-way consensus mutual selection is used to filter the feature point pairs again, and the point pairs with reverse matching errors or reverse nonmatching points are eliminated. The MSAC algorithm is used for fine matching and the final feature point is determined. Finally, the proposed method is compared with the other two algorithms, and the experimental results are shown in Figure 8.

Figure 8(a) shows the result of image fusion by the traditional gradual weighted fusion algorithm. From Figure 8(a), it is visible that there are two obvious stitching gaps, and the chromatic aberration is serious after the fusion. The red circle has been marked. Figure 8(b) shows that the stitching gap of the image to be stitched is obviously eliminated. After the fusion, there is still a slight chromatic aberration. Figure 8(c) shows that the proposed method effectively solves the stitching and chromatic aberration.

(2) Different Angles. A camera at a fixed position in the indoor hall at different rotation is used to take images of the indoor hall with overlapping areas, as shown in Figure 9. Figure 9(a) shows the angle shot facing the hall, and Figure 9(b) shows the image obtained by rotating the camera 30° left and right and 15° up and down.

The Hessian matrix is used to extract feature points from two images to be stitched at different angles, and the feature points are matched according to Section 2.2 of this article to eliminate mismatched points. Finally, the proposed method is compared with the other two algorithms, and the experimental results are shown in Figure 10.

There are obvious stitching seams by the traditional gradual-out algorithm, as shown in the blue mark in Figure 10(a). From Figure 10(a), it is clearly visible that there is a fusion gap. Because the weighting coefficient of this method shows a nonlinear trend, it can solve the seam problem well. Figure 10(c) is the fusion result obtained by the proposed method. It is similar to the fusion algorithm in the literature [23] in terms of visual effects, but the stitching efficiency will be greatly improved, as mentioned as follows.

(3) Different Resolutions. The images to be matched in this experiment are from the Internet, and the original images to be stitched are set to different resolutions, as shown in

Table 2: Matching precision of different methods.

Different methods	Outdoor parking		Indoor hall	
	Precision (%)	Time (s)	Precision (%)	Time (s)
Traditional SURF algorithm	74.07	3.59	65.52	3.58
Proposed method	85.18	2.01	76.80	2.02

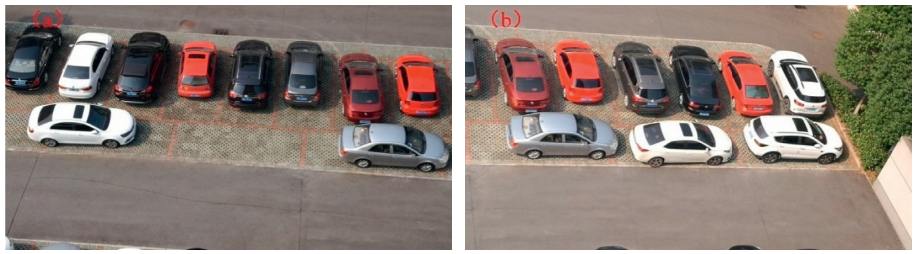


Figure 7: Two original images with different brightness: (a) the first image; (b) the second image.

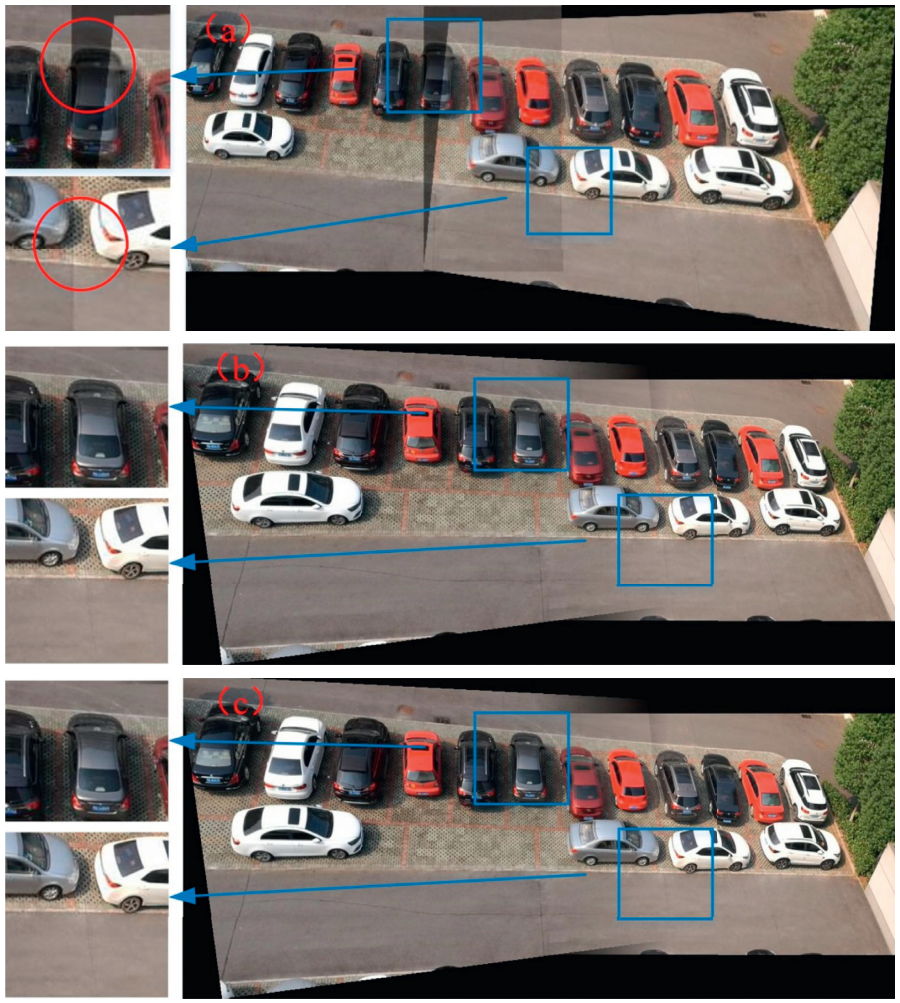


Figure 8: Comparison of the results of three fusion algorithms on different brightness. (a) Gradual and progressive weighted fusion algorithm. (b) Fusion algorithm in [23]. (c) The proposed method.

Figure 9: Two original images with different angles: (a) the first image; (b) the second image.

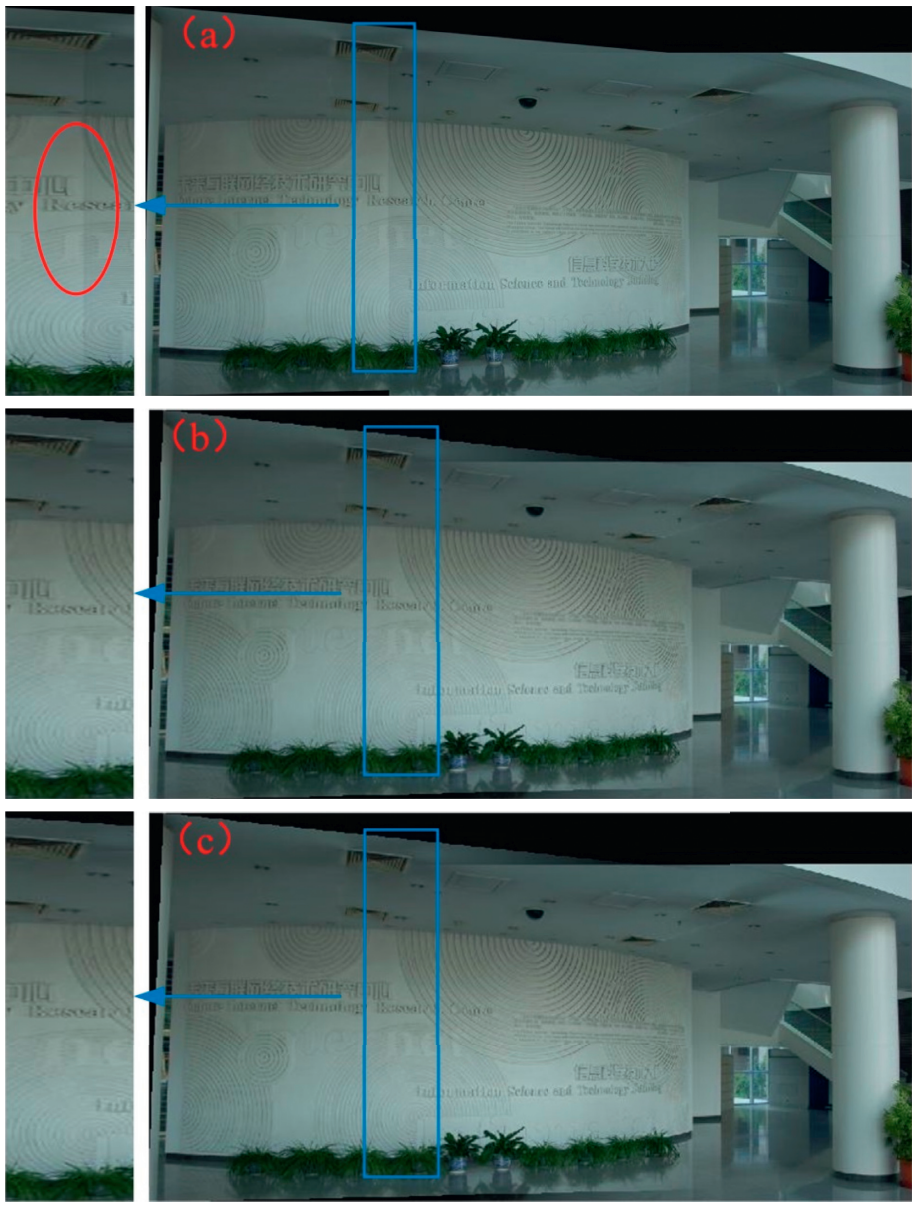


Figure 10: Comparison of the results of three fusion algorithms for different angles. (a) Gradual and progressive weighted fusion algorithm. (b) Fusion algorithm in [23]. (c) Fusion proposed method.

Figure 11. The resolution of Figure 11(a) is $720 * 568$, and the resolution of Figure 11(b) is $650 * 513$.

The image to be stitched is processed for image matching using the content of the first chapter of this article, and the wrong matching point pairs are eliminated. Then, image fusion is performed according to the method proposed in Section 2 of this article, and the fusion result is shown in Figure 12.

Figure 11: Two original image sequences: (a) the first image; (b) the second image.

Figure 12: Comparison of the results of three fusion algorithms on images with different resolutions. (a) Progressive and progressive weighted fusion algorithm. (b) Fusion algorithm in [23]. (c) The proposed method.

Figure 12(a) shows three stitching seams by the traditional gradual fusion algorithm. The blue square mark on the upper left is enlarged to show that the roof shows that there is obvious misalignment. The bottom left and right of the blue square mark shows that there is a wide stitching gap. The above errors are marked with red circles. Figure 12(b) shows the effect of the fusion algorithm used in [23]. From the left blue box, it can be seen that the algorithm eliminates the problem of misalignment and stitching. Figure 12(c) shows the result of the proposed method, which can also effectively eliminate the problem of misalignment and stitching gaps.

(4) Different Scales (Different Heights). This article uses drones to obtain images to be stitched with overlapping areas by setting different flying heights, as shown in Figure 13. Figure 13(a) is taken at a height of 90 m, and Figure 13(b) is taken at a height of 100 m. Therefore, these two images have different scale, and they will have different resolution. Different resolution images have different pixel

Figure 13: Two original image sequences of different scales. (a) Image 1. (b) Image 2.

size. Therefore, the pixels of the two different resolution images cannot merge together. The different resolution of the two images will cause matching seams.

The original image taken by the drone is first preprocessed to solve the effect of noise generated during the shooting process. Two images with overlapping areas are selected for feature point extraction and matching according to the matching algorithm proposed in this paper. Then, the processed image is fused using the gradual gradual-out algorithm, the fusion algorithm used in literature [23], and the proposed method. The comparison of the experimental results is shown in Figure 14.

Figure 14 shows the stitching effect of different scale pictures taken by drones. Figure 14(a) is the effect of the traditional gradual weighted fusion algorithm. Enlarging the area in the red box, it can be seen that the fused image is slightly misaligned. There are obvious gaps inside the blue box. There are slight gaps in the red area, but the stitching gaps at the blue mark are effectively eliminated, and there is no dislocation. Figure 14(c) is the result of the proposed method. The problem of stitching misalignment in the red area is effectively solved, and the stitching gap in the blue area is effectively eliminated, which solves the ghost or stitching caused by the incorrect weight ratio during the fusion process.

4.2.2. Quality Analysis of the Fusion Image. From the image stitching of different brightness, different angles, different resolutions, and different scales, it is impossible to clearly distinguish that the improved proposed method is better than the fusion algorithm [23]. Therefore, the square error and information entropy are used to evaluate the effectiveness of the proposed method.

(1) Time consumption: as shown in Figure 15, since the weighting coefficients are linearly related, the algorithm will take less time. After 4 sets of experiments, it can be seen that the total time consumed by the algorithm used in [23] is more than 2 seconds longer than that of the asymptotic algorithm. This is because the weighting coefficients used in [23] are more complicated. Since the proposed method uses cell to accelerate the weighting coefficient of the power function, it will double the calculation of the weighting coefficient of the power function. Therefore, the total time consumption of the proposed method is more than 2 s less than that of the fusion algorithm in the literature [23], and the total consumption of the gradual-out weighted fusion algorithm time is relatively close.

(2) Mean squared error (MSE): the smaller the MSE is, the closer it is to the original image and the better the fusion effect is. As shown in the histogram of Figure 16, the MSE of the proposed method is significantly smaller than the gradual-out weighted fusion method and slightly smaller than the algorithm used in the literature [24]. Compared with the four sets of experimental data, it is reduced by 1.32%, 1.48%, 1.39%, and 1.33%. The algorithm used in [23] is only about 1% smaller than the gradual-out algorithm. From the comparison of the mean square error data, it can be seen that the proposed method not only repairs the fusion seam but also significantly reduces the error with the original image.

(3) Information entropy: it is used to describe the average amount of information of an image and measure the richness of image information [25]. The greater the information entropy, the greater the amount of information contained in the fusion result image.

$$E = -\sum_{i=0}^{L-1} P(i)\log_2 P(i), \tag{12}$$

where $P(i)$ is the ratio of the number of pixels with the gray value i to the total number of images and L is the total number of gray levels.

Figure 17 shows that the information entropy of the gradual-out weighted fusion algorithm is almost equal to that of the algorithm used in [23]. Comparing with the other two algorithms, it can be concluded that, compared with the gradual-out algorithm, the proposed method has increased by 1.70%, 0.98%, 1.42%, and 1.04% in the four groups of experimental information entropy comparisons. Comparing with the gradual-out weighted fusion algorithm, the information entropy of the algorithm in [23] is increased by 0.17%, 0.16%, 0.09%, and 0.13%. After the image is stitched, the proposed method improves the image information and

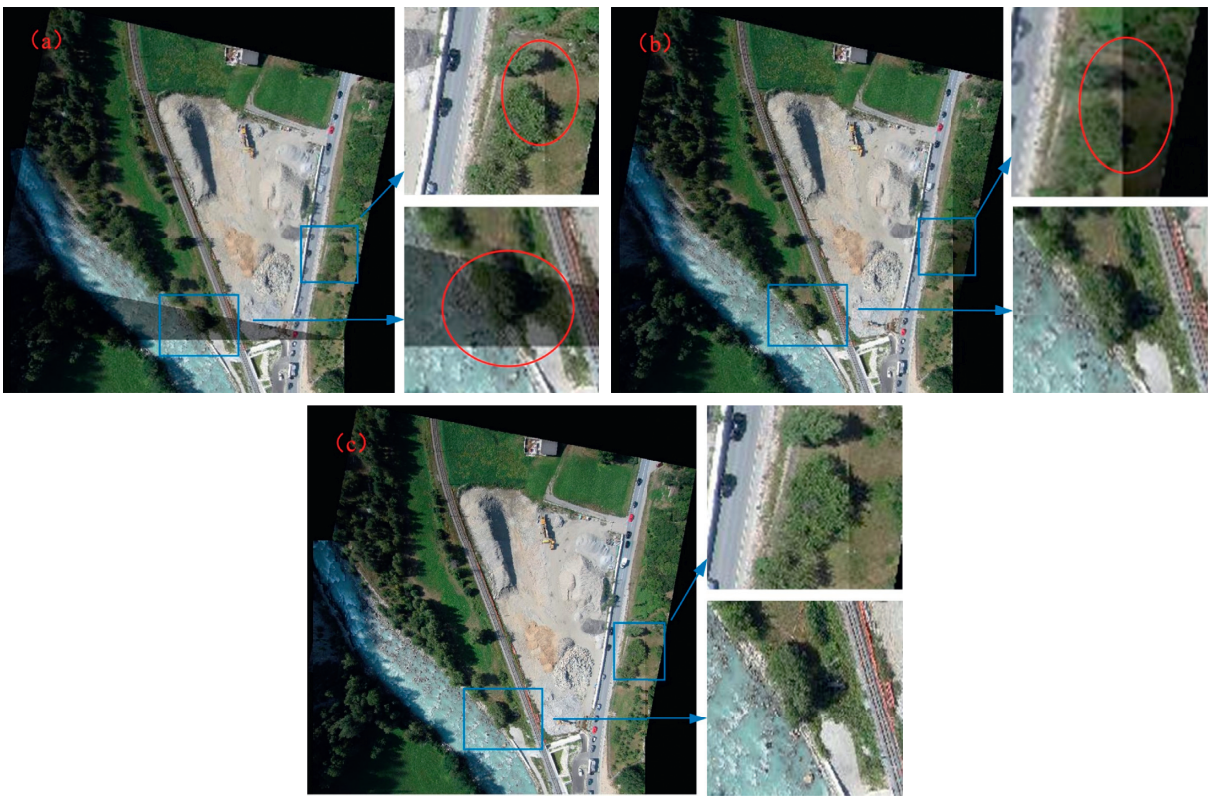


Figure 14: Comparison of the results of three fusion algorithms on images of different scales. (a) Gradual and progressive weighted fusion algorithm. (b) Fusion algorithm in [23]. (c) The proposed method.

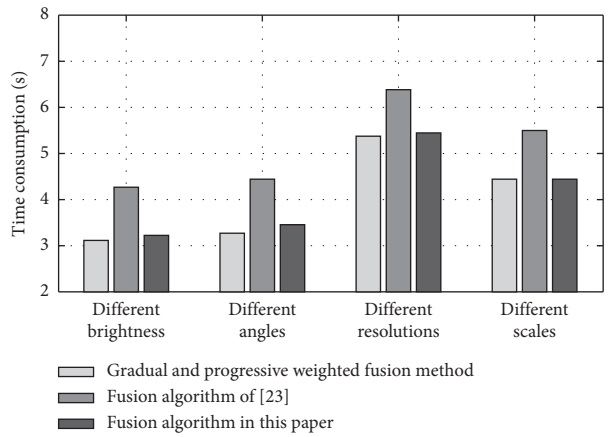


Figure 15: Time consumption of 4 groups of experiments.

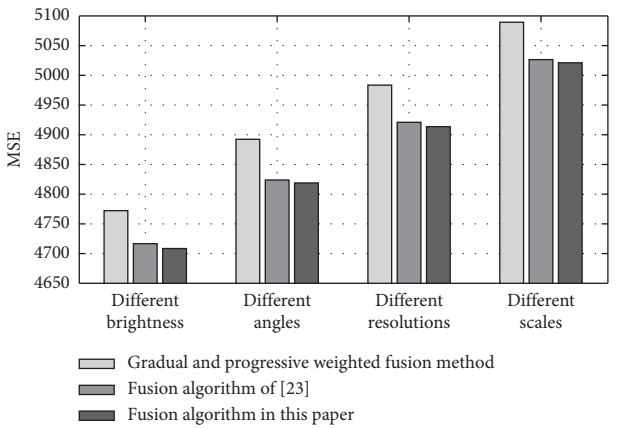


Figure 16: MSE of different methods.

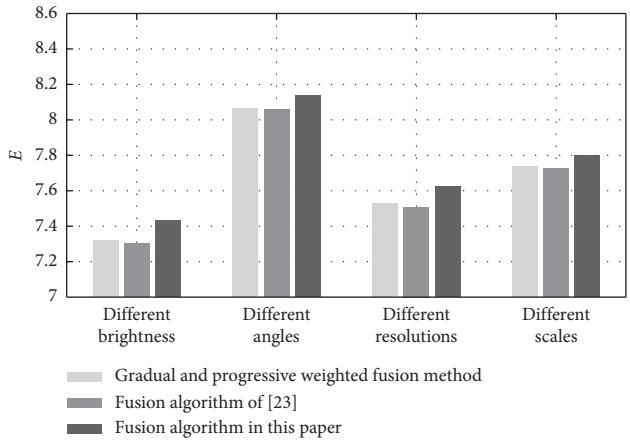


Figure 17: Information entropy of different methods.

eliminates ghosting. Simultaneously, the overall image information is significantly improved.

According to the above analysis, the proposed method can not only be applicable to changes in brightness, angle, resolution, and scale but also solve the problems of ghosting and stitching that often occur in traditional gradual-out algorithms. The proposed method is less time-consuming and has better detailed descriptions and richer information content. Compared with the other two fusion algorithms, the proposed method can obtain higher stitching and fusion quality.

5. Conclusion

The proposed method uses cosine similarity, two-way consistency selection, and MSAC algorithm to match feature points, which effectively solves the misidentification of matching points. In the image fusion stage, this paper proposes a power function-weighted fusion algorithm based on cell acceleration. It not only solves the ghosting and stitching in traditional fusion algorithms but also improves the efficiency of overall image stitching. Through the comparison of several sets of experimental data results, it can be shown that the proposed method has a higher correct rate of feature point matching than the traditional SURF algorithm, and it takes less time. In the image fusion stage, the three indicators of total time consumption, mean square error, and information entropy are used for evaluation. Compared with the progressive out weighted fusion algorithm, the total time consumed by the proposed method is at least 2 s, and the mean square error is reduced by about 1.32%~1.48%. Information entropy increased by about 0.98%~1.70%. The proposed method has better performance in matching accuracy and fusion effect. Simultaneously, it has better stitching image quality and universality than traditional fusion algorithms.

In the future, the registration and fusion of very low-resolution images should be focused on. Simultaneously, the key point of image stitching is the determination of image feature points. The effectiveness of the determination of image feature points will directly determine the speed of image stitching. Additionally, the precision of image feature points will influence the precision of image registration. Therefore, we will conduct the research of fast and accurate determination of image feature points and the fast stitching of video streams.

Data Availability

All the data generated or analyzed during this study are included within this article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was financially supported by the Youth Science Foundation Project (51706165).

References

[1] W. Yao and H. Hu, "Adaptive rotating image seamless stitching algorithm based on SIFT," *Journal of Nanjing University of Science and Technology: Natural Science Edition*, vol. 43, no. 2, pp. 169–194, 2019.

[2] Z. L. Lan, Z. Li, and W. Li, "Research on mosaic of lane image sequence based on SURF and optimal stitching," *Journal of Chongqing Jiaotong University (Natural Science Edition)*, vol. 38, no. 10, pp. 13–18, 2019.

[3] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[4] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[5] Y. Pang, W. Li, Y. Yuan, and J. Pan, "Fully affine invariant SURF for image matching," *Neurocomputing*, vol. 85, no. 1, pp. 6–10, 2012.

[6] S. Q. Gao, X. J. Tan, and C. X. Huang, "An improved image registration algorithm based on SURF," *Journal of PLA University of Science and Technology (Natural Science Edition)*, vol. 4, pp. 372–376, 2014.

[7] Y. D. Wang and X. H. Li, "Improved SURF fast image matching," *Journal of Changchun University of Technology*, vol. 37, no. 2, pp. 43–46, 2016.

[8] G. Wang, L. Liu, and Y. Zhang, "Research on scalable real-time image mosaic technology based on improved SURF," *Journal of Physics Conference Series*, vol. 1069, Article ID 012162, 2018.

[9] P. F. Liu and R. X. Gao, "Image stitching method based on phase correlation method and improved SURF algorithm," *Software Guide*, vol. 18, no. 11, pp. 157–160, 2019.

[10] J. Fernández-Fabeiro, A. Gonzalez-Escribano, and D. R. Llanos, "Distributed programming of a hyperspectral image registration algorithm for heterogeneous GPU clusters," *Journal of Parallel and Distributed Computing*, vol. 151, no. 6, pp. 86–93, 2021.

[11] A. Pasha Hosseinbor, R. Zhdanov, and A. Ushveridze, "An unsupervised 2D point-set registration algorithm for unlabeled feature points: application to fingerprint matching," *Pattern Recognition Letters*, vol. 100, pp. 137–143, 2017.

[12] S. Jian, Y. Z. Hong, B. X. Zong, and Z. H. Chong, "Poisson image fusion based on Markov random field fusion model," *Information Fusion*, vol. 14, no. 3, pp. 241–254, 2013.
[13] B. X. Du, "Video image stitching based on relative orientation and small area fusion," *Journal of Computer Applications*, vol. 35, no. 1, pp. 220–223, 2015.
[14] P. B. Durga and D. H. Ravindra, "Multi-filtering based edge preserving image fusion technique," *International Journal of Services Technology and Managemen*, vol. 23, no. 4, pp. 49–60, 2017.
[15] X. Q. Gao, F. Yang, H. R. Fan et al., "S-type fusion stitching of remote sensing images under low error matching strategy," *Progress in Laser and Optoelectronics*, vol. 56, no. 21, pp. 328–336, 2019.
[16] K. Jung and J. Hong, "Quantitative assessment method of image stitching performance based on estimation of planar parallax," *IEEE Access*, vol. 9, no. 1, pp. 6152–6163, 2021.
[17] P. K. Chilukuri, P. Padala, P. Padala, V. S. Desanamukula, and P. R. Pvgd, "L, r-stitch unit: encoder-decoder-CNN based image-mosaicing mechanism for stitching non-homogeneous image sequences," *IEEE Access*, vol. 9, no. 1, pp. 16761–16782, 2021.
[18] L. Z. Xue, Z. J. Juan, and L. R. Hai, "A high-quality stitching algorithm based on fisheye images," *Optik*, vol. 238, Article ID 166520, 2021.
[19] A. Vishwakarma and M. K. Bhuyan, "Image mosaicking using improved auto-sorting algorithm and local difference-based harris features," *Multimedia Tools and Applications*, vol. 79, no. 1, pp. 23599–23616, 2020.
[20] H. Z. Xu, S. C. Li, Y. H. Ji et al., "Research on panoramic camera image mosaic method based on feature point matching," *Journal of Agricultural Machinery*, vol. 1, pp. 2034–2039, 2019.
[21] T. Zheng, *Research on Image Contrast Enhancement Algorithm Based on Human Visual Perception*, University of Electronic Science and Technology, Chengdu, China, 2016.
[22] X. Y. Zhang, W. S. He, H. Y. Duan, and S. T. Wei, "Improved SURF feature extraction and matching algorithm," *Mechanical Design and Manufacturing Engineering*, vol. 47, no. 11, pp. 58–62, 2018.
[23] J. L. Li and P. Q. Jiang, "Image stitching combining deformation function and power function weights," *Journal of Computer Applications*, vol. 39, no. 10, pp. 3060–3064, 2019.
[24] H. Y. Cai, X. Y. Wu, L. R. Zhuo et al., "Fast SIFT image mosaic method combined edge detection," *Infrared and Laser Engineering*, vol. 47, no. 11, pp. 449–455, 2018.
[25] D. Y. Tsai, Y. Lee, and E. Matsuyama, "Information entropy measure for evaluation of image quality," *Journal of Digital Imaging*, vol. 21, no. 3, pp. 338–347, 2018.

Hindawi
Complexity
Volume 2021, Article ID 5560555, 10 pages
https://doi.org/10.1155/2021/5560555

Research Article

An Efficient Universal Generating Function-Based Analyzing Approach for Multistate System with Imperfect Coverage Failure

Zheng Li [1,2] **and Jinlei Qin** [1,2]

[1]*Department of Computer, North China Electric Power University, Baoding, China*
[2]*Engineering Research Center of Intelligent Computing for Complex Energy Systems, Ministry of Education, Baoding, China*

Correspondence should be addressed to Jinlei Qin; jlqin717@163.com

Received 31 January 2021; Revised 18 May 2021; Accepted 4 June 2021; Published 14 June 2021

Academic Editor: Miaomiao Wang

A system with more than two states is called a multistate system (MSS), and such systems have already become a general trend in the arena of complex industrial products and/or systems. Fault-tolerant technology often plays a very important role in improving the reliability of an MSS. However, the existence of imperfect coverage failure (ICF) in a work-sharing group (WSG) decreases the reliability of MSS. A method is proposed to assess the reliability and sensitivity of an MSS with ICF. The components in a WSG can cooperate so as to improve overall efficiency by increasing performance levels. Using the technique of the universal generating function (UGF), a component's UGF expression with ICF can be incorporated in two steps. During the computation of the system's UGF, an algorithm based on matrix (ABM) is developed to reduce the computational complexity. Consequently, indices of reliability can be easily calculated based on the UGF expression of an MSS. Sensitivity analysis can help engineers judge which WSG should be eliminated first under various resource limitations. Examples illustrate and validate this method.

1. Introduction

The fault-tolerant system is a high-reliability system designed by incorporating redundant components for critical elements in order to prevent the overall system from failing even when some of its individual elements fail. Fault-tolerant systems are often used in life-critical applications such as flight control, nuclear plant monitoring, and space missions and in mission-critical computer monitoring systems and data storage systems [1]. In addition to redundancy, the implementation of fault tolerance also requires automatic recovery and reconfiguration mechanisms. That is to say, even if sufficient redundancy exists, if the system cannot adequately detect, locate, and recover from internal faults and/or errors that have occurred, the entire system or one of its subsystems can fail [2]. The degree of fault tolerance is determined by the proportion of faults from which a system automatically recovers, and these faults are said to be covered by the recovery strategy [3]. Therefore, the reliability analysis of such systems must take into account the process by which faults and errors are detected and recovered from, as well as the complex system structure.

The problem of determining a system's fault tolerance and its trend have been intensively explored, with special methods and analysis techniques to evaluate the system reliability having been put forward [4]. A new, simple, and efficient approach is presented for incorporating imperfect fault coverage into a combinational model [5]. The optimal design of some systems subjected to the imperfect fault coverage has been formulated for *G*: (k/n) structures [6]. Furthermore, the optimal reliability of systems subjected to the same reason, imperfect fault coverage, has already been generalized to more systems including parallel-series, series-parallel, parallel, etc., and the non *s*-identical component case [7]. According to the type of fault-tolerant techniques, the appropriate modeling method of multifault coverage is suggested for evaluating reliability indices in fault-tolerant system, and it is easy to be used in the case of hierarchy structure of fault-level coverage groups [8]. An approach based on a binary decision diagram is proposed to analyze

the reliability and sensitivity of a benchmark network [9]. Based on the total probability theorem and a divide-and-conquer strategy, a new combinational approach to handling functional dependency has been put forward in the reliability analysis of imperfect coverage systems [10]. A long-neglected issue is that an initially relevant component could become irrelevant after the failures of other components, and this issue has been modeled through the coverage of irrelevant components in the system with imperfect fault coverage [11].

The MSS reliability theory and related conceptions have been investigated from many aspects. In an MSS, the system and its elements can function at a range of different performance levels, e.g., from perfect operation to complete failure. A universal moment generating function is extended to incorporate common cause failure into MSS reliability estimation [12]. A new second-order reliability method without parabolic approximation of the fitted quadratic surface is presented to improve the accuracy of reliability analysis [13]. The reliability modeling of MSS with preventive maintenance and customer demand is proposed to improve the reliability of MSS [14]. Employing Kalman filtering approach to estimate system and degraded sensor state is the base to calculate reliability for making a dynamic maintenance decision [15]. A life cycle cost reliability model with copula is established for system with multiple dependent degradation processes and environmental influence [16]. Simulation is often used to evaluate MSS because of its computational complexity directly by analytical method [17]. Similarly, an efficient simulation method based on survival signature has been proposed for system reliability analysis [18]. A complex framework based on integrated direct partial logic derivative (DPLD), whose computational complexity correlates with the number of system components and does not dependent on the structural complexity of MSS, is developed for qualitative and quantitative analysis focusing on component criticality [19]. Calculating critical system state with DPLD, multivalued logic method is used for the reliability analysis of MSS [20]. Application of structural function is presented in time-dependent reliability analysis, and DPLDs can be also used to find formulae for computation of most commonly known time-dependent importance measures [21]. Besides those approaches mentioned above for reliability, it is also critical to determine the impact of each component on MSS performance. Due to the nonbinary state of MSS and their components as well as dependences among different states of the same component, the performability analysis of MSS becomes more difficult and multivalued decision diagram is presented as an efficient algorithm to analyze the MSS [22]. Degraded performance is a common phenomenon in industrial products, and reliability is a vital quality of the MSS for providing the required performance level. An integrated routing risk model is constructed and risk control performance is also proved by simulated algorithm [23]. Other risk assessment model and risky multicriteria decision-making steps are developed [24, 25]. Focusing on performance analysis and optimization, a wireless sensor network framework is constructed [26, 27]. Regarding industrial cyberphysical system, safety control and performance monitoring are also focused on to answer how urgent they are and what degrees of fault tolerance and fault recovery are needed [28–30]. Although many researchers devote efforts on different aspects of system reliability and performance, the issue of computational complexity for an MSS is often notorious one because of the curse of "dimension explosion."

An MSS can also be subjected to imperfect coverage failures that lead to the entire failure of the whole system or its subsystems. Based on the ordered binary decision diagram, an efficient algorithm for the reliability evaluation of an MSS with a combinational performance requirement subject to imperfect fault cover has been proposed [31]. Considering the importance of the system-component state of an MSS, Griffith's importance measures and reliability are evaluated by the combination of conditional probability methods to find solutions for the multistate imperfect fault coverage model [32]. An optimal structure of an MSS with uncovered failure can attain maximum reliability through a proper balance between two types of task parallelization: parallel task execution with work sharing and redundant task execution [33]. The MSS with three different types of imperfect fault coverage, element-level coverage, fault-level coverage, and performance-dependent coverage, has been modeled for evaluating the reliability of the system [34]. Hereafter, the similar two types of parallelization of MSS with multifault coverage have been studied to obtain the optimal trade-off based on various settings of fault coverage factor [35]. Stochastic multivalue models are proposed to predict the reliability of a multistate phase-mission system with three different imperfect fault coverage conditions, and the efficiency of this model is compared with the universal generating function (UGF) technique [36]. Within a subsystem of MSS, when the effectiveness of recover mechanism depends on the entire performance level of the subsystem, the MSS performance distribution can be obtained by a recursive procedure based on the UGF [37]. A modification of generalized reliability block diagram method is suggested for the reliability assessment of MSS with imperfect multifault coverage [38]. However, the computational time of these models and methods is proportional to the number of components and, hence, is of high computational complexity. UGF is a tool for efficiently dealing with the performance distribution of a complex MSS [39], and many reliability assessment studies of MSS have adopted it as the primary analysis tool [40–43].

The novelty of this paper lies in that a revised approach based on the UGF technique is proposed to assess the reliability of an MSS with imperfect coverage failures (ICFs). Here the ICF is modeled by incorporating a scheme of working-share group (WSG), which is similar to the parallel structure. However, they are two different types of data-transmitting schemes. During the assessment, an algorithm based on matrix (ABM) is developed to significantly reduce the computation complexity. This approach is compared with other methods on time consumption. Furthermore, the sensitivity analysis can be handled by the engineers to decide which work group with ICF should be eliminated first under the limited conditions.

The remainder of this paper is organized as follows. Section 2 describes the imperfect coverage failure model. In Section 3, the UGF technique is revised to incorporate the ICF in order to perform the reliability and sensitivity analysis for the MSS with ICF. In Section 4, this approach is illustrated by several applications. Section 5 includes the authors' conclusions and suggests future research dealing with the reliability of MSS.

2. Modeling of Imperfect Coverage and MSS

2.1. Imperfect Coverage Model. There are many models developed for the ICF, such as element-level coverage models and fault-level coverage models. Here, the focused structure of the imperfect coverage failure model with single point failure can be modeled as shown in Figure 1 for reliability analysis [10, 44]. The single point failure is also refereed as single fault at a component where its coverage probability is solely depended on the properties of the failed component. The entry point to the model signifies an occurrence of a single point failure with three types of possible outcomes. They are three types of exits—R, C, S—that signify, respectively, different possible outcomes with different exit probabilities r, c, and s. Exit R indicates that the offending failure is transient and can be handled without discarding the component. Exit C implies an occurrence of a permanent covered failure in the component that needs to be discarded for the normal operation of the system. Exit S is a single point failure signifying that the uncovered failure without discarding the component leads to the system failure. When a single point failure occurs, there are the three types of possible outcomes. So, they consist of a partitioning of the event space, and the three exit probabilities sum to one, namely, $r + c + s = 1$. For the multipoint failures, it is to descript system failure because of concurrent faults on multiple components. And it will be modeled and studied in the future.

In order to improve the reliability of a fault-tolerant system with imperfect coverage failures, a scheme of a work-sharing group (WSG) is introduced [33]. A WSG usually consists of several identical components that are connected in parallel and form a set. Every component will fulfill a different subtask so that its WSG can finish the entire task. For a multichannel data transmission system from A to B, as shown in Figure 2, the components $C1$ and $C2$ make up a WSG to share the work under the reconfiguration of a data exchange management system (DEMS). Here, the DEMS can be assumed that it does not fail because its form of existence can be implemented as built-in software or other forms. The WSG is parallel with component $C3$ so that they can undertake the data transmission task together. Because of the existence of DEMS, the same tasks that the data packages are transmitted by $C3$ as DP3 are divided into two sections, DP1 and DP2, according to the performance of components $C1$ and $C2$. Under this assumption, if one component in the WSG, such as $C1$, fails, the DEMS can discover it and will not assign data packages to it again. This type of failure can be defined as covered failure. The DEMS will then reassign the data transmission task to $C2$. The WSG

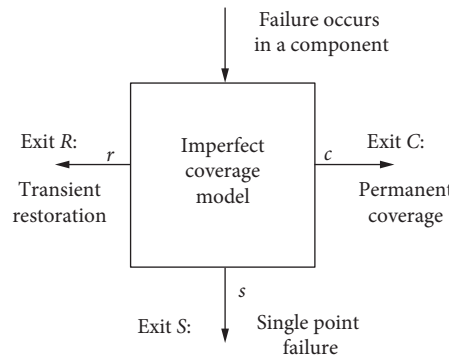


Figure 1: Imperfect coverage model with single point failure of a component.

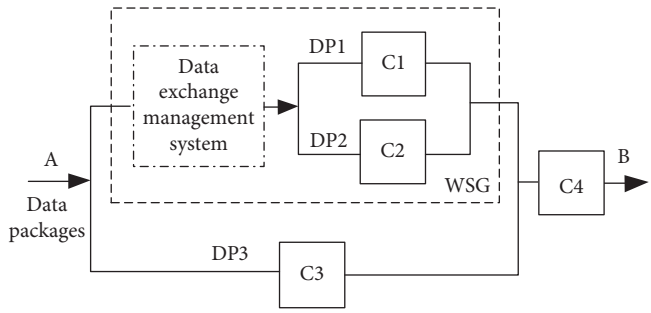


Figure 2: The working scheme of WSG.

will not completely fail, but its performance will decrease consequently. However, if $C1$ fails but the DEMS cannot discover it and continues to assign data transmission to $C1$, $C1$ will be in a state of exit s, namely, single point failure. The failure occurred in this situation can be defined as uncovered failure and the WSG cannot finish the task of data transmission.

The sole purpose of a parallel structure is to improve the reliability of the system. However, if there is no DEMS, the data packages will not be divided, and the efficiency of the system cannot be improved to transmit the data packages. Introducing the WSG into the system will increase efficiency, but due to the change in the system structure, reliability may be affected.

2.2. Modeling of MSS. For the purpose of modeling an MSS, the characteristics of its elements must first be defined. Generally speaking, any element i in MSS can have $m_i + 1$ different states corresponding to the performance levels that can be represented by the set $\mathbf{a}_i = \{a_{i0}, a_{i1}, \ldots, a_{im_i}\}$, where a_{is} is the performance level of element i in the state s, $s \in \{0, 1, \ldots, m_i\}$. The current performance level A_i of element i at any instant in time is a discrete random variable that takes value from $\mathbf{a}_i$: $A_i \in \mathbf{a}_i$. The probabilities of each different state or performance level for element i can be denoted by the set $\mathbf{p}_i = \{p_{i0}, p_{i1}, \ldots, p_{im_i}\}$, where

$$p_{is} = \Pr\{A_i = a_{is}\}. \quad (1)$$

Furthermore, the entire states of one element constitute a whole set whose states can be seen as mutually exclusive

events. That is to say, the element i will always be in one and only one of $m_i + 1$ states, such that

$$\sum_{s=0}^{m_i} p_{is} = 1. \tag{2}$$

The performance level distribution of element i will be determined completely by the collection of pairs

$$\mathrm{CP}_i = (a_{is}, p_{is}), \quad s = 0, 1, \ldots, m_i. \tag{3}$$

The system elements have certain performance levels corresponding to their respective states at one instant in time. The modeling of an MSS performance level is completely determined by CP_i, $i = 1, 2, \ldots, n$. Therefore, the states of an MSS are determined completely by the states of its components. Now, suppose that the MSS has $k + 1$ different states and the performance level corresponding to one state of the MSS at a certain moment can be represented by w_j, $j = 0, 1, \ldots, k$. In this case, the MSS performance level is a random variable denoted by W that takes values from the set

$$W = \{w_0, w_1, \ldots, w_k\}. \tag{4}$$

The probability mass function (PMF) of the MSS performance levels can be obtained as follows:

$$q_j = \Pr\{W = w_j\}, \quad j = 0, 1, \ldots, k. \tag{5}$$

Using the Cartesian product operation, we can define the space of all possible combinations of performance levels for all system components as

$$C^n = \prod_{i=1}^{n} \{a_{i0}, a_{i1}, \ldots, a_{im_i}\}. \tag{6}$$

The MSS structure function is naturally introduced as

$$\Phi(A_1, A_2, \ldots, A_n): C^n \longrightarrow W. \tag{7}$$

The function is to map the space of the component's performance levels into the space of the performance levels of an MSS.

From the above analysis, we can see that the model of an MSS includes two parts: PMF of performance levels for all system components and the structure function of the system. These can be rewritten as follows:

$$\begin{cases} \mathbf{a}_i, \mathbf{p}_i, & 0 \le i \le n, \\ W = \Phi(A_1, A_2, \ldots, A_n). \end{cases} \tag{8}$$

For an MSS whose performance level is defined as task completion time, its reliability can be expressed as the probability that the system satisfies the maximum allowed completion time of δ. From (5), one can obtain

$$R(\delta) = \sum_{j=0}^{M} q_j \Delta(w_j, \delta), \tag{9}$$

where

$$\Delta(w_j, \delta) = \begin{cases} 1, & \text{if } w_j \le \delta, \\ 0, & \text{if } w_j > \delta. \end{cases} \tag{10}$$

Another important measure is the conditional expected performance $\Omega(\delta)$. This expresses the system's expected performance under the condition that the MSS is in an acceptable state. Having the system reliability $R(\delta)$, this measure can be calculated by

$$\Omega(\delta) = \frac{\left(\sum_{j=0}^{M} q_j w_j \Delta(w_j, \delta)\right)}{R(\delta)}. \tag{11}$$

3. UGF Technique and Reliability Evaluation

3.1. UGF Technique. UGF, also called as u-function or universal z-transform [45], has been proven to be an effective method for solving high-dimensional combinatorial problems. UGF of a multistate component i associated with its performance level PMF can be defined as a polynomial

$$u_i(z) = \sum_{h_i=0}^{m_i} p_{ih_i} z^{a_{ih_i}}, \quad i = 1, 2, \ldots, n. \tag{12}$$

The essential property of UGF enables the entire UGF for an MSS, whose components are connected in series or parallel, to be obtained using simple algebraic operations corresponding to the individual UGF of a multistate component. To represent the PMF of the stochastic variable $W = \Phi(A_1, A_2, \ldots, A_n)$, the composition operator $\otimes_\phi$ is defined in the following equation:

$$\begin{aligned} U(z) &= \otimes_\phi (u_1(z), u_2(z), \ldots, u_n(z)) \\ &= \otimes_\phi \left(\sum_{h_1=0}^{m_1} p_{1h_1} z^{a_{1h_1}}, \sum_{h_2=0}^{m_2} p_{2h_2} z^{a_{2h_2}}, \ldots, \sum_{h_n=0}^{m_n} p_{nh_n} z^{a_{nh_n}} \right) \\ &= \sum_{h_1=0}^{m_1} \sum_{h_2=0}^{m_2} \cdots \sum_{h_n=0}^{m_n} \left(\prod_{i=1}^{n} p_{ih_i} z^{\phi\left(a_{1h_1}, a_{2h_2}, \ldots, a_{nh_n}\right)} \right). \end{aligned} \tag{13}$$

Note that the polynomial $U(z)$ represents all possible mutually exclusive combinations of individual independent components' UGFs. The function $\phi(a_{1h_1}, a_{2h_2}, \ldots, a_{nh_n})$ is determined according to the physical nature of the interaction between the performances of components.

Indeed, the derivation of $U(z)$ for various types of systems is usually a difficult computational task. As shown in [45], from the two perspectives of computational simplicity and derivation clarity, representing $U(z)$ in a recursive form is beneficial. In particular, when an MSS has a complex configuration, the entire system can be represented as the composition of subsystems corresponding to the subsets of multistate components. This property can be defined by

$$\begin{aligned} &\otimes_\phi (u_1(z), \ldots, u_k(z), u_{k+1}(z), \ldots, u_n(z)) \\ &= \otimes_\phi \left(\otimes_\phi (u_1(z), \ldots, u_k(z)), \otimes_\phi (u_{k+1}(z), \ldots, u_n(z)) \right). \end{aligned} \tag{14}$$

The configuration of any MSS can always be represented as a composition of independent subsystems containing only components connected in parallel or in series. For any components connected in parallel or in series in the MSS, the composition operator can be applied recursively in order to obtain the UGF of the intermediate pure parallel or pure series structures.

Consider one type of MSS system, a task processing computer system with its performance level defined as task completion time. For components connected in series, the system's total completion time is the sum of completion time of all its components. If two independent components (i and j) work in series, the total completion time is the sum of their individual completion times. The function ϕ_s should calculate the sum of corresponding parameters. The performance of the pair of components in this case is defined as

$$U_{ij}(z) = u_i(z) \otimes_{\phi_s} u_j(z) = \sum_{h_i=0}^{m_i} \sum_{h_j=0}^{m_j} p_{ih_i} p_{jh_j} z^{\mathrm{sum}\left(a_{ih_i}, a_{jh_j}\right)}. \tag{15}$$

For components connected in parallel, the total completion time is decided by the component with the shortest completion time. The function ϕ_p should obtain the minimum of all parameters. Therefore, the UGF in this case should take the following form:

$$U_{ij}(z) = u_i(z) \otimes_{\phi_p} u_j(z) = \sum_{h_i=0}^{m_i} \sum_{h_j=0}^{m_j} p_{ih_i} p_{jh_j} z^{\min\left(a_{ih_i}, a_{jh_j}\right)}. \tag{16}$$

For components connected in parallel with WSG, the task processing can be divided in proportion to their processing speeds. That is to say, components can share the task according to their performance levels. The function ϕ_{pw} should be the inverse of its performance, and the UGF for this pair of components can be determined by the function

$$U_{ij}(z) = u_i(z) \otimes_{\phi_{pw}} u_j(z) = \sum_{h_i=0}^{m_i} \sum_{h_j=0}^{m_j} p_{ih_i} p_{jh_j} z^{\mathrm{inv}\left(a_{ih_i}, a_{jh_j}\right)}, \tag{17}$$

where $\mathrm{inv}(a_{ih_i}, a_{jh_j}) = (a_{ih_i}^{-1} + a_{jh_j}^{-1})^{-1}$.

3.2. Reliability Analysis of MSS with ICF. For the MSS with ICF, its single point failure can usually be modeled as state 0 with assigned performance level [31]. The assigned performance level can coincide with component performance in the permanent coverage. The specific performance level is related to the system property, and here it might as well be assigned to x. According to (8), it can be expanded as

$$u_i(z) = p_{i0} z^x + \sum_{h_i=1}^{m_i} p_{ih_i} z^{a_{ih_i}} = p_{i0} z^x + \widehat{u}_i(z), \tag{18}$$

where the second item $\widehat{u}_i(z)$ represents the UGF of component i except for its single point failure state. If the component is not of the single point failure, $u_i(z) = \widehat{u}_i(z)$.

For the WSG with the single point failure, its UGF U_w can be obtained in three steps:

(1) $\widehat{u}_i(z)$ $(i \in \alpha)$ of every component can be expressed according to (18), where α is the set of components in the WSG.

(2) $\widehat{U}_w$ can be obtained based on the operation of components' $\widehat{u}_i(z)$ by (17).

(3) UGF U_w will be calculated as follows:

$$U_w(z) = \widehat{U}_w(z) + \left(1 - \prod_{i \in \alpha} (1 - a_{i0})\right) * z^x. \tag{19}$$

Once the UGF of WSG with ICF is obtained, it can be seen as a common component to be operated with other components according to the structure of the system. The UGF of the entire system can now be expressed easily. By applying (9) and (11), the reliability indices can be calculated.

During the calculation, the computation complexity of determining the UGF of two components is always high if it is only calculated manually. This is why the algorithm based on matrix (ABM) is developed. The seven steps of ABM, along with substeps, are as follows:

Step 1: the UGF of two components can be expressed in matrices A and B, respectively. Both A and B are composed by two rows and n columns. The first row is the performance level, and the second row is its corresponding probability. The number of columns in A and B may be different.

Step 2: define a matrix C with two rows. The number of C's columns is equal to the product of A's columns multiplied by B's columns. Matrix C is used to store the primitive values of the multiplication of matrixes A and B according to specific calculation rules.

Step 3: by applying the dual-iteration method, the elements in matrix C can be obtained from the following rules:

(i) Define the first column of A and B as the outer and inner iteration variables, respectively. Let k represent the column order of matrix C and $k = 1$ at beginning.

(ii) The value of matrix C's k column and first row should be calculated by the corresponding element of matrices A and B according to the function sum, min, or inv, based on (15)–(17).

(iii) The probability value of matrix C's k column and second row can be obtained by multiplying the corresponding element values of matrices A and B. The value in $C(2, k)$ is equal to the product of $A(2, 1)$ and $B(2, 1)$.

(iv) Assign k with $k + 1$, viz., $k = k + 1$.

(v) Modify the inner iteration variable to the second column of B and repeat the above three steps until the end of B's columns.
(vi) Modify the outer iteration variable to the second column of A, and repeat the above four steps until the end of A's columns.

Step 4: sort the values of C's first rows in ascending order, i.e., performance level. The elements of the second row of matrix C can be reordered correspondingly.

Step 5: abstract the unique elements from the first row of matrix C and form row vector E.

Step 6: define a result matrix D whose first row is equal to E. The second row of D is constructed by the following iteration steps:

(i) Define the first column of matrix D as the iteration variable.
(ii) From the first column of matrix C's first row, namely, $C(1, 1)$, locate for the same value with the first column of D's first row, namely, $D(1, 1)$.
(iii) Add the corresponding probabilities at the location determined by the previous step to the first column of D's second row, namely, $D(2, 1)$.
(iv) Modify the iteration variable to the second column, and then repeat the above two steps until the end of the number of D's columns.

Step 7: the results of the operation of matrices A and B are stored in the matrix D with just the performance level and probability. According to this matrix D, it can be rewritten in the UGF expression.

In the following examples, the ABM can be used frequently.

4. Illustrative Examples

4.1. Reliability Evaluation. Consider a task processing system with the structure as shown in Figure 2, in which components $C1$ and $C2$ comprise the WSG. After $C3$ is connected with the WSG in parallel, $C4$ is connected in series. These components are statistically independent. The data package can be transmitted through $C3$ and $C4$ in 15 seconds and 20 seconds with the probability of 0.7 and 0.75, respectively. The performance levels of $C1$ and $C2$ are the same. Their completion time and probability are both 15 seconds and 0.7. Their detected and undetected failures have equal probabilities, namely, 0.15. These parameters are listed in Table 1.

In this case, WSG $\alpha = \{C1, C2\}$ is affected by the ICF. According to the steps above, the UGF of these components' performance distribution can be represented as follows.

Based on the above analysis, the UGF of every component can be represented as

$$\begin{cases} u_{c1}(z) = u_{c2}(z) = 0.7 * z^{15} + 0.15 * z^{\infty} + 0.15 * z^{x}, \\ u_{c3}(z) = 0.7 * z^{15} + 0.3 * z^{\infty}, \\ u_{c4}(z) = 0.75 * z^{20} + 0.2 * z^{25} + 0.05 * z^{\infty}, \end{cases} \tag{20}$$

where $\widehat{u}_{c1}(z) = \widehat{u}_{c2}(z) = 0.7 * z^{15} + 0.15 * z^{\infty}$. Because the WSG consists of $C1$ and $C2$, they finish the data transmission task by sharing. Applying (17), such that

$$\begin{aligned} \widehat{U}_w(z) &= \widehat{u}_{c1}(z) \otimes_{\phi_{pw}} \widehat{u}_{c2}(z) \\ &= \left(0.7 * z^{15} + 0.15 * z^{\infty}\right) \otimes_{\phi_{pw}} \left(0.7 * z^{15} + 0.15 * z^{\infty}\right) \\ &= 0.49 * z^{7.5} + 0.21 * z^{15} + 0.0225 * z^{\infty}. \end{aligned} \tag{21}$$

A component's single point failure also leads to its lowest performance level, namely, infinity of task completion time $x = \infty$. According to (19),

$$\begin{aligned} U_w(z) &= \widehat{U}_w(z) + \left(1 - \prod_{i \in \{C1, C2\}} (1 - a_{i0})\right) * z^{x} \\ &= \widehat{U}_w(z) + \left(1 - (1 - 0.15)^2\right) * z^{\infty} \\ &= \widehat{U}_w(z) + 0.2775 * z^{\infty} \\ &= 0.49 * z^{7.5} + 0.21 * z^{15} + 0.3 * z^{\infty}. \end{aligned} \tag{22}$$

Based on the reliability block diagram, because the WSG is connected with $C3$ in parallel, (16) is adopted first. Component $C4$ is connected by series, and then (15) needs to be applied. The UGF of the system can be expressed as follows:

$$\begin{aligned} U_{\text{sys}}(z) &= \left(U_w(z) \otimes_{\phi_p} u_{c3}(z)\right) \otimes_{\phi_s} u_{c4}(z) \\ &= \left(0.49 * z^{7.5} + 0.42 * z^{15} + 0.09 * z^{\infty}\right) \otimes_{\phi_s} u_{c4}(z) \\ &= 0.3675 * z^{27.5} + 0.098 * z^{32.5} + 0.315 * z^{35} \\ &\quad + 0.084 * z^{40} + 0.1355 * z^{\infty}. \end{aligned} \tag{23}$$

According to (9) and (11), the probability that system can transmit data packages within $\delta = 40$ s and the conditional expected performance are

$$\begin{cases} R(40) = 1 - 0.1355 = 0.8645, \\ \Omega(40) = \dfrac{(0.3675 * 27.5 + 0.098 * 32.5 + 0.315 * 35 + 0.084 * 40)}{0.8645} \\ = 32.0142. \end{cases} \tag{24}$$

Table 1: Parameters of components.

Component	States	Probability	Performance level (time: s)	With or without ICF (Y/N)
$C1/C2$	0	0.15	∞	Y
	0	0.15	∞	N
	1	0.7	15	—
$C3$	0	0.3	∞	N
	1	0.7	15	—
$C4$	0	0.05	∞	N
	1	0.2	25	—
	2	0.75	20	—

If there is no WSG in this system, that is to say, if components $C1$, $C2$, and $C3$ are connected in parallel, then the UGF of $C1$ and $C2$ is expressed as

$$u_{c1}'(z) = u_{c2}'(z) = 0.7 * z^{15} + 0.3 * z^{\infty}. \tag{25}$$

The UGF of the system can be calculated by

$$\begin{aligned} U_{\text{sys}}'(z) &= \left(u_{c1}'(z) \otimes_{\phi_p} u_{c2}'(z) \otimes_{\phi_p} u_{c3}(z)\right) \otimes_{\phi_s} u_{c4}(z) \\ &= 0.72975 * z^{35} + 0.1946 * z^{40} + 0.07565 * z^{\infty}. \end{aligned} \tag{26}$$

Similarly, the reliability and conditional expected performance can be calculated as

$$\begin{cases} R'(40) = 0.92435, \\ \Omega'(40) = \dfrac{(0.72975 * 35 + 0.1946 * 40)}{0.92435} \\ = 36.0526. \end{cases} \tag{27}$$

From the comparison of (24) and (27), it can be seen that the reliability of the system with ICF is decreased. However, its expected task completion time is less than that of the system without WSG. In other words, the system with WSG has higher efficiency to complete the transmission of data packages. However, because of the existence of ICF, its reliability has declined.

In order to compare the efficiency of the above analyzing method with multivalued decision diagram (MDD) derived from reference [22], the system in Figure 2 and some parameters in Table 1 are still used for the reliability calculation. The other assumptions and parameters needed in MDD are in concert with those in the mentioned reference. Here, the focus is to compare the time consumption during the calculation for reliability. In the same configuration of computing environments, those results calculated by MDD are approaching to ones in (24). There is a distance on the time consuming between those two methods. From the validation listed in Table 2, it shows the ABM is less time consuming than MDD.

4.2. Sensitivity Analysis. Figure 3 shows a more complex system, composed of three subsystems. Within each subsystem, a WSG is configured to improve task processing efficiency.

Table 2: Comparison of results for two different methods.

Method	R (40)	Ω (40)	Time consumption (s)
ABM	0.8645	32.0142	1.0385
MDD	0.8539	32.1325	5.7862

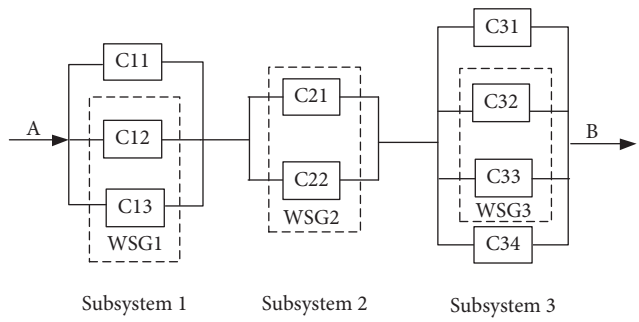


Figure 3: Sensitivity analysis diagram.

Because the components are of several states in each subsystem, the performance levels of task processing time and other parameters of these components are listed in the form of a five-element tuple, as shown in Table 3.

Based on the above method and these parameters, the system reliability satisfying the maximum allowed completion time can be calculated by (9). Consider four different cases:

A: the WSG in each subsystem is operating and works well

B: the WSG1 is eliminated and components $C12$ and $C13$ work in parallel

C: the WSG2 is eliminated and components $C21$ and $C22$ work in parallel

D: WSG3 is eliminated and components $C32$ and $C33$ work in parallel

The system reliability curves for these four cases are shown in Figure 4. Figure 4(a) shows the entire curve of the four cases. From Figure 4(b), which is a partial curve scaled in [25], [55], it can be seen that case A has the lowest reliability of the four cases. The sequence of reliability improvement for the other three cases, from low to high, is D, B, and C.

For the purpose of quantitative analysis of reliability improvements, those results are listed in Table 4. It can be

Table 3: Five-element tuple of components.

Component	(States, probability, performance level (time: s), with or without ICF (Y/N))				
	First	second	Third	Fourth	Fifth
C11	(0, 0.15, ∞, N)	(1, 0.8, 10, −)	(2, 0.05, 15, −)	—	—
C12	(0, 0.05, ∞, Y)	(0, 0.05, ∞, N)	(1, 0.7, 8, −)	(2, 0.15, 10, −)	(3, 0.05, 13, −)
C13	(0, 0.06, ∞, Y)	(0, 0.05, ∞, N)	(1, 0.7, 15, −)	(2, 0.1, 18, −)	(3, 0.09, 20, −)
C21	(0, 0.05, ∞, Y)	(0, 0.1, ∞, N)	(1, 0.8, 5, −)	(2, 0.05, 10, −)	—
C22	(0, 0.02, ∞, Y)	(0, 0.03, ∞, N)	(1, 0.7, 5, −)	(2, 0.2, 12, −)	(3, 0.05, 17, −)
C31	(0, 0.1, ∞, N)	(1, 0.75, 10, −)	(2, 0.1, 13, −)	(3, 0.05, 16, −)	—
C32	(0, 0.08, ∞, Y)	(0, 0.06, ∞, N)	(1, 0.65, 8, −)	(2, 0.1, 15, −)	(3, 0.11, 18, −)
C33	(0, 0.10, ∞, Y)	(0, 0.06, ∞, N)	(1, 0.7, 9, −)	(2, 0.1, 12, −)	(3, 0.04, 15, −)
C34	(0, 0.2, ∞, N)	(1, 0.7, 10, −)	(2, 0.1, 13, −)	—	—

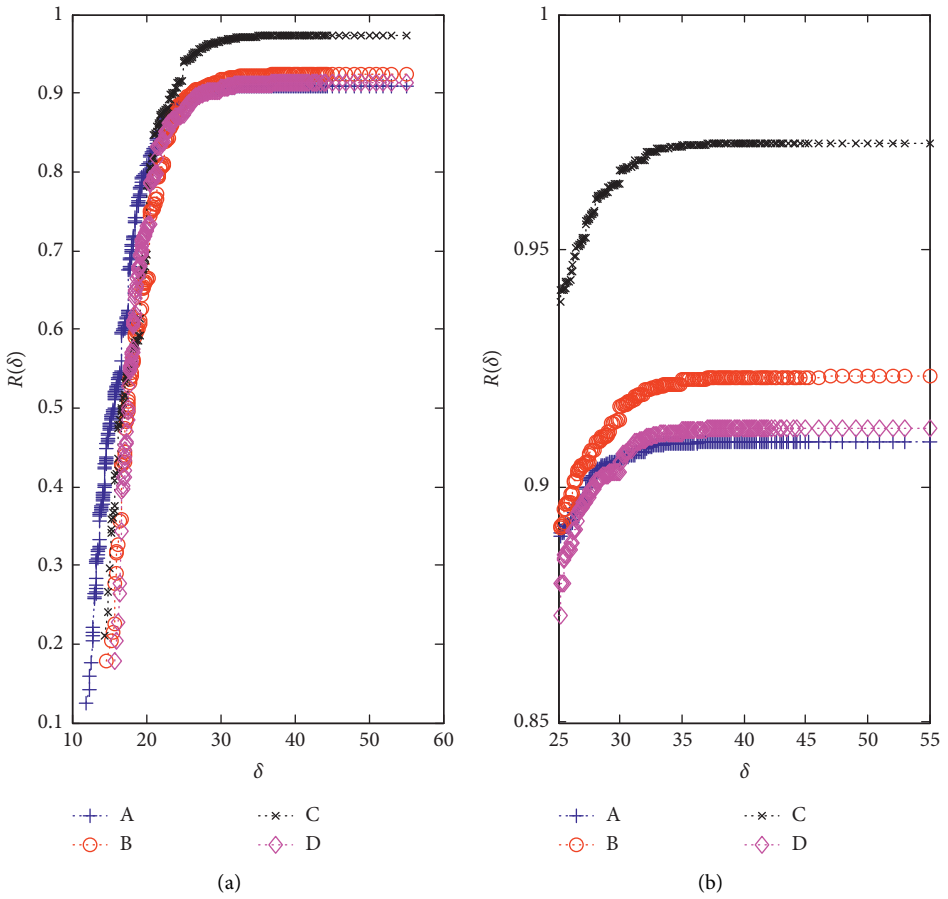


Figure 4: Reliability curves for four cases.

Table 4: Quantitative analysis.

Case	R (55)	Improvement (%)	Ω (55)	Reduction (%)
A	0.909552	—	15.85966	—
B	0.923215	1.50%	18.21348	14.84%
C	0.97277	6.95%	17.87951	12.74%
D	0.912349	0.31%	18.2138	14.84%

seen that Case C improves the reliability the most. At the same time, its expected performance is the lowest of the four cases. Although Cases B and D have the almost same reduction in expected performance, Case B shows higher improvement of reliability than Case D. It can be also seen that the results of reliability improvement from the quantitative analysis are in agreement with those from the above figures.

By applying the above analysis, the reliability engineer can choose the optimal scheme according to the indices of reliability or expected performance. Moreover, in this example, the sensitivity analysis is aimed only at the single WSG. In fact, combination of several WSGs is also a suitable mechanism for further sensitivity analysis based on this approach.

5. Concluding Remarks

In this paper, a reliability and sensitivity analysis approach is proposed for an MSS with ICF. In order to improve transmission efficiency, the components within a WSG can share a task such as data transmission according to its processing time. During the task sharing, the system may fail to detect the failure of one component, in which case, system reliability will be reduced and could even lead to the failure of task transmission. Based on the technique of UGF, the impact of ICF is incorporated in the expression of its UGF and its calculation methods for reliability and sensitivity analysis are also suggested. During the computation, the ABM algorithm is developed in order to significantly reduce the computational complexity. Two examples in this paper illustrate the application of the suggested approach. A reliability engineer can easily apply this approach to decide the optimal scheme under limited resources. In the above analysis, only the series-parallel MSS is considered. More complex MSS topologies, such as bridge and G: (k/n) structures, should be the focus. For other ICF models with different structures, they will be also explored in the future research.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was supported financially in part by a grant from Fundamental Research Funds for the Central Universities (nos. 2020MS120 and 2018MS076).

References

[1] W. G. Bouricius, W. C. Carter, D. C. Jessep, P. R. Schneider, and A. B. Wadia, "Reliability modeling for fault-tolerant computers," *IEEE Transactions on Computers*, vol. C-20, no. 11, pp. 1306–1311, 1971.

[2] J. B. Dugan, "Fault trees and imperfect coverage," *IEEE Transactions on Reliability*, vol. 38, no. 2, pp. 177–185, 1989.

[3] T. F. Arnold, "The concept of coverage and its effect on the reliability model of a repairable system," *IEEE Transactions on Computers*, vol. C-22, no. 3, pp. 251–254, 1973.

[4] S. Amari, A. Myers, A. Rauzy, and K. Trivedi, "Imperfect coverage models: status and trends," in *Handbook of Performability Engineering*, K. B. Misra, Ed., pp. 321–348, Springer-Verlag, Berlin, Germany, 2008.

[5] S. V. Amari, J. B. Dugan, and R. B. Misra, "A separable method for incorporating imperfect fault-coverage into combinatorial models," *IEEE Transactions on Reliability*, vol. 48, no. 3, pp. 267–274, 1999.

[6] S. V. Amari, H. Pham, and G. Dill, "Optimal design of$k$-out-of-n:G subsystems subjected to imperfect fault-coverage," *IEEE Transactions on Reliability*, vol. 53, no. 4, pp. 567–575, 2004.

[7] S. V. Amari, J. B. Dugan, and R. B. Misra, "Optimal reliability of systems subject to imperfect fault-coverage," *IEEE Transactions on Reliability*, vol. 48, no. 3, pp. 275–284, 1999.

[8] G. L. S. V. Amari, "Reliability analysis of fault tolerant systems with multi-fault coverage," *International Journal of Performability Engineering*, vol. 3, no. 4, pp. 441–451, 2007.

[9] L. Liudong Xing, "An efficient binary-decision-diagram-based approach for network reliability and sensitivity analysis," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 38, no. 1, pp. 105–115, 2008.

[10] L. Xing, B. A. Morrissette, and J. B. Dugan, "Efficient analysis of imperfect coverage systems with functional dependence," in *Reliability and Maintainability Symposium (RAMS), 2010 Proceedings-Annual*, San Jose, CA, USA, January 2010.

[11] J. Xiang, F. Machida, K. Tadano, and Y. Maeno, "An imperfect fault coverage model with coverage of irrelevant components," *IEEE Transactions on Reliability*, vol. 64, no. 1, pp. 320–332, 2015.

[12] G. Levitin, "Incorporating common-cause failures into nonrepairable multistate series-parallel system analysis," *IEEE Transactions on Reliability*, vol. 50, no. 4, pp. 380–388, 2001.

[13] X. Huang, Y. Li, Y. Zhang, and X. Zhang, "A new direct second-order reliability analysis method," *Applied Mathematical Modelling*, vol. 55, no. 2018, pp. 68–80, 2018.

[14] J. Qin and Z. Li, "Reliability modeling for multistate system with preventive maintenance under customer demand," *Complexity*, vol. 2020, Article ID 3165230, , 2020.

[15] B. Liu, P. Do, B. Iung, and M. Xie, "Stochastic filtering approach for condition-based maintenance considering sensor degradation," *IEEE Transactions on Automation Science and Engineering*, vol. 17, no. 1, pp. 177–190, 2020.

[16] B. Liu, X. J. Zhao, G. Q. Liu, and Y. Q. Liu, "Life cycle cost analysis considering multiple dependent degradation processes and environmental influence," *Reliability Engineering & System Safety*, vol. 197, no. 10, 2020.

[17] J. L. Qin, Z. Li, Y. G. Niu, and G. Q. Meng, "Simulated method for reliability evaluation of multi-state coherent system," *Iranian Journal of Science and Technology Transaction A-Science*, vol. 42, no. A3, pp. 1363–1371, 2018.

[18] E. Patelli, G. Feng, F. P. A. Coolen, and T. Coolen-Maturi, "Simulation methods for system reliability using the survival signature," *Reliability Engineering & System Safety*, vol. 167, pp. 327–337, 2017.

[19] M. Kvassay, E. Zaitseva, and V. Levashenko, "Importance analysis of multi-state systems based on tools of logical differential calculus," *Reliability Engineering & System Safety*, vol. 165, pp. 302–316, 2017.

[20] M. Kvassay, J. Rabcan, and P. Rusnak, "Multiple-valued logic in analysis of critical states of multi-state system," in *Proceedings of the 2017 International Conference on Information and Digital Technologies (IDT)*, Zilina, Slovakia, July 2017.

[21] P. Rusnak, J. Rabcan, M. Kvassay, and V. Levashenko, "Time-dependent reliability analysis based on structure function and logic differential calculus," in *Contemporary Complex Systems*

*and Their Dependability*Springer International Publishing, Cham, Switzerland, 2019.

[22] S. V. Amari, L. Liudong Xing, A. Shrestha, J. Akers, and K. S. Trivedi, "Performability analysis of multistate computing systems using multivalued decision diagrams," *IEEE Transactions on Computers*, vol. 59, no. 10, pp. 1419–1433, 2010.

[23] B. Fan, Y. Zeng, K. Jiang, and L. Tang, "A controllable chaotic immune algorithm for risk-aware routing in DiffServ networks," *Computer Communications*, vol. 76, pp. 67–76, 2016.

[24] C. B. Li, J. H. Yuan, and X. Feng, "Linguistic assessment information risky multi-criteria decision-making about wind power investment," *Journal of Intelligent & Fuzzy Systems*, vol. 30, no. 5, pp. 3017–3023, 2016.

[25] C. Li, J. Yuan, D. Liu, W. Li, and S. Lin, "Risk group assessment based on the demands of important consumers with standby electric source," *Journal of Intelligent & Fuzzy Systems*, vol. 33, no. 4, pp. 2177–2188, 2017.

[26] Y. Ju, J. Y. Yan, and H. Xu, "Performance optimization based on compressive sensing for wireless sensor networks," *Wireless Personal Communications*, vol. 95, no. 3, pp. 1927–1941, 2017.

[27] Z. X. Chen, Y. F. Jing, D. S. Han, and L. J. Wang, "Performance analysis of dual-media cooperative communication based on wireless and power line under hybrid fading," *International Journal of Distributed Sensor Networks*, vol. 15, no. 5, p. 11, 2019.

[28] Y. Jiang, S. Yin, and O. Kaynak, "Data-driven monitoring and safety control of industrial cyber-physical systems: Basics and beyond," *IEEE Access*, vol. 6, pp. 47374–47384, 2018.

[29] S. Yin, J. J. Rodriguez-Andina, and Y. Jiang, "Real-time monitoring and control of industrial cyberphysical systems: with integrated plant-wide monitoring and control framework," *IEEE Industrial Electronics Magazine*, vol. 13, no. 4, pp. 38–47, 2019.

[30] Y. Jiang, S. Yin, and O. Kaynak, "Performance supervised plant-wide process monitoring in industry 4.0: a roadmap," *IEEE Open Journal of the Industrial Electronics Society*, vol. 2, pp. 21–35, 2021.

[31] Y.-R. Chang, S. V. Amari, and S.-Y. Kuo, "Reliability evaluation of multi-state systems subject to imperfect coverage using OBDD," in *Proceedings of the 2002 Pacific Rim International Symposium on Dependable Computing*, Tsukuba, Japan, Decembet 2002.

[32] Y.-R. Yung-Ruei Chang, S. V. Amari, and S.-Y. Sy-Yen Kuo, "OBDD-based evaluation of reliability and importance measures for multistate systems subject to imperfect fault coverage," *IEEE Transactions on Dependable and Secure Computing*, vol. 2, no. 4, pp. 336–347, 2005.

[33] G. Levitin, "Optimal structure of multi-state systems with uncovered failures," *IEEE Transactions on Reliability*, vol. 57, no. 1, pp. 140–148, 2008.

[34] G. Levitin and S. Amari, "Three types of fault coverage in multi-state systems," in *Proceedings of the 2009 8th International Conference on Reliability, Maintainability and Safety*, Chengdu, China, July 2009.

[35] R. Peng, H. Mo, M. Xie, and G. Levitin, "Optimal structure of multi-state systems with multi-fault coverage," *Reliability Engineering & System Safety*, vol. 119, no. 2013, pp. 18–25, 2013.

[36] X. Song, Z. Zhai, P. Zhu, Y. Guo, and Y. Zhang, "A stochastic approach for evaluating the reliability of multi-stated phased-mission systems with imperfect fault coverage," in *Proceedings of the 2017 Prognostics and System Health Management Conference (PHM-Harbin)*, Harbin, China, July 2017.

[37] G. Levitin and S. V. Amari, "Multistate systems with static performance- dependent fault coverage," *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, vol. 222, no. 2, pp. 95–103, 2008.

[38] G. Levitin and S. V. Amari, "Multi-state systems with multi-fault coverage," *Reliability Engineering & System Safety*, vol. 93, no. 11, pp. 1730–1739, 2008.

[39] I. A. Ushakov, "Optimal standby problems and a universal generating function," *Soviet Journal of Computer and Systems Sciences*, vol. 25, no. 4, pp. 79–82, 1987.

[40] G. Levitin, "Reliability of multi-state systems with common bus performance sharing," *IIE Transactions*, vol. 43, no. 7, pp. 518–524, 2011.

[41] H. Yu, J. Yang, and H. Mo, "Reliability analysis of repairable multi-state system with common bus performance sharing," *Reliability Engineering & System Safety*, vol. 132, pp. 90–96, 2014.

[42] J. Qin, Y. Niu, and Z. Li, "A combined method for reliability analysis OF multi-state system OF minor-repairable components," *Ekspolatacja I Niezawodnosc-Maintenance and Reliability*, vol. 18, no. 1, pp. 80–88, 2016.

[43] J. L. Qin and Z. Li, "Reliability and sensitivity analysis method for a multistate system with common cause failure," *Complexity*, vol. 2019, no. 8, 2019.

[44] L. Xing, "Reliability evaluation of phased-mission systems with imperfect fault coverage and common-cause failures," *IEEE Transactions on Reliability*, vol. 56, no. 1, pp. 58–68, 2007.

[45] G. Levitin, *The Universal Generating Function in Reliability Analysis and Optimization*, Springer-Verlag, London, UK, 2005.

Hindawi
Complexity
Volume 2021, Article ID 9976473, 18 pages
https://doi.org/10.1155/2021/9976473

WILEY | Hindawi

Research Article

Optimal Control Strategy of Platform Load Oriented to Network and Load Cooperation

Xuesong Shao, Gaoying Cui, Xiao Chen, Xinrong Ji, and Yongxian Yi

State Grid Jiangsu Marketing Service Center(Metrology Center), Nanjing, Jiangsu Province 210016, China

Correspondence should be addressed to Xuesong Shao; zhao_yuanliang@163.com

Received 1 April 2021; Revised 9 May 2021; Accepted 28 May 2021; Published 8 June 2021

Academic Editor: Xiao Ling Wang

In recent years, with the continuous growth of China's power peak load and the rapid development of renewable energy, a large number of renewable energy sources are connected to the power grid, increasing the uncertainty of power grid operation and posing new major challenges to the power system regulation capacity. Flexible load has the characteristics of wide distribution, fast response, and high economy, which is an important control resource for the future power system. Based on the flexible load of commercial buildings and residential users, this paper studies the resource characteristics and response characteristics, clarifies the resource characteristics and demand response characteristic indexes of commercial and residential users, and establishes the response characteristics model of commercial buildings and residential users. Considering the influence of weather, holidays, incentive mechanism, and other factors on the response of flexible load, the quantitative analysis method of flexible load resource regulation potential for regional power grid dispatching was studied, and the feasibility of flexible load resources directly participating in the load control system was analyzed. Based on the uncertainty and mathematical characterization method of the active response of flexible loads, the optimal combination control strategy of demand response resources was proposed to eliminate the problems of heavy load and overload of regional power grid equipment by using the active response ability of flexible loads. Finally, the IEEE 14-node system is selected for simulation verification, which provides a theoretical basis for alleviating the power grid operation pressure in the peak load period of the power grid in the urban core area, improving the safety and economic operation level of regional power grid dispatching and the utilization rate of power grid equipment assets.

1. Introduction

With the rapid development of China's economy and the increasing load, the contradiction between power supply and demand is further deepened. Flexible load has the advantages of fast response and high economy and reflects different users' willingness to use electricity. Flexible load resources are important interactive resources in the framework of smart grid. Although a single load is not controllable at any time, a large number of small- and medium-sized loads will show a certain overall controllability when gathered together.

In order to make the flexible loads of commercial and residential users better participate in the system scheduling through the demand response project, it is necessary to master the response characteristics when they participate in the demand response project. Domestic and foreign scholars have done a lot of research on this. The literature [1] studies the load regulation technology of a large number of large apartment air conditioning load clusters, studies the regulation characteristics of the air conditioning load cluster, and proposes a regulation method of state cohort based on this characteristic. The literature [2] proposes a new temperature setting adjustment method based on the load regulation characteristics of air conditioning regulated by the state queue method proposed by Lee et al., in order to improve problems such as state absence and load fluctuation in regulation. The literature [3] establishes the first-order

model of a single air conditioner and establishes the power aggregation model of air conditioning load by using the Latin hypervertical method. After the establishment of the aggregation model, the controllable capacity of air conditioning load is evaluated, and the expression between the controllable capacity and the control temperature and outdoor temperature was obtained. The literature [4] compares three different ways to contract design in which demand response participates in power system regulation for incentive payments. The literature [5] mainly analyzes the relationship between the installation capacity of residents' air conditioning load and residents' characteristic variables. In the literature [6], the status quo of energy consumption of public buildings at home and abroad and the development process of load calculation methods are systematically reviewed and analyzed, the methods of investigation and field measurement of energy consumption of public buildings and the possible problems are discussed, and some suggestions are given. The literature [7] divides the air conditioning load into three industries, which are commercial users, administrative and public institution users, and residential users, and obtains the industry load through the ratio relationship between the industry electric quantity and the number of hours of the maximum load utilization. The literature [8] proposes a demand response scheme using a hopping scheme for consumption scheduling of appliances which can protect the privacy of electricity consumers. The literature [9] proposes a novel demand-side response model, which can be integrated into system scheduling model straightway and enables the optimal scheduling of energy/power consumption of a heterogeneous population of TCLs. The literature [10] analyzes and studies the load control technology of central air conditioning based on demand response. The literature [11] analyzes the response characteristic that the diversity of the load group was destroyed after the air conditioning load group participated in the regulation, proposes the control strategy to maintain the diversity of the air conditioning load group, analyzes the phenomenon of "secondary peak load" after the air conditioning load group was controlled, and puts forward the time-divided load recovery strategy. The literature [12] studies the response characteristics of refrigerated and nonrefrigerated central air conditioning systems used in public buildings. The literature [13] proposes a simple dynamic model on the basis of the analysis, estimation, and control of a variable-output compressor air conditioner (VOCAC).

The literature [14] shows that with the development of smart power grid, the interaction ability of "source-network-load" is improved, and the participation of flexible loads in power grid scheduling makes the management of line overload more flexible and economical. Scholars at home and abroad have studied how to mobilize flexible loads in courts to eliminate heavy load on lines and formed a heavy-load elimination mechanism based on electricity price and incentive demand responses. The literature [15] sorts out and analyzes the lessons of the restructuring of the American energy industry based on the natural gas and electric power accidents in California. In the literature [16], a model predictive control (MPC) of distributed air conditioning loads is proposed in order to balance fluctuations in solar power. The literature [17–19] considers the impact of load demand elasticity on power grid power flow from the perspective of users and aims at maximizing users' benefits by establishing reasonable electricity price to guide load transfer from peak period to trough period to eliminate line overload. The literature [20–23] invents interruptible loads to alleviate the problem of overloaded lines. In order to minimize the interruption load, interruption compensation price, and power purchase cost of users, the power flow of heavy-duty lines can be reduced by making reasonable load reduction plan. The literature [24] firstly analyzes different types of power loads, builds user load characteristic quantity that reflects the power consumption characteristics of loads, and then obtains user demand response load curve after the implementation of time-of use electricity price. Then, according to the principle of economics, the demand response behavior of users under the environment of TOU electricity price is analyzed, and the corresponding load reduction and transfer model is constructed. The literature [25] encourages users to respond by sharing the benefits brought by reducing grid congestion with users. However, these results have only been verified in a single-time cross-section and have yet to be verified in a continuous-time cross-section. At the same time, simple electricity price or incentive mechanism is difficult to fully mobilize the active response potential of flexible loads, while the coordinated application of electricity price and incentive mechanism can increase the diversity of choices on the user side, which is conducive to the maximum utilization of the demand response potential of flexible loads and effectively reduce the power flow of heavy-load line.

The agglomeration effect caused by a large number of flexible load access will affect the existing dispatching operation mode of network in different degrees. At present, there are some research studies on flexible load participating in dispatching operation control of power grid at home and abroad which mainly focusing on the active response characteristics of flexible loads, load characteristics of electric vehicles, demand response scheduling strategies participating in wind power consumption and the influence on rotating reserve capacity, etc. However, there is little research on the model and algorithm of regional power grid optimization reconfiguration with flexible load. Therefore, it is urgent to carry out research on the key technologies of flexible load participating grid dispatching operation to support the integration of demand response dispatching and current dispatching methods.

Based on the load response characteristics of commercial buildings and residential users, this paper analyzes the potential of flexible load regulation. On this basis, the flexible load optimal combination control model based on the maximum comprehensive benefit is established, and the medium- and long-term control, day-ahead control, and real-time control methods of flexible load are studied. And the feasibility of flexible load resources directly participating in the load control system was analyzed. Based on the uncertainty and mathematical characterization method of

the active response of flexible loads, the optimal combination control strategy of demand response resources was proposed to eliminate the problems of heavy load and overload of regional power grid equipment by using the active response ability of flexible loads. In this paper, the complementary potential of the hybrid energy system and the response priority mechanism of the integrated energy system equipment are analyzed from the aspects of technical and economic factors. The design model of the integrated energy coordinated control is optimized. Finally, the IEEE 14-node system is selected for simulation verification. It provides a theoretical basis for improving the dynamic balance ability of the power grid, alleviating the operating pressure of the power grid and improving the economic operating level of the power grid.

2. An Analysis of Load Response Characteristics of Commercial Buildings and Residential Users Oriented to Courts

2.1. An Analysis of Load Response Characteristics of Commercial Building. In the central air conditioning system, the operating power of the refrigerator takes the largest proportion of the operating power of the central air conditioning system. Therefore, only the refrigerators in the central air conditioning system are regulated and mathematically modeled.

Objective function:

$$\max\left(P_{\text{pot}}\right) = \max\left(P_{\text{air}} \cdot \text{num\_pot}\right)\left(t = T_{e\_\text{start}}, \ldots, T_{e\_\text{end}}\right). \tag{1}$$

In the above formula, T_{e_start} and T_{e_end} represent the start and end moments of emergency demand response events, respectively. During the duration of emergency demand response events, the load reduction of central air conditioning load group may be different at each moment, and the minimum load reduction over this period is defined as the regulation potential P_{pot}. P_{air} is the operating power of a single central air conditioning refrigerator; num_pot is a decision variable to be optimized, and its significance is the number of central air conditioning refrigerators actually reduced corresponding to the regulation potential. Therefore, the significance of the objective function is to maximize the regulation potential of the central air conditioning load group during the emergency demand response event.

The constraints are as follows:

(1) Load reduction constraints during emergency demand response events

$$P_{\text{base}}(t) - \sum_{k=1}^{n} P_{\text{air}} \cdot S(k,t) \geq P_{\text{pot}}\left(t = T_{e\_\text{start}}, \ldots, T_{e\_\text{end}}\right). \tag{2}$$

In equation (2), $P_{\text{base}}(t)$ represents the original overall operating power of the central air conditioning group at time t, that is, the load baseline of the central air conditioning load group; $\sum_{k=1}^{n} P_{\text{air}} \cdot S(k,t)$ represents the overall actual operating power after the central air conditioning load group is controlled at time t, and n represents the number of controlled central air conditioning in the central air conditioning load group; $S(k,t)$ represents the operating state of the refrigerator of the k-th central air conditioning at time t, $S(k,t) = 1$ means that the refrigerator of the k-th central air conditioning is in operation at time t, and $S(k,t) = 0$ means that the refrigerator of the k-th central air conditioning is in shutdown state at time t.

(2) Room temperature time-varying equation

According to the energy conservation principle, during the refrigeration period of the central air conditioning system, the indoor energy conservation equation can be obtained:

$$\rho_a \cdot V_k \cdot c_a \cdot \mathrm{d}T_{\text{in}}^{\text{on}}(t) = Q_{t1} \cdot \mathrm{d}t + Q_{\text{er}} \cdot \mathrm{d}t + Q_{\text{new}} \cdot \mathrm{d}t - Q_{t2} - Q_z^{\text{on}} \cdot \mathrm{d}t, \tag{3}$$

where ρ_a is air density; V_k is indoor volume; c_a is the air's specific heat capacity; $T_{\text{in}}^{\text{on}}(t)$ is the indoor temperature within the period of t in the refrigeration period; and $\rho_a \cdot V_k \cdot c_a \cdot \mathrm{d}T_{\text{in}}^{\text{on}}(t)$ ρ represents the heat obtained in the room during the dt period. By solving the first-order differential equation, the time-varying equation of indoor temperature in the refrigeration period can be obtained, shown as follows:

$$T_{\text{in}}^{\text{on}}(t) = \frac{\alpha_3 - Q_p}{\alpha_2} - \left[\frac{\alpha_3 - Q_p}{\alpha_2} - T_{\text{in}}^{\text{on}}(0)\right] \cdot e^{-(\alpha_2/\alpha_1)\cdot t}. \tag{4}$$

When the central air conditioning system is in shutdown period, its indoor energy conservation equation is

$$\rho_a \cdot V_k \cdot c_a \cdot \mathrm{d}T_{\text{in}}^{\text{off}}(t) = Q_{t1} \cdot \mathrm{d}t + Q_{\text{er}} \cdot \mathrm{d}t + Q_{\text{new}} \cdot \mathrm{d}t - Q_{t2} - Q_z^{\text{off}} \cdot \mathrm{d}t. \tag{5}$$

By solving the above first-order differential equation, the indoor temperature change equation during the shutdown period can be obtained, as follows:

$$T_{\text{in}}^{\text{off}}(t) = \frac{\alpha_3}{\alpha_2} + \frac{\alpha_4}{a \cdot \alpha_1 - \alpha_2} \cdot e^{-a \cdot t} - \left[\frac{\alpha_4}{a \cdot \alpha_1 - \alpha_2} + \frac{\alpha_3}{\alpha_2} - T_{\text{in}}^{\text{off}}(0)\right] \cdot e^{-(\alpha_2/\alpha_1)\cdot t}. \tag{6}$$

Specific expressions of some variables in the above equation are as follows:

$$\begin{cases} \alpha_1 = \rho_a \cdot V_k \cdot c_a + k_s \cdot S_{\text{wall}}^{\text{in}}, \\ \alpha_2 = 1.8 \cdot \left(k_{\text{top}} \cdot S_{\text{top}} + k_{\text{wall}} \cdot S_{\text{wall}}\right), \\ \alpha_3 = \alpha_2 \cdot T_{\text{out}} + Q_{\text{er}}, \\ \alpha_4 = a \cdot m_z \cdot c_w \cdot \left(T_{\text{w-out}} - T_{\text{w-in}}\right), \end{cases} \tag{7}$$

where V_k is the indoor volume; c_a is the specific heat capacity of air; k_s is the heat storage coefficient of the inner wall; $S_{\text{wall}}^{\text{in}}$ is the area of the inner wall; k_{top} and k_{wall} are the coefficients of heat conduction of the roof and wall, respectively; S_{top} and S_{wall} are the area of the roof and wall; m_z is the quality of frozen water; c_w is the specific heat capacity of frozen water; $T_{\text{w-in}}$ and $T_{\text{w-out}}$ are, respectively, the inlet and outlet water temperature of frozen water; and Q_{er} is the total cooling load of indoor equipment, lighting, and personnel.

(3) Upper and lower limits on room temperature changes

$$\begin{cases} T_{\text{in}}(k,t) = T_{\text{in}}^{\text{on}}(k,t) \cdot S(k,t) + T_{\text{in}}^{\text{off}}(k,t) \cdot [1 - S(k,t)], \\ T_{\min}(t) \le T_{\text{in}}(k,t) \le T_{\max}(t), \end{cases} \quad \left(t = T_{p\_\text{cool}}, \ldots, T_{e\_\text{start}}, \ldots, T_{e\_\text{end}}\right), \tag{8}$$

where $T_{\max}(t)$ and $T_{\min}(t)$ represent the upper and lower limits of the temperature range at each time after the load group of central air conditioning is regulated by load aggregators. T_{p_cool} represents the start time of the advance notice period. Therefore, if the load aggregators use precooling control strategy, the advance notice period is $t = T_{p\_\text{cool}}, \ldots, (T_{e\_\text{start}} - 1)$; otherwise, no advance notice is required, and at this time, $T_{p\_\text{cool}} = T_{e\_\text{start}}$.

(4) Minimum operating time constraint of refrigerator

$$\begin{cases} \sum_{t=T_{p\_\text{cool}}}^{T_{p\_\text{cool}}+G_k-1} [1 - S(k,t)] = 0, \\ \sum_{t=m}^{m+UT_k-1} S(k,t) \ge UT_k [S(k,n) - S(k,n-1)] \quad \left(\forall m = T_{p\_\text{cool}} + G_k, \ldots, T_{e\_\text{end}} - UT_k + 1\right), \\ \sum_{t=n}^{T_{e\_\text{end}}} \{S(k,t) - [S(k,t) - S(k,t-1)]\} \ge 0 \quad \left(\forall n = T_{e\_\text{end}} - UT_k + 2, \ldots, T_{e\_\text{end}}\right), \\ G_k = \min\left\{\left(T_{e\_\text{end}} - T_{p\_\text{cool}} + 1\right), \left[UT_k - U_k^0\right] S(k,0)\right\}. \end{cases} \tag{9}$$

In equation (9), UT_k represents the minimum operating time of the k-th central air conditioning refrigerating machine; U_k^0 represents the initial operating time of the k-th central air conditioning refrigerating machine; $S(k,0)$ represents the initial operating state of the k-th central air conditioning refrigerating machine; and G_k refers to the required operating time at least of the refrigerating machine in order to keep continuous with the operating state before control during the initial period after the start of the control. The third expression in equation (9) is to ensure that the operation state of the central air conditioning unit remains continuous before the end of the controlled state.

(5) Minimum downtime constraint of the refrigerator

$$\begin{cases} \sum_{t=T_{p\_\text{cool}}}^{T_{p\_\text{cool}}+L_k-1} S(k,t) = 0, \\ \sum_{t=m}^{m+DT_k-1} [1 - S(k,t)] \ge \text{DT}_k [S(k,n-1) - S(k,n)] \quad \left(\forall m = T_{p\_\text{cool}} + L_k, \ldots, T_{e\_\text{end}} - \text{DT}_k + 1\right), \\ \sum_{t=n}^{T_{e\_\text{end}}} \{1 - S(k,t) - [S(k,t-1) - S(k,t)]\} \ge 0 \quad \left(\forall n = T_{e\_\text{end}} - DT_k + 2, \ldots, T_{e\_\text{end}}\right), \\ L_k = \min\left\{\left(T_{e\_\text{end}} - T_{p\_\text{cool}} + 1\right), \left[\text{DT}_k - Q_k^0\right] [1 - S(k,0)]\right\}. \end{cases} \tag{10}$$

In the above formula, DT_k represents the minimum downtime of the k-th central air conditioning refrigerator; Q_k^0 represents the initial shutdown time of the k-th central air conditioning refrigerating machine; L_k refers to the required shutdown time at least of the refrigerating machine in order to keep continuous with the operating state before control during the initial period after the start of the control. The third expression in equation (10) is to ensure that the operation state of the central air conditioning units remains continuous before the end of the controlled state.

(6) Constraints on the value range of decision variables

$$\begin{cases} S(k,t) \in \{0,1\} \quad \left(t = T_{\text{pre-cooling}}, \ldots, T_{e\_\text{start}}, \ldots, T_{e\_\text{end}}\right), \\ 0 \le \text{num\_pot} \le \min\left(\dfrac{P_{\text{base}}(t)}{P_{\text{air}}}\right) \quad \left(t = T_{e\_\text{start}}, \ldots, T_{e\_\text{end}}\right). \end{cases} \tag{11}$$

The goal of the above central air conditioning load group regulation potential optimization model is to find the maximum number (num_pot) of central air conditioning refrigerators to be cut corresponding to the regulation potential in the objective function (1) and the operating state ($S(k,t)$) of each central air conditioning refrigerator at every moment controlled by the load aggregator under the condition that the constraints are satisfied.

The refrigerator's start-stop state matrix $S(k,t)$ contains the advance notice period of all central air conditioning refrigerators when the precooling regulation strategy is adopted and the state information of the emergency demand response event duration period. According to the state matrix obtained by optimization, the load aggregator can realize the regulation potential by adjusting the operating state of each central air conditioning refrigerator in these two periods. In addition, the room temperature will drop to the lower limit of room temperature at the end of the advance notice period to realize prerefrigeration; during the duration of emergency demand response events, reduce the air conditioning load according to the regulation potential obtained by optimization.

After the emergency demand response event, the central air conditioning refrigerators will be artificially and orderly controlled to gradually enter the operation state, so that the room temperature will return to the original room temperature range. When the room temperature enters the original room temperature range, the artificial control of the running state of the refrigerator will be cancelled.

In the model established above, only the relevant constraints of the refrigerator are considered; that is, the refrigerator is mainly used to achieve load reduction, which is a reasonable control method based on the consideration of communication and control complexity. If greater reduction is required, the air supply unit can be considered for regulation. When optimizing the air supply unit, the model is as follows.

The objective function:

$$\max\left(P_{\text{pot}}\right) = \max\left(\min\left(P_{\text{base}}(t) - \left(P_{\text{air}} \cdot \sum_{k=1}^{n} S_{\text{air}}(k,t) + P_{\text{fan}} \cdot \sum_{k=1}^{n} S_{\text{fan}}(k,t)\right)\right)\right) \quad \left(t = T_{e\_\text{start}}, \ldots, T_{e\_\text{end}}\right). \tag{12}$$

where $P_{\text{base}}(t)$ is the baseline load taking into account the power of the forced draught blower; P_{air} and P_{fan} are, respectively, the power of refrigerator and air supply system; $S_{\text{air}}(k,t)$ and $S_{\text{fan}}(k,t)$ are, respectively, the operating states of the refrigerating machine and air supply system in the k-th building at time t. Therefore, the operating power of commercial buildings' air conditioning load group in time period t is jointly determined by the operating states of refrigerators and air supply systems in each building, as shown in the following formula:

$$P_{\text{total}} = P_{\text{air}} \cdot \sum_{k=1}^{n} S_{\text{air}}(k,t) + P_{\text{fan}} \cdot \sum_{k=1}^{n} S_{\text{fan}}(k,t). \tag{13}$$

Taking into account the regulation potential P_{pot} of air conditioning load group of commercial buildings controlled by the air supply system, the expression is as follows:

$$P_{\text{pot}} = \min\left(P_{\text{base}}(t) - \left(P_{\text{air}} \cdot \sum_{k=1}^{n} S_{\text{air}}(k,t) + P_{\text{fan}} \cdot \sum_{k=1}^{n} S_{\text{fan}}(k,t)\right)\right). \tag{14}$$

It is the minimum value of load reduction of commercial building air conditioning load group during emergency demand response event.

The constraints are as follows:

(1) Room temperature time-varying equation

Indoor energy conservation equation:

$$\rho_a \cdot V_k \cdot c_a \cdot \mathrm{d}T_{\mathrm{in}}^{\mathrm{off}}(t) = Q_{t1} \cdot \mathrm{d}t + Q_{\mathrm{er}} \cdot \mathrm{d}t + Q_{\mathrm{new}} \cdot \mathrm{d}t - Q_{t2}. \tag{15}$$

By solving the first-order differential equation, the room temperature time-varying equation under this condition can be obtained:

$$T_{\mathrm{in}}^{\mathrm{off2}}(t) = \frac{\alpha_3}{\alpha_2} - \left[\frac{\alpha_3}{\alpha_2} - T_{\mathrm{in}}^{\mathrm{off}}(0)\right] \cdot e^{-(\alpha_2/\alpha_1)\cdot t}. \tag{16}$$

(2) Coupling constraints of the operating state of refrigerating unit and air supply system

$$S_{\mathrm{air}}(k,t) - S_{\mathrm{fan}}(k,t) \le 0. \tag{17}$$

It can be seen from the above formula that when the refrigerator is in operation state ($S_{\mathrm{air}}(k,t) = 1$), the air supply system must also be in operation state ($S_{\mathrm{fan}}(k,t) = 1$); when the refrigerator is off ($S_{\mathrm{air}}(k,t) = 0$), the air supply system can be in either running state ($S_{\mathrm{fan}}(k,t) = 1$) or off state ($S_{\mathrm{fan}}(k,t) = 0$).

(3) Constraint of running time of the air supply system

$$\sum_{t=T_{\mathrm{pre-cool}}}^{T_{\mathrm{end}}} \left(S_{\mathrm{fan}}(k,t) - S_{\mathrm{air}}(k,t)\right) \le T_{\mathrm{fan\_on}} \cdot \mathrm{Times}_{\mathrm{on-off}}. \tag{18}$$

The operating time of the air supply system is related to the change of the residual cooling capacity of chilled water. For the convenience of modeling, the running time of the air supply system is simplified to a fixed time $T_{\mathrm{fan\_on}}$; that is, the running time of the air supply system shall not exceed $T_{\mathrm{fan\_on}}$ during the shutdown period of the central air conditioning refrigerator.

2.2. Analysis of Load Response Characteristics of Residential Users. The objective function is consistent with that of air conditioning load in the commercial buildings. By controlling the air conditioning load of residential users, the maximum load reduction goal is achieved under the condition of satisfying the corresponding constraints, so that the potential of the air conditioning cluster can be fully developed.

In addition to one of the above objective functions, due to the variety and quantity of residents' air conditioning loads, in order to reduce the impact on users, the corresponding number of air conditioning units should be reduced as little as possible under the constraints of the maximum load reduction. The specific objective function is shown as follows:

$$\min(x_{\mathrm{run}}) = \min\left(\sum_{k=1}^{n} S(k,t)\right) \quad (t = T_{e\_\mathrm{statrt}}, \ldots, T_{e\_\mathrm{end}}). \tag{19}$$

The meanings of the above parameter are consistent with those in the air conditioning load model of commercial buildings where x_{run} is the newly set variable representing the number of running machines at each time, and the objective function is added to make it affect the air conditioning load as little as possible under the condition of the maximum load reduction.

For constraint, it also includes the above constraints, namely, load reduction constraints during emergency demand response events, room temperature time-varying equations, constraints on upper and lower limits of room temperature variation, constraints on minimum operating time of air conditioning, constraints on minimum downtime of air conditioning, constraints on the value range of decision variables, etc. Most of the constraints contain the same contents as the air conditioning cluster load in commercial buildings. The main difference is that the start-stop state matrix $S(k,t)$ of the air conditioning load only contains the state information of all residents' air conditioners during the duration of emergency demand response events but does not include the state in the advance notice period when the prerefrigeration control strategy is used. This is mainly because prerefrigeration measures for residential air conditioning load do not have its practicality. Similarly, according to the state matrix obtained by optimization, the load aggregator regulates the operating state of each resident's air conditioner during this period, so as to realize the regulation potential and reduce the air conditioning load according to the regulation potential obtained by optimization during the duration of emergency demand response events.

3. Establish an Optimal Control Model for the Flexible Load Participating in the Power Grid Dispatching

3.1. Analysis Method of Flexible Load Regulation Potential. In this paper, Xinjiekou Zhuanjinying Line # 1(hereinafter referred to as Jinying # 1) is selected as a typical representative of the power characteristics of the courts to analyze the load characteristics.

Table 1 shows the linear regression model summary of the load under the meteorological factor Jinying # 1 feeder in summer (June, July, and August), focusing on the value of the determination coefficient R squared. This coefficient indicates the percentage of the variance explained by the independent variable in the total variance. The greater the value, the greater the role of this factor. In the following table, the determination coefficient R reaches 71.7%, indicating that the regression relationship can explain the 71.7% variation of the dependent variable and has a good fitting degree.

Table 1: Summary table of the linear regression model of Jinying # 1 line load in summer.

Model	R	Adjusted R	Error in standard estimates
1	0.717	0.703	18.423

An analysis of variance was conducted on the model. The result of the analysis of variance showed that the F value was 54.974, and the P value was less than 0.05. Therefore, the model was statistically significant or valuable for predicting variables.

The Table 2 gives the estimated value and test results of the constant term and regression coefficient in the regression equation. It can be seen that the constant term = −31.001, and the four regression coefficients are 139.900, 30.094, 12.826, and 7.430, respectively. The regression equation can be written as follows:

$$\begin{aligned}\text{load value} = {} & -31.001 + 139.900 \times \text{temperature factor} + 30.094 \times \text{humidity factor} \\ & + 12.826 \times \text{precipitation factor} + 7.430 \times \text{wind speed factor}.\end{aligned} \tag{20}$$

Table 3 shows a summary of the linear regression model of winter meteorological factor Jinying # 1 load. The R value in the model abstract is 0.859, indicating that the regression relationship can explain 85.9% variation of the dependent variable. In Table 4, Sig value is 0, less than 0.05, indicating that most independent variables in the model are correlated with dependent variables, indicating that the model has statistical significance.

Table 5 gives the estimated value and test results of the constant term and regression coefficient in the regression equation. The regression equation can be written as follows:

$$\begin{aligned}\text{load value} = {} & 103.295 - 32.514 \times \text{temperature factor} - 2.414 \times \text{humidity factor} \\ & - 5.539 \times \text{precipitation factor} - 0.884 \times \text{wind speed factor}.\end{aligned} \tag{21}$$

According to the above construction method of the multifactor load characteristic regression model, each feeder can be independently analyzed, and the corresponding regression model of each feeder can be obtained, so that the feeder load value can be accurately predicted according to meteorological factors. According to different models in spring and summer, the load baseline mean values of feeders in spring and summer can be obtained. According to the difference between the two, the theoretical maximum load reduction potential can be obtained. According to this value, different load reduction potential targets can be formulated. After a large number of simulation calculations, the load reduction potential of each feeder under different prerefrigeration duration, different control duration, and different room temperature can be obtained. Finally, the flexible load regulation potential table can be obtained, as shown in Table 6.

3.2. Establish an Optimal Combination Control Model of Flexible Load Based on the Maximum Comprehensive Benefit. When the grid dispatching needs to regulate an interactive resource to complete a certain interactive task, the grid needs to pay the corresponding cost to the user, that is, the interactive cost of flexible load dispatching on grid side, no matter whether the price type load or the incentive load responds. At this time, with the change of the relationship between electricity price and power, the income of power grid will also change.

When the interactive resources on the flexible load side respond according to the price signal or incentive measures sent by the power grid, the electricity expense of the flexible load side will be different while the relationship between the electricity price and power changes. The variation of electricity expense is the interactive benefit of the flexible load side load response. Accordingly, the inverse value of the interaction benefit of load response represents the interactive cost of load response on the flexible load side. At the same time, the mode of electricity consumption will also change accordingly, which will affect the satisfaction of the mode of electricity consumption.

C_{inter}, the comprehensive interaction cost of flexible load demand response, can be expressed as

$$\begin{cases} C_{\text{inter}} = \sum\limits_{t \in n_T} \left(\sum\limits_{i \in n_p} C_{i,t}^{p} + \sum\limits_{j \in n_I} C_{j,t}^{I} + \sum\limits_{m \in n_G} C_{m,t}^{g} \right), \\ C_{m,t}^{g} = a_m \left(P_{m,t}^{g} \right)^2 + b_m P_{m,t}^{g} + \lambda_{m,t} c_m, \end{cases} \tag{22}$$

where C_{inter} is the comprehensive interaction cost of flexible load demand response; $C_{i,t}^{p}$ is the scheduling cost of power-price type flexible load i in time period t; $C_{j,t}^{I}$ is the dispatching cost of incentive flexible load j in time period t; n_p is the quantity of flexible load of electricity price type; n_I is the quantity of excitation type flexible load; n_G is the amount of generators in the system; $C_{m,t}^{g}$ is the generating cost of the generator m in the system during the period of time t; $P_{m,t}^{g}$ is the active power output of the generator m during the period

Table 2: Table of coefficient of the linear regression model of Jinying # 1 line load in summer.

Model		Nonstandard coefficient		Normalized coefficient	T	Sig
		B	Standard error	Beta		
1	(Constant)	−31.001	12.083	—	−2.566	0.012
2	Temperature factor	139.900	9.809	1.397	14.262	0.000
3	Humidity factor	30.094	3.719	0.564	8.092	0.000
4	Precipitation factor	12.826	1.500	0.705	8.550	0.000
5	Wind speed factor	7.430	2.295	0.207	3.238	0.002

Table 3: Summary of the linear regression model of Jinying # 1 line load in winter.

Model	R	Adjusted R	Error in standard estimates
1	0.859	0.817	11.033

Table 4: Variance diagram of the line load linear regression model of Jinying # 1 in winter.

Model	Quadratic sum	Df	Mean square	F	Sig (b)
Regression	6484.935	4	1621.234	13.319	0.000
Residual	5112.456	42	121.725		
Total	11597.391	46			

of time t; a_m, b_m, and c_m are the power generation cost model coefficient of generator m; and $\lambda_{m,t}$ is a variable of 0-1, which is 1 when the generator is put into operation in time period t, and 0 otherwise.

In this paper, the optimal dispatching is carried out aiming at maximizing the comprehensive benefits of "source-network-load." Since the numerical value of the comprehensive satisfaction degree of electricity consumption (R) and the comprehensive interaction cost (C_{inter}) do not belong to the same order of magnitude, it is necessary to normalize the C_{inter} in order to ensure the optimization effect.

The objective function is as follows:

$$\text{Max } F = \text{Max}\left(\frac{1}{n_p}\sum_{i\in n_p} R_i - \frac{C_{\text{inter}}}{\sum_{t\in n_T}\sum_{k\in n_f}\left(P_{k,t}^0 \times p_0\right)}\right), \tag{23}$$

where R_i is the comprehensive electricity consumption satisfaction of power-price type load i; n_f is the total number of flexible loads; p_0 is the original unified power price of the system; and $P_{k,t}^0$ is the original load of the flexible load k during the period of time t. There are mainly three constraints, which are unit constraint, system constraint, and load constraint, and the details are as follows:

(1) Unit constraints

The upper and lower limits of unit output:

$$P_{g,m}^{\min} \le P_{m,t}^g \le P_{g,m}^{\max}, \quad m \in n_G, \tag{24}$$

where $P_{g,m}^{\min}$ and $P_{g,m}^{\max}$ indicate the lower and upper limits of the output of generator $m$separately.

Unit climbing constraint:

$$\begin{cases} \lambda_{m,t}P_{m,t}^g - \lambda_{m,t-1}P_{m,t-1}^g \le U_{g,m}, \\ \lambda_{m,t-1}P_{m,t-1}^g - \lambda_{m,t}P_{m,t}^g \le D_{g,m}, \end{cases} \tag{25}$$

where $U_{g,m}$ and $D_{g,m}$, respectively, represent the upward and downward climbing constraint values of generator m.

Minimum start-stop time constraint of unit:

$$\begin{cases} \left(\lambda_{m,t} - \lambda_{m,t-1}\right)\cdot\left(T_{t-1} - T_{\text{on\_}m}\right) \ge 0, \\ \left(\lambda_{m,t} - \lambda_{m,t-1}\right)\cdot\left(T_{t-1} - T_{\text{off\_}m}\right) \le 0, \end{cases} \tag{26}$$

where T_{t-1} is the duration of period $t-1$ and $T_{\text{on\_}m}$ and $T_{\text{off\_}m}$ are the minimum start-stop time of generator m, respectively.

(2) System constraints

Active power balance constraint:

$$\sum_{k\in n_f} P_{k,t} = \sum_{m\in n_G} P_{m,t}^g, \tag{27}$$

where $P_{k,t}$ is the load of the flexible load k during the period of time t after participating in the demand response.

Network constraints:

$$\begin{aligned} &\left|-b_{xy}\times\left(\theta_{x,t} - \theta_{y,t}\right)\right| \le L_{xy}, \quad \forall t \in n_T, \\ &\qquad \forall x, y \in \{1, 2, \ldots, N\}, \quad \coprod x \ne y, \end{aligned} \tag{28}$$

where b_{xy} is the electrical susceptance of line $x-y$; $\theta_{x,t}$ and $\theta_{y,t}$ separately represent the phase angles of nodes x and y; L_{xy} is the maximum active power transmission capacity for line $x-y$; and N is the amount of nodes in the system.

(3) Load constraint

There are mainly two load constraints which are load maximum adjustment constraints and excitation response invocation time constraints, and the details are as follows:

$$\begin{aligned} P_k^{\min} \le P_{k,t} \le P_k^{\max}, \quad k \in n_f, \\ 0 \le \text{num}_j \le n_{I\_j}^{\max}, \quad j \in n_I, \end{aligned} \tag{29}$$

where $P_k^{\min}$ and $P_k^{\max}$ are the minimum and maximum of the incentive flexible load j, respectively; num_j is the number of times that the incentive flexible load j is actually called; and $n_{I\_j}^{\max}$ is the maximum number of times that the incentive flexible load j is called.

Table 5: Table of coefficient of the linear regression model of Jinying # 1 line load in winter.

	Model	Nonstandard coefficient		Normalized coefficient	T	Sig
		B	Standard error	Beta		
1	(Constant)	103.295	2.247	—	45.978	0.000
2	Temperature factor	−32.514	5.042	−0.698	−6.448	0.000
3	Humidity factor	−2.414	1.520	−0.170	−1.588	0.120
4	Precipitation factor	−5.539	3.026	−0.191	−1.830	0.004
5	Wind speed factor	−0.884	2.205	−0.042	−0.401	0.691

Table 6: Flexible load regulation potential table.

Outdoor temperature 31–33°C				
Precooling duration (min)	Duration of event (min)	Upper limit of room temperature (°C)	Load reduction potential (MW)	Compensation fees (yuan)
0	30	26	55	25740
		27	99	54054
		28	115.5	72072
		29	115.5	81081
	40	26	44	29568
		27	77	60368
		28	115.5	103488
		29	115.5	116424
10	30	26	115.5	59459.4
		27	115.5	69369.3
		28	115.5	79279.2
		29	115.5	89189.1
	40	26	60.5	44721.6
		27	99	85377.6
		28	115.5	113836.8
		29	115.5	128066.4

The dispatcher can get the optimal nodal power price, load reduction period and load reduction amount of each incentive flexible load, and the output arrangement of each power generating unit in each period through optimization, so as to issue scheduling instructions.

4. A Comprehensive Optimization Control Strategy of Flexible Load for Courts

Through the above analysis, the comprehensive optimization control strategy of flexible load under different time scales can be obtained to eliminate the line overload problem.

4.1. Medium- and Long-Term Regulation of Flexible Loads. For medium- and long-term load dispatching planning, the main consideration is the demand response model of load under the electricity price incentive of time-of-use electricity price and interruptible load. Its model is mainly manifested as response modeling of cumulative electric quantity at various periods of peak and valley level:

$$\begin{cases} q_{\mathrm{Ap}} = \sum_{i=1}^{n} q_{\mathrm{Aip}}, q_{\mathrm{Am}} = \sum_{i=1}^{n} q_{\mathrm{Aim}}, q_{\mathrm{Av}} = \sum_{i=1}^{n} q_{\mathrm{Aiv}}, \\ q_{\mathrm{Aip}} = q_{\mathrm{Aip0}} + \dfrac{\varepsilon_{\mathrm{ppi}} q_{\mathrm{Aip0}} \left(p_{\mathrm{TOUip}} - p_{\mathrm{TOUip0}}\right)}{p_{\mathrm{TOUip0}}} + \dfrac{\varepsilon_{\mathrm{pvi}} q_{\mathrm{Aip0}} \left(p_{\mathrm{TOUiv}} - p_{\mathrm{TOUiv0}}\right)}{p_{\mathrm{TOUiv0}}} - \theta_i \Delta L_{\mathrm{IL}i} \left(t_{2i} - t_{1i}\right), \\ q_{\mathrm{Aim}} = q_{\mathrm{Aim0}}, \\ q_{\mathrm{Aiv}} = q_{\mathrm{Aiv0}} + \dfrac{\varepsilon_{\mathrm{vvi}} q_{\mathrm{Aiv0}} \left(p_{\mathrm{TOUiv}} - p_{\mathrm{TOUiv0}}\right)}{p_{\mathrm{TOUiv0}}} + \dfrac{\varepsilon_{\mathrm{vpi}} q_{\mathrm{Aiv0}} \left(p_{\mathrm{TOUip}} - p_{\mathrm{TOUip0}}\right)}{p_{\mathrm{TOUip0}}}, \end{cases} \tag{30}$$

Table 7: Load expectation and variance at different time scales.

Time scale	Electricity price incentives	Expectation	Variance
Medium and long term	Time-of-use electricity price + interruptible load	$\begin{cases} q_{\mathrm{Ap}} \\ q_{\mathrm{Am}} \\ q_{\mathrm{Av}} \end{cases}$	$\begin{cases} \sigma^2_{\mathrm{Ap}} = \sigma^2_{\mathrm{AMC}} \\ \sigma^2_{\mathrm{Am}} = 0.264 \ln(\Delta q_{\mathrm{Am}} + 1) \\ \sigma^2_{\mathrm{Av}} = 0.264 \ln(\Delta q_{\mathrm{Av}} + 1) \end{cases}$
Day-ahead	Day-ahead real-time electricity price + interruptible load	P_{Ct}	$\sigma^2_C = \begin{cases} \sigma^2_{\mathrm{CMCt}} & t_1 \le t \le t_2 \\ 0.264 \ln(\Delta L_{\mathrm{Ct}} + 1) & \text{else} \end{cases}$
Real-time	Real-time electricity	P_{Dt}	$\sigma^2_D = 0.264 \ln(\Delta L_{\mathrm{Dt}} + 1)$

Note: σ^2_{AMCt}, σ^2_{BMCt}, and σ^2_{CMCt} represent the load variance values obtained by Monte Carlo simulation.

where n represents the number of loads under a certain node; θ_i is the probability of the i-th load signing an interruptible contract; $\Delta L_{\mathrm{IL}i}$ is the amount of interruptible load of load i; and $t_{1i} \sim t_{2i}$ represents the start-stop time of the interruptible load for load i.

4.2. Day-Ahead Regulation of Flexible Load. For day-ahead load dispatching planning, the main consideration is the demand response model under day-ahead real-time power price and interruptible load:

$$\begin{aligned} P_{\mathrm{Ct}} &= \sum_{i=1}^{n} P_{\mathrm{Cit}}, \\ P_{\mathrm{Cit}} &= P_{\mathrm{Cit}}(p_{\mathrm{SPit}}) - \Delta P_{\mathrm{IL}i}. \end{aligned} \tag{31}$$

4.3. Real-Time Regulation of Flexible Load. For real-time load dispatching planning, the demand response model of load under the incentive of real-time power price is considered. Its main features are shown as follows:

(1) Price elasticity of demand will be different while power price changes.

(2) The load at the next moment is the result of superposition of load response components based on the load at the previous moment.

$$\begin{aligned} P_{\mathrm{Dt+1}} &= \sum_{i=1}^{n} P_{\mathrm{Dit+1}}, \\ P_{\mathrm{Dit+1}} &= P_{\mathrm{Dit}} + \frac{\varepsilon_i(p_{\mathrm{SPit+1}}) P_{\mathrm{Dit}}(p_{\mathrm{SPit+1}} - p_{\mathrm{SPit}})}{p_{\mathrm{SPit}}}, \\ \varepsilon_i(p_{\mathrm{SPit+1}}) &= \frac{\partial P_i(p_{\mathrm{SPit+1}})}{\partial p_{\mathrm{SPit+1}}}. \end{aligned} \tag{32}$$

According to the above analysis, the expectation and variance of load under different time scale responses can be obtained, as shown in Table 7.

5. Example Analysis

The IEEE 33-node system in the literature [26] was selected for example analysis. The IEEE 33-node system is shown in Figure 1. There are 5 generators in the system, which are located on nodes 8, 12, 17, 19, and 22. Tables 8–10 show the system's nodes, active power output, circuit, and load parameters. The unit of cost coefficient a, b, c are \$/MW2, \$/MW, and \$, respectively, and the units of impedance and ground capacitance are Ω.

In Table 10, all users can participate in price-based demand response. Users 1 to 10 participate in incentive demand response, which is divided into translational, reducible, and transferable classes. Take $\mu_1 = \mu_2 = 0.5$, take $K = 1, \alpha = 0.15$, and the original price $p_0 = 42.7$ in the electricity price response model. When the breaker breaking capacity constraint is not taken into account, the optimal power flow is adopted to calculate and obtain the active power output data of each generator in 24 periods, as shown in Table 11. The active power flow of each branch is shown in Table 12.

The maximum breaking capacity of circuit breakers on each line is 300 MW. It can be seen from Table 12 that the power flow of lines 7-8 and 7–9 exceeds the limit in some time periods. To eliminate the network congestion of lines 7-8 and 7–9, four different scenarios are constructed in the literature [26]. Scenario settings are shown in Table 13.

In the four scenarios, MATLAB was used for simulation calculation. The computer CPU was Intel(R) Core(TM) I5-4590, the main frequency was 3.3 GHz, and the memory was 8 GB. The optimized simulation time in each scenario is 10.57 s, 24.07 s, 28.1 s, and 57.03 s, respectively.

The optimized active power flow curves of lines 7-8 and 7–9 are shown in Figures 2 and 3.

In Scenario 1, the output of the generator is adjusted without considering the demand response. The optimized output of each generator node is shown in Table 14.

As can be seen from Figures 2 and 3, in Scenario 1, the power flow of the blocked line can only be reduced to 340 MW by adjusting the output distribution of the power generation side to the minimum, and the problem of power flow beyond the limit cannot be eliminated. At the same time, it can be seen from Table 14 that when the overload of line active power flow is eliminated, the variation of active power output of generator nodes 17 and 19 is too large, which is difficult to be realized in the actual power grid.

However, in Scenarios 2, 3, and 4, due to the implementation of integrated dispatching for power generation and consumption side at the same time, the off-limit power flow is under ideal control, and under this dispatching mode, the blocked line still has room for power flow to fall.

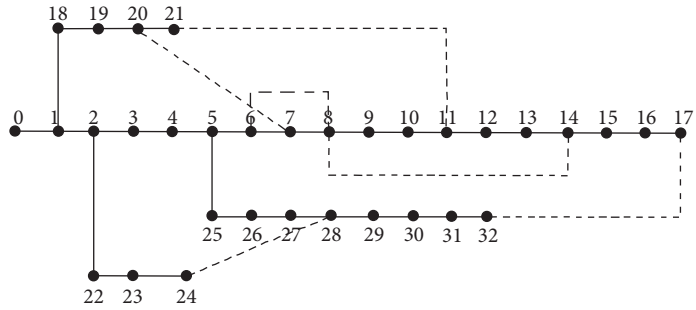


Figure 1: IEEE 33-node system.

Table 8: Active power output data of the IEEE 33-node system.

Node number	Upper limit of active power output/(MW)	Lower limit of active power output/(MW)	Cost coefficient		
			a	b	c
8	300	0	0	1	300
12	300	0	0	2	300
17	350	0	0	3	350
19	350	0	0	6	350
22	500	0	0	8	500

5.1. Interactive Cost Analysis of Flexible Loads. The interaction costs of flexible loads in each scenario are shown in Table 15.

The comparison of unit power generation costs in each scenario is shown in Table 16:

Since the network congestion cannot be eliminated only through generation side dispatching in Scenario 1, the interaction cost in Scenario 1 is not analyzed.

As can be seen from Table 15, in Scenario 4, when electricity price and incentive comprehensive response scheduling are adopted, the total interactive cost is the minimum. This is mainly because the collaborative scheduling on the generating and consuming side can effectively reduce the output of high-cost units on the basis of ensuring the power consumption and increase the output of the low-cost units, so as to effectively reduce the total cost of power generation and effectively reduces the cost of electricity price response and incentive response of load under the comprehensive action of electricity price and incentive at the same time.

And in Scenario 3, although the interaction response cost on the load side (electricity price response cost + incentive response cost) is the lowest, it calls more high-cost units (the total output of units 8 and 12 is 4756 MW in Scenario 3, only 3961 MW in Scenario 4), making its unit power generation cost (from Table 16) higher, pushing up the total cost of the interaction. In addition, it can be found from the data that, under the same dispatching adjustment effect, the total interactive cost of electricity price response is obviously smaller than that of incentive response.

5.2. Comparison of Comprehensive Satisfaction with Electricity Consumption. Under different scenarios, the comprehensive satisfaction of electricity consumption is shown in Table 17:

Table 9: Line parameters of the IEEE 33-node system.

Branch circuit	Branch impedance	Ground capacitance
0-1	$0.0922 + j0.047$	0.0412
1-2	$0.01938 + j0.05917$	0.0528
2-3	$0.04699 + j0.19797$	0.0438
3-4	$0.05811 + j0.17632$	0.0374
4-5	$0.05403 + j0.22304$	0.0492
5-6	$0.05695 + j0.17388$	0.0340
6-7	$0.06701 + j0.17103$	0.0346
7-8	$0.00000 + j0.17615$	0.0000
7-9	$0.00000 + j0.11001$	0.0000
8-9	$1.0440 + j0.7400$	0.0000
9-10	$0.1966 + j00650$	0.0000
10-11	$0.3744 + j0.1238$	0.0000
11-12	$1.4680 + j1.550$	0.0000
12-13	$0.5416 + j0.7129$	0.0000
13-14	$0.5910 + j0.5620$	0.0000
14-15	$0.7463 + j0.5450$	0.0000
15-16	$1.2890 + j1.7210$	0.0000
1-18	$0.09498 + j0.1989$	0.0000
18-19	$0.12291 + j0.25581$	0.0000
19-20	$0.06615 + j0.13027$	0.0000
20-21	$0.12711 + j0.27038$	0.0000
2-22	$0.08205 + j0.19207$	0.0000
22-23	$0.22092 + j0.19988$	0.0000
23-24	$0.17093 + j0.34802$	0.0000
5-25	$0.2030 + j0.1034$	0.0000
25-26	$0.2842 + j0.1447$	0.0000
26-27	$1.0590 + j0.9337$	0.0000
27-28	$0.8042 + j0.7006$	0.0000
28-29	$0.5075 + j0.2585$	0.0000
29-30	$0.9744 + 0.9630$	0.0000
30-31	$0.3105 + j0.3619$	0.0000
31-32	$0.3410 + j0.5362$	0.0000

As can be seen from the chart, since there is no electricity response in Scenarios 1 and 3, electricity comprehensive satisfaction is 1, and the comprehensive satisfaction with

Table 10: Load demand response attributes of each user.

Load number	Maximum /average load	Load start time	Load end time	Minimum running time	Minimum /average load	Minimum downturn time	Incentive scheduling cost (first stage)	Incentive scheduling cost (second stage)	Incentive scheduling cost (third stage)	Stand-by cost	Total limit calls number of stimulated DR	Maximum reserve capacity of incentive DR	PDR/IDR	Load type
1	1.15	1	24	4	0.75	4	5	10	15	3	8	200	PDR&IDR	Translational
2	1.15	1	24	3	0.75	3	5	10	20	3	8	30	PDR&IDR	Transferable
3	1.15	1	24	2	0.75	2	5	10	25	3	8	30	PDR&IDR	Reducible
4	1.15	1	24	2	0.75	2	5	15	30	3	8	40	PDR&IDR	Reducible
5	1.15	1	24	1	0.75	0	50	50	50	20	1	80	PDR&IDR	Reducible
6	1.15	1	24	2	0.75	2	4	9	14	3	8	200	PDR&IDR	Translational
7	1.15	1	24	3	0.75	3	4	9	19	2	8	30	PDR&IDR	Transferable
8	1.15	1	24	2	0.75	2	4	9	24	2	8	30	PDR&IDR	Reducible
9	1.15	1	24	2	0.75	2	4	14	29	2	8	40	PDR&IDR	Reducible
10	1.15	1	24	1	0.75	0	50	50	50	20	1	80	PDR&IDR	Reducible
11	1.15	1	24	0	0.75	0	1000	1000	1000	1000	16	0	PDR	
12	1.15	1	24	0	0.75	0	1000	1000	1000	1000	8	0	PDR	
13	1.15	1	24	0	0.75	0	1000	1000	1000	1000	8	0	PDR	
14	1.15	1	24	0	0.75	0	1000	1000	1000	1000	8	0	PDR	
15	1.15	1	24	0	0.75	0	1000	1000	1000	1000	1	0	PDR	

Table 11: Active power output of each generator without considering line overload constraint.

Generator	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
8	25.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	75.0	150.0	225.0	289.4	232.9	157.9	82.9	7.9	0.0	75.0	150.0	225.0	225.0	150.0	75.0	0.0
12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	24.5	91.5	166.5	241.5	221.3	296.3	225.0	150.0	75.0	0.0	74.7	149.7	100.2	166.2	91.2	16.2	0.0
17	62.5	0.0	87.5	0.0	2.8	90.3	177.8	265.3	350.0	340.8	350.0	350.0	350.0	350.0	262.5	175.0	247.1	320.1	350.0	350.0	350.0	324.2	236.7	149.2
19	225.1	255.4	276.0	188.5	149.8	175.0	262.5	350.0	322.7	328.1	350.0	350.0	350.0	350.0	332.8	297.0	350.0	262.5	350.0	350.0	350.0	262.5	175.0	87.5
22	125.0	250.0	375.0	476.8	500.0	407.2	500.0	500.0	500.0	500.0	500.0	500.0	500.0	500.0	500.0	500.0	500.0	500.0	500.0	500.0	500.0	500.0	382.4	396.7

Table 12: Active power flow in 24 periods of each line without considering line capacity constraint.

Circuit	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
0-1	9.2	18.6	23.1	8.9	2.3	−13.6	−4.5	7.4	−24.0	−35.4	−77.3	−89.6	31.2	38.5	49.4	57.3	71.7	92.4	115.2	131.6	91.7	66.3	62.9	45.2
1-2	−44.4	10.0	37.9	30.6	26.1	28.6	45.2	55.5	43.6	11.6	−21.9	−43.1	−35.8	1.3	30.4	56.6	57.6	54.8	25.9	9.7	43.4	73.4	39.6	26.7
2-3	−33.3	−7.3	−0.3	5.8	9.6	8.3	2.7	2.6	−19.2	−56.3	−94.7	−114.8	−105.5	−67.9	−28.0	3.2	4.3	−1.7	−40.8	−64.5	−49.4	−10.8	5.6	8.0
3-4	37.8	70.8	95.8	66.0	41.1	54.1	91.0	118.4	117.4	88.9	62.6	48.6	56.2	85.5	105.3	117.6	118.0	124.3	102.8	81.7	95.6	114.8	89.1	54.5
4-5	−34.0	−10.3	−3.7	13.5	26.8	20.9	10.3	6.3	−16.9	−57.8	−99.6	−120.5	−109.6	−69.4	−26.3	3.9	5.4	1.7	−39.8	−68.4	−56.3	−15.1	13.3	19.3
5-6	−27.5	−12.8	−13.3	−3.0	3.5	0.9	−11.9	−15.5	−39.5	−76.1	−114.0	−132.6	−123.1	−87.6	−46.2	−15.2	−14.0	−20.8	−61.2	−86.1	−78.1	−38.8	−6.3	1.1
6-7	−78.8	−92.6	−114.6	−62.5	−20.0	−41.1	−94.8	−130.6	−153.2	−162.4	−175.1	−180.5	−178.0	−170.5	−149.0	−132.0	−131.1	−142.2	−160.0	−165.0	−168.6	−148.5	−89.5	−43.3
7-8	−30.9	86.4	198.7	323.6	449.9	430.5	426.5	417.5	414.3	407.6	386.7	379.9	386.7	400.7	401.6	404.7	405.3	416.5	405.7	396.7	406.4	413.3	421.0	425.8
7-9	−21.9	−75.2	−146.3	−195.2	−252.8	−250.7	−272.0	−283.0	−300.7	−324.1	−349.3	−357.6	−354.1	−340.7	−298.0	−275.1	−275.5	−286.2	−317.2	−331.2	−326.6	−301.1	−262.0	−247.9
8-9	28.8	−9.5	−39.6	−68.7	−97.7	−83.8	−92.2	−90.5	−92.5	−72.3	−53.4	−43.1	−49.3	−71.0	−80.7	−79.0	−80.4	−92.7	−86.3	−69.2	−87.2	−96.8	−81.8	−75.9
9-10	−75.2	−28.5	5.5	80.0	149.3	129.8	100.6	76.4	50.4	14.3	−24.0	−46.3	−31.8	2.6	41.0	75.9	73.9	72.1	35.0	5.6	22.1	55.6	101.9	129.7
10-11	−33.0	−26.2	−27.8	−9.5	5.1	−1.9	−17.4	−28.8	−42.3	−60.7	−80.4	−90.6	−84.3	−68.6	−45.3	−27.3	−28.2	−31.1	−51.5	−65.5	−58.3	−40.4	−14.8	−1.4
11-12	10.8	21.9	22.1	2.0	−18.1	−6.1	−5.5	4.1	−25.0	−64.6	−104.9	−129.2	−114.8	−77.2	−29.7	15.1	13.3	1.1	−46.9	−76.0	−64.7	−21.2	7.7	2.5
12-13	−65.2	−63.8	−74.1	−37.4	−7.8	−24.7	−53.7	−77.0	−86.5	−96.2	−108.2	−110.6	−110.9	−103.2	−88.2	−81.4	−80.1	−82.0	−93.7	−97.9	−96.7	−85.4	−55.6	−28.6
13-14	−30.2	−34.5	−45.5	−38.5	−33.5	−37.7	−51.9	−65.3	−71.6	−79.0	−85.7	−91.5	−88.4	−81.5	−71.4	−61.4	−63.7	−66.3	−75.1	−81.5	−78.9	−71.0	−49.7	−36.1
14-15	−71.5	−78.1	−101.3	−78.8	−62.5	−75.5	−108.1	−139.2	−153.7	−170.3	−188.8	−196.4	−194.3	−179.9	−153.7	−132.4	−137.0	−142.6	−163.1	−174.7	−170.1	−152.2	−104.5	−73.1
15-16	0.0	−13.5	−38.7	−64.3	−91.9	−83.2	−90.0	−84.2	−102.9	−113.9	−127.5	−131.3	−133.6	−120.7	−99.6	−79.1	−76.1	−92.3	−118.4	−117.8	−128.3	−102.4	−79.7	−75.5
16-17	−30.4	−43.9	−76.2	−88.6	−106.0	−105.2	−123.9	−139.6	−153.5	−170.0	−189.3	−196.0	−195.5	−180.3	−150.6	−124.7	−130.5	−139.7	−162.2	−175.2	−170.8	−152.4	−113.3	−96.5
1-18	44.6	38.0	36.5	3.5	−25.4	−9.5	5.8	23.2	23.4	26.1	29.7	30.0	29.4	28.6	25.6	28.4	28.4	23.9	23.1	26.0	21.7	22.8	10.5	−6.1
18-19	−7.9	−6.8	−7.8	−2.1	2.2	−0.9	−4.0	−7.2	−8.5	−9.8	−13.4	−10.9	−13.5	−12.9	−8.8	−7.7	−7.8	−8.2	−10.1	−9.6	−9.9	−8.4	−4.5	−1.4
19-20	−58.8	−59.1	−74.4	−47.0	−27.0	−42.1	−67.8	−92.7	−99.1	−110.1	−123.7	−126.6	−126.3	−118.2	−99.8	−90.3	−93.1	−92.7	−102.5	−112.4	−105.1	−98.0	−67.2	−42.2
20-21	−32.6	−57.2	−63.1	−71.0	−93.4	−112.2	−134.5	−142.8	−167.3	−172.1	−181.6	−194.5	−52.1	−66.3	−68.0	−76.2	−80.2	−86.3	−91.2	−99.4	−83.8	−78.4	−63.2	−55.6
2-22	22.4	11.2	7.6	2.1	−19.6	−11.6	11.2	27.4	32.1	38.9	44.7	51.3	−130.2	−113.5	−101.1	−92.6	−77.3	−62.1	−57.9	−48.5	−39.2	−27.9	−25.3	−22.1
22-23	−52.9	14.0	26.9	37.2	55.4	68.2	48.1	32.4	22.3	9.2	−17.3	−35.2	−178.2	−152.4	−139.5	−122.9	−107.1	−97.4	−81.8	−63.5	−71.6	−77.4	−55.2	−38.5
23-24	44.3	51.2	65.7	82.9	96.4	116.3	137.2	132.6	117.4	78.3	72.1	62.9	−31.5	−29.3	−27.8	−22.9	−19.6	−17.4	−15.3	−14.0	−11.9	−8.2	−6.9	−3.0
5-25	−32.1	−28.9	−22.4	−13.6	−9.4	−3.8	5.6	10.5	−17.2	−44.3	−59.3	−88.9	−152.6	−134.7	−98.0	−92.3	−81.0	−114.9	−108.6	−97.2	−84.2	−62.9	−55.8	−42.6
25-26	−99.7	−103.2	−119.3	−88.4	−54.7	−45.9	−66.9	−78.2	−92.4	−127.3	−130.6	−164.2	−128.9	−110.2	−77.3	−48.6	−18.3	11.8	27.1	−36.0	−44.7	−57.0	−23.6	−16.7
26-27	4.2	22.8	39.6	99.7	118.4	132.6	100.7	83.4	79.3	54.2	21.9	13.4	49.2	55.6	63.9	59.1	78.2	94.3	71.3	99.8	82.5	38.4	43.2	37.1
27-28	−5.6	−23.4	−11.9	7.9	25.7	39.6	10.6	3.2	−17.7	−48.1	−66.2	−98.9	−112.8	−97.6	−88.3	−47.0	−13.9	23.7	13.1	−9.2	39.5	64.7	62.6	44.1
28-29	−46.5	−22.1	−3.4	5.9	14.6	−13.2	−19.4	−31.1	−39.3	−58.3	−72.0	−81.9	22.1	41.0	55.2	71.9	89.2	118.4	46.9	27.4	10.9	54.9	84.6	89.7
29-30	23.2	13.9	7.8	−3.4	−16.7	−20.6	−41.2	−65.9	−72.8	−99.3	−121.7	−128.4	−33.6	−27.4	−19.6	−9.4	8.7	24.9	33.5	42.9	46.4	37.1	36.8	26.5
30-31	−44.2	−34.1	−22.8	−13.2	−3.2	8.3	−4.9	−27.9	−34.1	−48.3	−63.9	−87.9	−182.0	−178.6	−150.4	−144.9	−127.3	−92.9	−75.3	−62.7	−77.9	−65.7	−56.2	−42.6
31-32	32.8	30.5	24.9	16.7	8.7	39.6	48.9	61.3	88.6	92.5	111.9	126.5	31.7	28.4	26.1	22.9	19.3	16.1	15.7	12.5	8.8	7.4	5.0	−5.9

Table 13: Four different scenarios for eliminating network congestion.

Scene sequence	Scene settings
Scenario 1	Eliminating network congestion only by generation side dispatching
Scenario 2	Adjust generation output and implement electricity price dispatch on load side to eliminate network congestion
Scenario 3	Adjust generation output and implement incentive dispatch on load side to eliminate network congestion
Scenario 4	Adjust generation output and implement electricity price and incentive dispatch on load side to eliminate network congestion (integrated dispatch of generation and consumption)

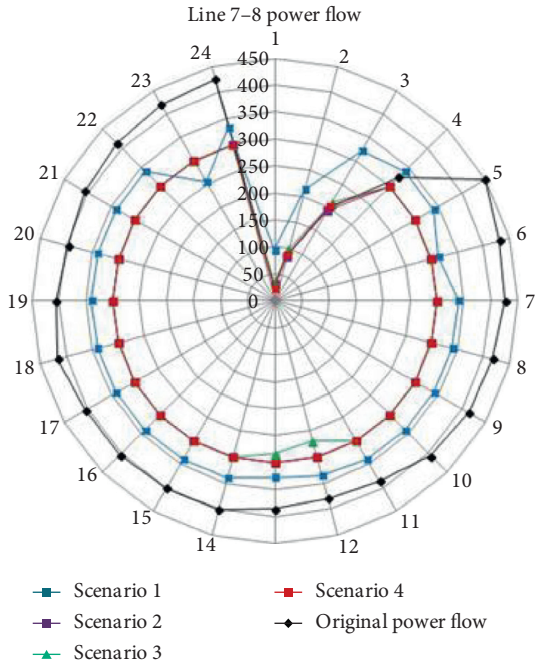


Figure 2: Active power flow curve of line 7-8 after optimized scheduling in different scenarios.

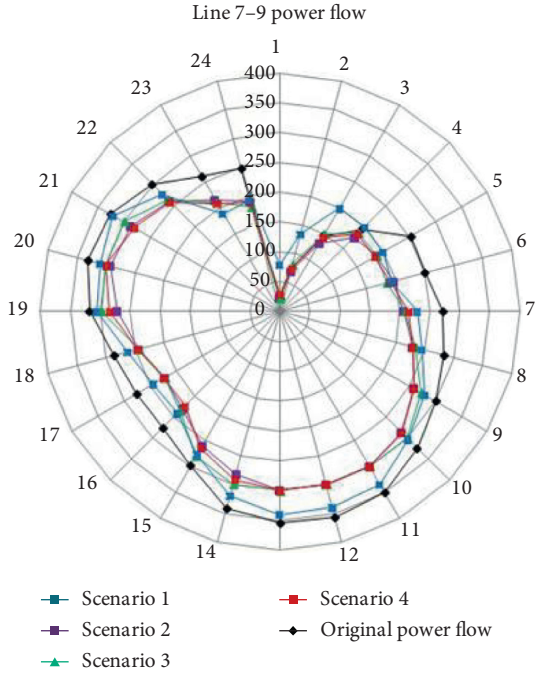


Figure 3: Active power flow curve of line 7–9 after optimized scheduling in different scenarios.

Table 14: The variation of active power output of each generator when only the output regulation of the generator is considered.

Generator	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
8	0.0	0.0	0.0	0.0	0.0	0.0	5.5	27.2	27.2	27.2	27.2	10.6	63.5	67.1	67.1	67.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	−24.5	−27.2	−27.2	−27.2	−10.6	−63.5	−67.1	−67.1	−67.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
17	112.6	230.4	199.4	199.4	187.2	84.7	84.7	84.7	−7.6	0.0	0.0	0.0	0.0	0.0	0.0	6.3	21.7	0.0	0.0	0.0	0.0	−61.7	−61.7	−61.7
19	−112.6	−230.4	−199.4	−188.5	−149.8	−87.5	−87.5	−87.5	7.6	0.0	0.0	0.0	0.0	0.0	0.0	−6.3	−21.7	0.0	0.0	0.0	0.0	61.7	61.7	61.7
22	0.0	0.0	0.0	−10.9	−37.4	2.8	−2.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 15: Comparison of load interaction costs under different scenarios.

Scenario	Total cost of interaction	Total cost of generation	Electricity price response cost	Incentive response cost
Scenario 1	733430	733430	0	0
Scenario 2	749350	733359	15991	0
Scenario 3	759660	751539	0	8121
Scenario 4	743190	731929	9462	1799

Table 16: Comparison of unit power generation cost after response scheduling in different scenarios.

Scenario	Total cost of generation	Gross generation	Unit generation cost
Scenario 1	733430	27440.43	26.73
Scenario 2	733359	27419.21	26.75
Scenario 3	751539	27506.64	27.32
Scenario 4	731929	27593.08	26.53

Table 17: Comparison of comprehensive satisfaction with electricity consumption after response scheduling in different scenarios.

Scenario	Satisfaction of power consumption mode	Interactive benefit satisfaction	Comprehensive satisfaction
Scenario 1	1	1	1
Scenario 2	0.815	1.043	0.929
Scenario 3	1	1	1
Scenario 4	0.987	0.996	0.992

electricity consumption in Scenario 4 is greater than that in Scenario 2. This is mainly due to the addition of incentive response mechanism in Scenario 4, and the incentive scheduling plan is formulated according to the load interruptible plan independently reported by users. In this case, the autonomous power consumption willingness of flexible load can be guaranteed to the maximum extent. Leading to a high degree of satisfaction with its power consumption mode. In Scenario 2, although the user obtains a certain income from electricity fee through the response of electricity price, the mode of electricity consumption changes greatly, which leads to a low degree of comprehensive satisfaction with electricity consumption.

It can be found from the comparison of comprehensive economic benefits and satisfaction that when network congestion occurs, in Scenario 4, the optimal scheduling mode of adjusting power generation output and simultaneously implementing electricity price and incentive scheduling on the load side to eliminate network congestion (integrated comprehensive scheduling of power generation and electricity consumption) is relatively the most appropriate.

6. Conclusions

In this paper, the mathematical model of air conditioning load in commercial buildings and residential users is established, and the response characteristics of commercial buildings and residential users are analyzed. On this basis, Xinjiekou Zhuanjinying Line # is selected as the typical representative of the power consumption characteristics of the courts to analyze the load characteristics, and the flexible load regulation potential table is obtained. A flexible load optimal combination control model based on the maximum comprehensive benefit was established, and through this model, the dispatcher can obtain the optimal node electricity price, load reduction period and load reduction amount of each incentive flexible load, and the output arrangement of each generator set in each period, so as to issue scheduling instructions. Finally, the comprehensive optimization control strategy of the flexible load oriented to the courts is proposed, and the medium- and long term-regulation, day-ahead regulation, and real-time regulation methods of the flexible load are studied, and the IEEE 14-node system is selected for example analysis.

In this paper, the complementary potential of the hybrid energy system and the response priority mechanism of the integrated energy system equipment are analyzed from the aspects of technical and economic factors. The design model of the integrated energy coordinated control is optimized, and the solution analysis is carried out by using the multitime scale cooperative control strategy. Through the analysis of the example, the optimal scheduling results of multitime scale of source storage load are obtained. This study is of great significance for the construction of the hybrid energy system, the promotion of new energy consumption, the increase of energy storage application in power grid, and the realization of more sustainable energy management.

Data Availability

The data used to support findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper

Acknowledgments

This study was supported by the Research and Application of Key Technologies for Accurate Control of Station Load and Smart Energy Consumption for Network-Load Cooperation (5400-202018421A-0-0-00).

References

[1] S. C. Lee and S. J. Kim, S. H. Kim, Demand side management with air conditioner loads based on the queuing system model," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 661–668, 2011.

[2] L. Zhou, L. Yang, and C. Gao, "Improvement of temperature regulation method and control strategy of convergent air conditioning load," in *Proceedings of the 2014 International Conference on Circuits Systems and Electronics Engineering*, Kuala Lumbur, Malaysia, March 2014.

[3] Y. Zhang, *Study on Evaluation and Application of Load Participation System for Peak Load Regulation*, Shandong University, Shadong, China, 2020.

[4] T. W. Haring, J. L. Mathieu, and G. Andersson, "Comparing centralized and decentralized contract design enabling direct load control for reserves," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 2044–2054, 2016.

[5] S. Xujun, J. Bei, L. Wenbo et al., "Investigation and analysis of air conditioning cooling load of residents in summer in Hefei," *Power Demand Side Mangement*, vol. 10, no. 4, pp. 27–29, 2008.

[6] Z. Lu, *Analysis of Load Characteristics and Energy Consumption of Air Conditioning in High-Rise Office Buildings in Chongqing*, Chongqing University, Chongqing, China, 2007.

[7] X. Mingwei, X. Ming, C. Chunhong et al., "Research and forecast modeling of air conditioning load characteristics in summer," *Method*, vol. 52, pp. 2007-2008, 2009.

[8] E. S. Parizy, H. R. Bahrami, and S. Choi, "A low complexity and secure demand response technique for peak load reduction," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 3259–3268, 2019.

[9] T. Vincenzo, T. Fei, and S. Goran, "Role and benefits of flexible thermostatically controlled loads in future low-carbon systems," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5067–5079, 2018.

[10] A. Zhen, H. Yuqi, Y. Xi et al., "Research on load control technology of central air conditioning based on demand response," *Power Communication Management & Smart Grid Communication Technology Forum Proceedings*, vol. 42,, 2012.

[11] M. Liu, X. Chu, W. Zhang et al., "A scheduling control strategy for air conditioning load group based on diversity preservation," in *Proceedings of the International Conference on Circuits Systems and Electronics Engineering, 2014*, Kuala Lumpur, Malaysia, March 2014.

[12] Q. Xu, C. Yang, and Q. Yan, "Day-ahead peak clipping strategy of power load considering thermal balance inertia of large-scale air conditioning," *Power System Technology*, vol. 40, no. 1, pp. 156–163, 2016.

[13] B. J. C. van Putten, N. Mahdavi, J. H. Braslavsky, and Braslavsky, "An analytical model for demand response of variable-speed air conditioners," *IFAC-Papers OnLine*, vol. 51, no. 28, pp. 426–431, 2018.

[14] X. Liu, B. Wang, L. Yang, and Y. Jianguo, "Generation day-ahead scheduling model with user-side interaction in smart grid," *Proceedings of International Conference on Circuits Systems and Electronics Engineering*, vol. 33, no. 1, pp. 30–38, 2013.

[15] L. Won-Woo, "US lessons for energy industry restructuring: based on natural gas and California electricity incidences," *Energy Policy*, vol. 32, pp. 237–259, 2011.

[16] M. Nariman, H. Braslavsky Julio, M. Seron Maria et al., "Model predictive control of distributed air-conditioning loads to compensate fluctuations in solar power," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 3055–3065, 2017.

[17] L. Weijia, W. Qiuwei, W. Fushuan et al., "Market mechanisms for electric vehicles and controllable loads to participate in distribution system congestion management," *Automation of Electric Systems*, vol. 38, no. 24, pp. 26–33, 2014.

[18] X. Ma, W. Chao, H. Xiao, L. Ying, and W. Hao, "Optimization strategy of double-layer charging for electric vehicles based on node blocking price," *Power System Technology*, vol. 40, no. 12, pp. 3707–3714, 2016.

[19] S. Qian, P. Jianchun, P. Juntao et al., "Blocking management of multi-time power demand response," *Power System Technology*, vol. 34, no. 9, pp. 139–143, 2010.

[20] S. Zhang, W. Qin, Z. Yao et al., "Blocking management of power bilateral transactions based on interruptible load auction model," *Power System Protection and Control*, vol. 39, no. 22, pp. 122–128, 2011.

[21] Q. Shi and L. Shidong, "Interruptible loads participate in blocking management model and characteristics analysis," *Proceedings of the CSUEPSA*, vol. 27, no. 7, pp. 48–53, 2015.

[22] P. Yan, Z. Binqiao, and X. Deng, "Study on transmission congestion management model in power market," *Telecom Power Technologies*, vol. 33, no. 4, pp. 36–38, 2016.

[23] L. Caihua, D. Songhuai, C. Zhiwen et al., "Optimal scheduling of interruptible load for blocking management under LMP mechanism," *Power System Protection and Control*, vol. 36, no. 4, pp. 67–71, 2008.

[24] X. Kong, Y. Qun, M. Yunfei et al., "Analysis method of user load demand response under time-of-use electricity price," *Proceedings of the CSUEPSA*, vol. 27, no. 10, pp. 75–80, 2015.

[25] X. Jia, H. Zhong, Q. Xia et al., "Transmission congestion management method based on incentive coupon mechanism," *Power System Technology*, vol. 37, no. 5, pp. 1291–1298, 2013.

[26] X. Shao, Y. Zhou, and G. Cui, "Research on optimal control method of customer-side controllable load energy consumption," in *Proceedings of the 2020 Chinese Automation Congress, CAC 2020*, pp. 4365–4370, Shanghai, China, November 2020.

Hindawi
Complexity
Volume 2021, Article ID 7010413, 8 pages
https://doi.org/10.1155/2021/7010413

Research Article

The Extension of the GVW Algorithm to Valuation Domains

Dongmei Li and Licui Zheng

School of Mathematics and Computations, Hunan University of Science and Technology, Xiangtan 411201, Hunan, China

Correspondence should be addressed to Licui Zheng; zhenglicui@126.com

Received 6 April 2021; Accepted 24 May 2021; Published 7 June 2021

Academic Editor: Hou-Sheng Su

The GVW algorithm is an effective algorithm to compute Gröbner bases for polynomial ideals over a field. Combined with properties of valuation domains and the idea of the GVW algorithm, we propose a new algorithm to compute Gröbner bases for polynomial ideals over valuation domains in this study. Furthermore, we use an example to demonstrate the improvement of our algorithm.

1. Introduction

The notion of Gröbner basis was first put forward by Buchberger [1]. The theory of Gröbner has been widely applied in numerous fields such as engineering, signal processing, neuroscience, coding theory, complexity, and control of networked dynamical systems and so on. For example, in the theory of symbolic dynamic systems, the problems of determining whether there is a shift equivalence of lag from one nonnegative matrix to another can be transferred into solving large-scale equations, while the latter can be solved by the Gröbner basis theory [2–13].

Seeking more efficient algorithms for the computation of Gröbner bases is a problem in which many researchers cared about extremely [14–19]. Faugère [19] proposed a fast algorithm called F5 for computing Gröbner bases. In this algorithm, he introduced two notions of rewriting and signatures, which allow them to filter the useless S-polynomials in a rather convenient way. A new algorithm named G2V [20] for computing Gröbner bases is presented by Gao et al., which is an algorithm of incremental signature and based on a simple theory. A few months later, they gave an extend version named the GVW algorithm [21]. We are particularly interested in the GVW algorithm which not only matches the original algorithm given by Buchberger in simplicity but also more effective than F5 under some term orders. The algorithms mentioned above are applied to polynomial ideals over fields.

Several algorithms have been widely investigated for Gröbner bases to rings, such as Euclidean domain, principle ideal domain, and valuation rings that may contain zero divisors [22–24].

In this study, we aim to extend the GVW algorithm to valuation domains and present a signature-based algorithm to compute Gröbner bases for ideals in $R = V[x_1, \ldots, x_n]$, where V is a valuation domain. In this algorithm, we study relations between J pairs and propose a new concept named factor, which allows us to filter the useless J pairs in a rather convenient way.

The structure of the study is arranged as follows: some basic concepts of the Gröbner basis theory is given in Section 2. In Section 3, we propose theory for the GVW algorithm over valuation domains and obtain main results of this study. Then, we present the new algorithm and demonstrate the improvement clearly by an example in Section 4.

2. Preliminaries

Let V be a valuation ring. For any two nonzero elements $a, b \in V$, there always exists $a|b$ or $b|a$. The term order for monomials in $R = V[x_1, \ldots, x_n]$ is arbitrary throughout this section. The form of the monomial in R is

$$x^{\alpha} = x_1^{a_1}, x_2^{a_2}, \ldots, x_n^{a_n}, \tag{1}$$

where $\alpha = (a_1, a_2, \ldots, a_n) \in \mathbb{N}^n$. The definition of the leading monomial (abbreviated as lm), the leading term (abbreviated as lt), and the leading coefficient (abbreviated as lc) of a given polynomial is as usual.

A nonzero polynomial set G_w in an ideal I is named as (weak) Gröbner basis for I if

$$\langle \mathrm{lt}(G_w) \rangle = \mathrm{lt}(I). \tag{2}$$

This does not imply that, for each $f \in I$, there exists some polynomial $g \in G_w$, so that $\mathrm{lt}(g) | \mathrm{lt}(f)$. For example, $I = \langle 2x, 3y \rangle \subset \mathbb{Z}[x, y]$ has a Gröbner basis $G = \{2x, 3y\}$, and $xy = -y(2x) + x(3y) \in I$, but xy is not divisible by $2x$ nor $3y$ in $\mathbb{Z}[x, y]$.

Definition 1. A set G_s is called a strong Gröbner basis for I if $\forall f \in I$, and there is a polynomial $g \in G_s$, so that $\mathrm{lt}(g) | \mathrm{lt}(f)$, where G_s is formed by the nonzero polynomials from I.

For the above example, $\{2x, 3y, xy\}$ is a strong Gröbner basis of the ideal I, but the set $\{2x, 3y\}$ is not. This shows already some difference when dealing with polynomials over rings from those over fields.

Proposition 1. *Suppose V is a valuation domain. Then, every ideal in $V[x_1, \ldots, x_n]$ has a strong Gröbner basis.*

Proof. It can be easily obtained by properties of valuation ring.

We now follow the notations in [21]. Let $R = V[x_1, \ldots, x_n]$, where V denotes a valuation domain, and $h_1, \ldots, h_m \in R$ are the polynomials. Let

$$\begin{aligned} I &= \langle h_1, \ldots, h_m \rangle \\ &= \{ p_1 h_1 + \cdots + p_m h_m \colon (p_1, \ldots, p_m) \in R^m \} \subseteq R \end{aligned} \tag{3}$$

be an ideal, and its Gröbner basis is what we want to obtain. Vectors in R^m are denoted by bold letters, for example, $\mathbf{p} = (p_1, \ldots, p_m)$. Let $\mathbf{e}_j$ be the j^{th} unit vector in R^m for $1 \le j \le m$. Define an R submodule of $R^m \times R$:

$$M = \{ (\mathbf{p}, f) \in R^m \times R \colon \mathbf{p}\mathbf{h}^t = f, \text{that is, } p_1 h_1 + \cdots + p_m h_m = v \}. \tag{4}$$

Note that, as an R module, M is generated by

$$(\mathbf{e}_1, h_1), (\mathbf{e}_2, h_2), \ldots, (\mathbf{e}_m, h_m). \tag{5}$$

$\mathbf{H} = \{(p_1, \ldots, p_m) \in R^m \colon p_1 h_1 + \cdots + p_m h_m = 0\}$ is defined to be the syzygy module of $\mathbf{h} = (h_1, \ldots, h_m)$. We shall see that the big module M allows us to get the Gröbner bases for I and $\mathbf{H}$ in the same time and allows us to develop a criterion to detect useless S polynomials.

We define quasiordering in V: for arbitrary $a, b \in V$, we say $a \prec b$ if $a | b$. Following this definition, then a term order (throughout this study, by a monomial order, we mean a global ordering [21]) $\le$ on R is defined by $a_1 t_1 \le a_2 t_2$ iff

$$\begin{aligned} &t_1 \le t_2 \\ &\text{or } t_1 = t_2, \\ &a_1 | a_2. \end{aligned} \tag{6}$$

We assume R^m has a term order that is compatible with that of R. We refer the readers to [21] for several examples on how term order of R can be extended to R^m. Note that a term in R^m is the form as

$$x^{\alpha} e_i, \tag{7}$$

for some $\alpha \in \mathbb{N}^m$ and $1 \le i \le m$. For any nonzero $\mathbf{p} \in R^m$,

$$\mathrm{lt}(\mathbf{p}) = \mathrm{lc}(\mathbf{p}) \cdot \mathrm{lm}(\mathbf{p}), \tag{8}$$

where $\mathrm{lm}(\mathbf{p})$ is the leading monomial of $\mathbf{p}$, and $\mathrm{lc}(\mathbf{p}) \in V$ is the leading coefficient of $\mathbf{p}$. □

3. Theory of the Algorithm

In this section, we present the theory of our algorithm. First, some basic definitions are needed.

Definition 2. $\mathrm{lt}(\mathbf{p})$ is said to be the signature of $(\mathbf{p}, f)$, where $(\mathbf{p}, f) \in R^m \times R$.

Definition 3. We say $(\mathbf{p}_1, f_1) \in R^m \times R$ can be top-reduced by $(\mathbf{p}_2, f_2) \in R^m \times R (\mathrm{lt}(f_2) \ne 0)$ when they meet the following two conditions:

(i) $\mathrm{lt}(f_2)$ divides $\mathrm{lt}(f_1)$ (i.e., f_2 is top-divisible by (f_1)) and

(ii) $\mathrm{lt}(t \cdot \mathbf{p}_2) \prec \mathrm{lt}(\mathbf{p}_1)$, where $t = \mathrm{lt}(f_1)/\mathrm{lt}(f_2)$

Then, the relevant top-reduction is

$$(\mathbf{p}_1, f_1) - t(\mathbf{p}_2, f_2) = (\mathbf{p}_1 - t\mathbf{p}_2, f_1 - t f_2). \tag{9}$$

Hence, we can divide this reduction into two types; one is called regular when

$$\mathrm{lm}(t\mathbf{p}_2) \prec \mathrm{lm}(\mathbf{p}_1), \tag{10}$$

and the other is called super if

$$\frac{\mathrm{lt}(\mathbf{p}_1)}{\mathrm{lt}(\mathbf{p}_2)} = \frac{\mathrm{lt}(f_1)}{\mathrm{lt}(f_2)}. \tag{11}$$

Besides, there is another super top-reduction, that is, when $f_2 = 0$, we say that $(\mathbf{p}_1, f_1)$ is super top-reduced by $(\mathbf{p}_2, 0)$ if $\mathbf{p}_1$ and $\mathbf{p}_2$ are both nonzero and $\mathrm{lt}(\mathbf{p}_2)$ divides $\mathrm{lt}(\mathbf{p}_1)$. So when a pair $(\mathbf{p}_1, f_1)$ can be reduced by $(\mathbf{p}_2, 0)$, we just reduce the signature of $(\mathbf{p}_1, f_1)$ but without increasing $\mathrm{lt}(f_1)$ (even if $f_1 = 0$). What attracts more of our attention is that $(\mathbf{p}_1, 0)$ is never top-reduced by $(\mathbf{p}_2, f_2)$ when $f_2 \ne 0$.

Similar to Definition 1, we give the definition of strong Gröbner basis for M in the following.

Definition 4. Suppose G_s is a subset of M, G_s is said to be a strong Gröbner basis of M, if every pair $(\mathbf{p}, f) \in M$ can be top-reduced by at least one pair in G_s.

It is easy to draw a conclusion from this definition, that is, every pair in M can be top-reduced to 0 by its strong Gröbner basis.

Lemma 1. *If G_s is a strong Gröbner basis of M, where $G_s = \{(\mathbf{p}_1, f_1), \ldots, (\mathbf{p}_k, f_k)\}$, then*

(1) A strong Gröbner basis for the syzygy module of $\mathbf{h} = (h_1, \ldots, h_m)$ exists, which is $G_0 = \{\mathbf{p}_i : f_i = 0, 1 \le i \le k\}$

(2) The strong Gröbner basis for $I = \langle h_1, \ldots, h_m \rangle$ also exists, which is $G_1 = \{f_i : 1 \le i \le k\}$

Proof. Assume $\mathbf{p} = (p_1, \ldots, p_m)$ is an element from the syzygy module of $\mathbf{h}$, then $(\mathbf{p}, 0) \in M$, and there must exist some pair $(\mathbf{p}_i, f_i)$ in G_s that can top-reduce $(\mathbf{p}, 0)$ with $f_i = 0$. Thus, $\mathbf{p}_i \in G_0$ and $\mathrm{lt}(\mathbf{p})$ can be reduced by $\mathrm{lt}(\mathbf{p}_i)$. This tell us that G_0 is the set we need, which is a Gröbner basis for the syzygy module of $\mathbf{h}$.

For an arbitrary nonzero polynomial $v \in I$, there is $\mathbf{p} = (p_1, \ldots, p_m) \in R^m$, such that $\mathbf{p}\mathbf{h}^t = f$ according to the definition of R^m and then $(\mathbf{p}, f) \in M$. Among all such $\mathbf{p}$, we choose the minimum $\mathrm{lt}(\mathbf{p})$. By our assumption, there exists at least one pair in G which can top-reduce $(\mathbf{p}, f)$.

If $f = 0$, then $(\mathbf{p}, f)$ can be reduced by $(\mathbf{p}_i, 0)$ and get (p', f), but $p'\mathbf{h}^t = f$, which contradicts the minimality of $\mathrm{lt}(\mathbf{p})$ as $\mathrm{lt}(p') < \mathrm{lt}(\mathbf{p})$. So there exists $(\mathbf{p}_i, f_i) \in G_1$ with $f_i \neq 0$ and $\mathrm{lt}(f_i) | \mathrm{lt}(f)$. Hence, G_1 is a Gröbner basis of I. □

Definition 5. For any pair, $P_1 = (\mathbf{p}_1, f_1), P_2 = (\mathbf{p}_2, f_2) \in R^m \times R$, and $f_1 \neq 0, f_2 \neq 0$. Let

$$\begin{aligned} t &= \mathrm{lcm}(\mathrm{lm}(f_1), \mathrm{lm}(f_2)), \\ t_1 &= \frac{t}{\mathrm{lm}(f_1)}, \\ t_2 &= \frac{t}{\mathrm{lm}(f_2)}. \end{aligned} \tag{12}$$

(i) If $\max(\mathrm{lm}(t_1\mathbf{p}_1), \mathrm{lm}(t_2\mathbf{p}_2)) = t_1\mathrm{lm}(\mathbf{p}_1)$, define

$$J(P_1, P_2) = \begin{cases} t_1 P_1, & \text{if } \mathrm{lc}(f_2) | \mathrm{lc}(f_1), \\ \dfrac{\mathrm{lc}(f_2)}{\mathrm{lc}(f_1)} t_1 P_1, & \text{if } \mathrm{lc}(f_1) | \mathrm{lc}(f_2). \end{cases} \tag{13}$$

(ii) If $\max(\mathrm{lm}(t_1\mathbf{p}_1), \mathrm{lm}(t_2\mathbf{p}_2)) = t_2\mathrm{lm}(\mathbf{p}_2)$, define

$$J(P_1, P_2) = \begin{cases} \dfrac{\mathrm{lc}(f_1)}{\mathrm{lc}(f_2)} t_2 P_2, & \text{if } \mathrm{lc}(f_2) | \mathrm{lc}(f_1), \\ t_2 P_2, & \text{if } \mathrm{lc}(f_1) | \mathrm{lc}(f_2). \end{cases} \tag{14}$$

Remark 1. With notations as above, we do not define J pair for $P_1 = (\mathbf{p}_1, f_1)$ and $P_2 = (\mathbf{p}_2, f_2)$ when one of f_1 and f_2 is zero nor when $\mathrm{lm}(t_1\mathbf{p}_1) = \mathrm{lm}(t_2\mathbf{p}_2)$.

In order to study the relation between J pairs, we propose the following conception:

Definition 6. Suppose $t_1(\mathbf{p}_1, f_1), t_2(\mathbf{p}_2, f_2)$ are the J pairs formed from G which is a (finite) subset of M, and $t_1(\mathbf{p}_1, f_1)$ is called a factor of $t_2(\mathbf{p}_2, f_2)$ if

$$\begin{aligned} at_1\mathrm{lt}(\mathbf{p}_1)w &= t_2\mathrm{lt}(\mathbf{p}_2), \\ at_1\mathrm{lt}(f_1)w &\prec t_2\mathrm{lt}(f_2), \end{aligned} \tag{15}$$

for some monomial $aw, a \in V$.

The next result is very useful for reduction.

Lemma 2. *Assume $at(\mathbf{p}_j, f_j)$ can be regular top-reduced by $(\mathbf{p}_i, f_i)$ with neither of f_i, f_j is zero, where a is an element from V and t is a monomial from R; then, $c_1 t_1(\mathbf{p}_j, f_j)$ is the J pair of $(\mathbf{p}_i, f_i)$ and $(\mathbf{p}_j, f_j)$, where*

$$\begin{aligned} c_1 &= \frac{\mathrm{lcm}(\mathrm{lc}(f_j), \mathrm{lc}(f_i))}{\mathrm{lc}(f_j)}, \\ t_1 &= \frac{\mathrm{lcm}(\mathrm{lm}(f_j), \mathrm{lm}(f_i))}{\mathrm{lm}(f_j)}, \end{aligned} \tag{16}$$

where c_1 is a divisor of a, $t_1 | t$, and $(\mathbf{p}_i, f_i)$ can regular top-reduce $c_1 t_1(\mathbf{p}_j, f_j)$.

Proof. By our assumption, there exist $a_1 \in V$ and a monomial say $w_1 \in R$, so that

$$\begin{aligned} at\mathrm{lt}(f_j) &= a_1 w_1 \mathrm{lt}(f_i), \\ w_1\mathrm{lm}(\mathbf{p}_i) &\prec t\mathrm{lm}(\mathbf{p}_j). \end{aligned} \tag{17}$$

Set

$$\begin{aligned} c_2 &= \frac{\mathrm{lcm}(\mathrm{lc}(f_j), \mathrm{lc}(f_i))}{\mathrm{lc}(f_i)}, \\ t_2 &= \frac{\mathrm{lcm}(\mathrm{lm}(f_j), \mathrm{lm}(f_i))}{\mathrm{lm}(f_i)}, \end{aligned} \tag{18}$$

we have that

$$\begin{aligned} c_2 t_2 &= \frac{\mathrm{lcm}(\mathrm{lc}(f_j), \mathrm{lc}(f_i))}{\mathrm{lc}(f_i)} \cdot \frac{\mathrm{lcm}(\mathrm{lm}(f_j), \mathrm{lm}(f_i))}{\mathrm{lm}(f_i)} \\ &= \frac{\mathrm{lcm}(\mathrm{lt}(f_j), \mathrm{lt}(f_i))}{\mathrm{lt}(f_i)}. \end{aligned} \tag{19}$$

Then, some $a_2 \in V$ and monomial $w_2 \in R$ exist, such that

$$a_2w_2\mathrm{lcm}\left(\mathrm{lt}\left(f_j\right),\mathrm{lt}\left(f_i\right)\right)=at\mathrm{lt}\left(f_j\right)=a_1w_1\mathrm{lt}\left(f_i\right).$$

$$\begin{aligned}c_1t_1&=\frac{\mathrm{lcm}\left(\mathrm{lc}\left(f_j\right),\mathrm{lc}\left(f_i\right)\right)}{\mathrm{lc}\left(f_j\right)}\cdot\frac{lcm\left(\mathrm{lm}\left(f_j\right),\mathrm{lm}\left(f_i\right)\right)}{\mathrm{lm}\left(f_j\right)}\\&=\frac{lcm\left(\mathrm{lt}\left(f_j\right),\mathrm{lt}\left(f_i\right)\right)}{\mathrm{lt}\left(f_j\right)}.\end{aligned}\tag{20}$$

Thus,

$$\begin{aligned}a_1w_1\mathrm{lt}\left(f_i\right)&=a_2w_2c_2t_2\mathrm{lt}\left(f_i\right),\\at\mathrm{lt}\left(f_j\right)&=a_2w_2c_1t_1\mathrm{lt}\left(f_j\right),\\a\mathrm{lc}\left(f_j\right)&=a_2c_1\mathrm{lc}\left(f_j\right),\\a_1\mathrm{lc}\left(f_i\right)&=a_2c_2\mathrm{lc}\left(f_i\right),\\t\mathrm{lm}\left(f_j\right)&=t_1w_2\mathrm{lm}\left(f_j\right),\\w_1\mathrm{lm}\left(f_i\right)&=t_2w_2\mathrm{lm}\left(f_i\right).\end{aligned}\tag{21}$$

Then, we have that

$$\begin{aligned}a&=a_2c_1,\\a_1&=a_2c_2,\\c_1&|a,\\t&=t_1w_2,\\w_1&=t_2w_2.\end{aligned}\tag{22}$$

Hence, $c_1t_2\mathrm{lt}(\mathbf{p}_i)\prec c_2t_1\mathrm{lt}(\mathbf{p}_j)$ as $s\mathrm{lm}(\mathbf{p}_i)\prec t\mathrm{lm}(\mathbf{p}_j)$. Thus, $\max(t_2\mathrm{lm}(\mathbf{p}_i),t_1\mathrm{lm}(\mathbf{p}_j))=t_1\mathrm{lm}(\mathbf{p}_j)$. And the J pair of $(\mathbf{p}_j,f_j)$ and $(\mathbf{p}_i,f_i)$ is $c_1t_1(\mathbf{p}_j,f_j)$ by Definition 5. Note that

$$at\mathrm{lt}\left(f_1\right)=a_1w_1\mathrm{lt}\left(f_2\right),\tag{23}$$

and then,

$$a_2c_1t_1w_2\mathrm{lt}\left(f_j\right)=a_2c_2t_2w_2\mathrm{lt}\left(f_i\right).\tag{24}$$

Therefore,

$$\mathrm{lt}\left(f_i\right)|c_1t_1\mathrm{lt}\left(f_j\right).\tag{25}$$

We see that $(\mathbf{p}_i,f_i)$ can regular top-reduce $c_1t_1(\mathbf{p}_j,f_j)$.

Suppose S is a set formed by the pairs in $R^m\times R$, we say S can regular (super) top-reduce the pair $(\mathbf{p},f)\in R^m\times R$, if there is at least one pair in S which can regular (super) top-reduce the pair $(\mathbf{p},f)$. Furthermore, we implement a series of such reductions to $(\mathbf{p},f)$ until it cannot be regular top-reduced by this set anymore, but $(\mathbf{p},f)$ can be super top-reduced by S, and this reduction defined as $(\mathbf{p},f)$ is eventually super top-reduced by S. □

Theorem 1. *Suppose T is any term in R^m, and there always exists a pair $(\mathbf{p}_i,f_i)$ in G_s, a monomial $t\in R$ and $a\in V$, such that $T=at\mathrm{lt}(\mathbf{p}_i)$, where G_s is a subset of M. Then, G_s is a strong Gröbner basis for M if and only if, for the J pairs formed from G_s such as $(\mathbf{p},f)$, there always exists a pair $(\mathbf{p}_1,f_1)\in G_s$, so that $lt(\mathbf{p}_1)|lt(\mathbf{p})$ and $t\mathrm{lm}(f_1)\prec lm(f)$, where $t=lm(\mathbf{p})/lm(\mathbf{p}_1)$.*

Proof. Necessity: let $(\mathbf{p},f)$ be an arbitrary J pair formed from G_s; then, $(\mathbf{p},f)$ is in M, and it is top-reduced by G_s as the set G_s is a strong Gröbner basis for the module M. We can do the regular top-reductions to $(\mathbf{p},f)$ as much as possible until $(\mathbf{p},f)$ cannot be regular top-reduced any more, say to get $(\mathbf{p}_0,f_0)$. And $(\mathbf{p}_0,f_0)$ can be top-reduced by G_s as it is still in M; but now, the reduction can only be super reduction, say $(\mathbf{p}_0,f_0)$ can be super reduced by $(\mathbf{p}_1,f_1)\in G_s$.

(1) If $f_1=0$, then $\mathrm{lt}(\mathbf{p}_1)|\mathrm{lt}(\mathbf{p}_0)=\mathrm{lt}(\mathbf{p})$, and $tf_1=0$ is smaller than $\mathrm{lm}(f)$, and the conclusion is true.

(2) If $f_1\neq 0$,

$$\begin{aligned}t&=\frac{\mathrm{lm}\left(f_0\right)}{\mathrm{lm}\left(f_1\right)}=\frac{\mathrm{lm}\left(\mathbf{p}_0\right)}{\mathrm{lm}\left(\mathbf{p}_1\right)},\\&\frac{\mathrm{lc}\left(f_0\right)}{\mathrm{lc}\left(f_1\right)}\cdot\mathrm{lc}\left(\mathbf{p}_1\right)\prec\mathrm{lc}\left(\mathbf{p}_0\right).\end{aligned}\tag{26}$$

Combined with the definition of quasiordering given, we have that

$$\frac{\mathrm{lc}\left(f_0\right)}{\mathrm{lc}\left(f_1\right)}\cdot\mathrm{lc}\left(\mathbf{p}_1\right)|\mathrm{lc}\left(\mathbf{p}_0\right),\tag{27}$$

and then, $\mathrm{lc}(\mathbf{p}_1)|\mathrm{lc}(\mathbf{p}_0)$.

Since $(\mathbf{p}_0,f_0)$ is obtained by performed regular top-reduction to $(\mathbf{p},f)$, then $\mathrm{lm}(f_0)\prec\mathrm{lm}(f)$ and $\mathrm{lt}(\mathbf{p}_0)=\mathrm{lt}(\mathbf{p})$; the latter shows that

$$t\mathrm{lm}\left(f_1\right)=\mathrm{lm}\left(f_0\right)\prec\mathrm{lm}\left(f\right).\tag{28}$$

So, $\mathrm{lt}(\mathbf{p}_1)|\mathrm{lt}(\mathbf{p}_0)=\mathrm{lt}(\mathbf{p})$ and $t\mathrm{lm}(f_1)\prec\mathrm{lm}(f)$.

Sufficiency: suppose there are some pairs in M which cannot be top-reduced the set G_s, say $(\mathbf{p},f)\in M$ is such a pair. We prove that such a pair does not exist. Select the minimal signature $T=\mathrm{lt}(\mathbf{p})$ from all such pairs with $\mathrm{lt}(\mathbf{p})\neq\mathbf{0}$, and we choose a pair $(\mathbf{p}_j,f_j)$ from G_s, so that

(i) $T=at\mathrm{lt}(\mathbf{p}_j)$ with $a\in V$ and t is a monomial in R, and

(ii) $t\mathrm{lm}(f_j)$ is the smallest one among all $1\le j\le k$ satisfying (i)

In the following, we prove that $at(\mathbf{p}_j,f_j)$ should not be regular top-reduced by G_s. Suppose that $at(\mathbf{p}_j,f_j)$ could be regular top-reduced by a pair $(\mathbf{p}_i,f_i)$, $i\neq j$, then $f_j\neq 0, f_i\neq 0$. What we expect is to get a conclusion that contradicts condition (ii). From Lemma 2, we have that $c_1t_1(\mathbf{p}_j,f_j)$ is the J pair of $(\mathbf{p}_j,f_j)$ and $(\mathbf{p}_i,f_i)$; moreover, it can still be regular top-reduced by $(\mathbf{p}_i,f_i)$, where

$$c_1 = \frac{\mathrm{lcm}\left(\mathrm{lc}\left(f_i\right), \mathrm{lc}\left(f_j\right)\right)}{\mathrm{lc}\left(f_j\right)},$$
$$t_1 = \frac{\mathrm{lcm}\left(\mathrm{lm}\left(f_j\right), \mathrm{lm}\left(f_i\right)\right)}{\mathrm{lm}\left(f_j\right)}, \tag{29}$$
$$t = t_1 w, c_1 | a,$$

for some $c_1 \in V$ and monomial $w \in R$. □

Case 1. If $\mathrm{lc}(f_i) | \mathrm{lc}(f_j)$, then $c_1 = 1$, and the J pair of $(\mathbf{p}_i, f_i)$ and $(\mathbf{p}_j, f_j)$ is $t_1(\mathbf{p}_j, f_j)$. Note that $t_1(\mathbf{p}_j, f_j)$ is a J pair formed from G_s, so there must exists a pair $(\mathbf{p}_\ell, f_\ell) \in G_s$, so that $\mathrm{lt}(\mathbf{p}_\ell) | t_1 \mathrm{lt}(\mathbf{p}_j)$ and $t_3 \mathrm{lm}(f_\ell) \prec t_1 \mathrm{lm}(f_j)$ with $t_3 = t_1 \mathrm{lm}(\mathbf{p}_j)/\mathrm{lm}(\mathbf{p}_\ell)$. Since $\mathrm{lc}(\mathbf{p}_\ell) | \mathrm{lc}(\mathbf{p}_j)$, we set $a_3 = (\mathrm{lc}(\mathbf{p}_j)/\mathrm{lc}(\mathbf{p}_\ell))$, and then,

$$\begin{aligned} a \cdot a_3 t_3 w \mathrm{lt}(\mathbf{p}_\ell) = a \cdot t_1 w \mathrm{lc}\left(\mathbf{p}_j\right) \cdot \mathrm{lm}\left(\mathbf{p}_j\right) = a t_1 w \mathrm{lt}\left(\mathbf{p}_j\right) = a t \mathrm{lt}\left(\mathbf{p}_j\right) = T, \\ t_3 w \mathrm{lm}(f_\ell) \prec w t_1 \mathrm{lm}\left(f_j\right) = t \mathrm{lm}\left(f_j\right). \end{aligned} \tag{30}$$

This contradicts the condition (ii) for the selection of $(\mathbf{p}_j, f_j)$ in G_s.

Case 2. If $\mathrm{lc}(f_j) | \mathrm{lc}(f_i)$, then $c_1 = (\mathrm{lc}(f_i)/\mathrm{lc}(f_j))$, and the J pair of $(\mathbf{p}_i, f_i)$ and $(\mathbf{p}_j, f_j)$ is $(\mathrm{lc}(f_j)/\mathrm{lc}(f_i))t_1(\mathbf{p}_j, f_j)$. Note that $(\mathrm{lc}(f_i)/\mathrm{lc}(f_j))t_1(\mathbf{p}_j, f_j)$ is a J pair formed from G_s, and there must exist some pair say $(\mathbf{p}_\ell, f_\ell) \in G_s$, such that $\mathrm{lt}(\mathbf{p}_\ell) | (\mathrm{lc}(f_i)/\mathrm{lc}(f_j)) t_1 \mathrm{lt}(\mathbf{p}_j)$ and $t_3 \mathrm{lm}(f_\ell) \prec t_1 \mathrm{lm}(f_j)$ with $t_3 = t_1 \mathrm{lm}(\mathbf{p}_j)/\mathrm{lm}(\mathbf{p}_\ell)$. Since $\mathrm{lc}(\mathbf{p}_\ell) | (\mathrm{lc}(f_i)/\mathrm{lc}(f_j)) \mathrm{lc}(\mathbf{p}_j)$, set $a_3 = (\mathrm{lc}(f_i)/\mathrm{lc}(f_j)) \cdot (\mathrm{lc}(\mathbf{p}_j)/\mathrm{lc}(\mathbf{p}_\ell))$, and then,

$$\begin{aligned} \frac{a}{\left(\mathrm{lc}(f_i)/\mathrm{lc}\left(f_j\right)\right)} a_3 t_3 w \mathrm{lt}(\mathbf{p}_\ell) = \frac{a}{\left(\mathrm{lc}(f_i)/\mathrm{lc}\left(f_j\right)\right)} \frac{\mathrm{lc}(f_i)}{\mathrm{lc}\left(f_j\right)} t_1 w \mathrm{lc}\left(\mathbf{p}_j\right) \cdot \mathrm{lm}\left(\mathbf{p}_j\right) = a t_1 w \mathrm{lt}\left(\mathbf{p}_j\right) = a t \mathrm{lt}\left(\mathbf{p}_j\right) = T, \\ t_3 w \mathrm{lm}(f_\ell) \prec w t_1 \mathrm{lm}\left(f_j\right) = t \mathrm{lm}\left(f_j\right). \end{aligned} \tag{31}$$

This contradicts the condition (ii) for the selection of $(\mathbf{p}_j, f_j)$ in G_s.

Let

$$(\tilde{\mathbf{p}}, \tilde{f}) = (\mathbf{p}, f) - a t \left(\mathbf{p}_j, f_j\right), \tag{32}$$

and then, $\mathrm{lm}(\tilde{\mathbf{p}}) \prec \mathrm{lm}(\mathbf{p})$ and $\mathrm{lt}(f) \neq a t \mathrm{lt}(f_j)$. Otherwise, $(\mathbf{p}, f)$ would be top-reduced by $(\mathbf{p}_j, f_j)$ which contradicts with the selection of $(\mathbf{p}, f)$. So, $\tilde{f} \neq 0$ and $(\tilde{\mathbf{p}}, \tilde{f})$ can be top-reduced by G_s as $(\tilde{\mathbf{p}}, \tilde{f}) \in M$ and $\mathrm{lm}(\tilde{\mathbf{p}}) \prec \mathrm{lm}(\mathbf{p})$, $\mathrm{lt}(\tilde{\mathbf{p}}) \prec \mathrm{lt}(\mathbf{p})$, say it is top-reduced by $(\mathbf{p}_k, f_k)$. If $f_k = 0$, we use this type of pairs to reduce $(\tilde{\mathbf{p}}, \tilde{f})$ as much as possible and obtain a new pair $(p', \tilde{f})$ finally, which cannot be top-reduced by the same type of pairs (here, it refers the pairs whose f-part is zero) in G_s. Since $(p', \tilde{f}) \in M$ and $\mathrm{lm}(p') \prec \mathrm{lm}(\mathbf{p})$, $\mathrm{lt}(p') \prec \mathrm{lt}(\mathbf{p})$, and then $(p', \tilde{f})$ can be top-reduced by G_s, say by $(\mathbf{p}_k', f_k')$, $f_k' \neq 0$. For $\mathrm{lt}(f) \neq a t \mathrm{lt}(f_j)$, there are the following three cases that need to be considered:

(i) If $\mathrm{lm}(f) \prec t \mathrm{lm}(f_j)$, then $\mathrm{lt}(\tilde{f}) = a t \mathrm{lt}(f_j)$; but $\mathrm{lm}(p') \prec t \mathrm{lm}(\mathbf{p}_j)$, so there must exist some pairs in G_s that can regular top-reduce at $(\mathbf{p}_j, f_j)$; assume the pair in G_s is $(\mathbf{p}_k', f_k')$. This is impossible as $a t(\mathbf{p}_j, f_j)$ cannot be regular top-reduced by any pair in G_s.

(ii) If $\mathrm{lm}(f) \succ t \mathrm{lm}(f_j)$, then $\mathrm{lm}(\tilde{f}) = \mathrm{lm}(f)$, and $(\mathbf{p}, f)$ can be regular top-reduced by $(\mathbf{p}_k, f_k)$; this contradicts the choice of $(\mathbf{p}, f)$.

(iii) If $\mathrm{lm}(f) = t \mathrm{lm}(f_j)$ and $\mathrm{lc}(f) \neq a \mathrm{lc}(f_j)$, $(\tilde{\mathbf{p}}, \tilde{f})$ is top-reduced by $(\mathbf{p}_k', f_k')$, and then, $\mathrm{lt}(f_k') | \mathrm{lt}(\tilde{f})$; this means that $\mathrm{lm}(f_l') | \mathrm{lm}(\tilde{f})$ and $\mathrm{lc}(f_k') | \mathrm{lc}(\tilde{f}) = \mathrm{lc}(f) - a \mathrm{lc}(f_j)$.

From the property of valuation ring, we consider the relation between $a \mathrm{lc}(f_j)$ and $\mathrm{lc}(f)$ in the following three cases:

(a) If $a \mathrm{lc}(f_j) = b \mathrm{lc}(f)$ and b is not a nonzero unit, note that a valuation ring is a local ring, and b is in the unique maximal ideal. Since $\mathrm{lc}(f_k') | \mathrm{lc}(f) - a \mathrm{lc}(f_j) = (1 - b) \mathrm{lc}(f)$ and $1 - b$ is invertible, then $\mathrm{lc}(f_k') | \mathrm{lc}(f)$. Note that $(\mathrm{lm}(f)/\mathrm{lm}(f_l')) \mathrm{lm}(\mathbf{p}_k') \prec \mathrm{lm}(p') \prec \mathrm{lm}(\mathbf{p})$ and $(\mathbf{p}, f)$ are regular top-reduced by $(\mathbf{p}_k, f_k)$; this contradicts the choice of $(\mathbf{p}, f)$.

(b) If $a \mathrm{lc}(f_j) \cdot b = \mathrm{lc}(f)$ and b is not a nonzero unit, note that a valuation ring is a local ring, and b is in the unique maximal ideal. Since $\mathrm{lc}(f_k') | \mathrm{lc}(f) - a \mathrm{lc}(f_j) = (1 - b) a \mathrm{lc}(f_j)$ and $1 - b$ is invertible, then $\mathrm{lc}(f_k') | a \mathrm{lc}(f_j)$. Note that $(t \mathrm{lm}(f_j)/\mathrm{lm}(f_\ell')) \mathrm{lm}(\mathbf{p}_k') \prec \mathrm{lm}(p') \prec \mathrm{lm}(\mathbf{p}) = t \mathrm{lm}(\mathbf{p}_j)$, then $a t(\mathbf{p}_j, f_j)$ is regular top-reduced by $(\mathbf{p}_k, f_k)$, and this case is impossible.

(c) If $a \mathrm{lc}(f_j) = b \mathrm{lc}(f)$ and b is a nonzero unit, then $a \mathrm{lc}(f_j) | \mathrm{lc}(f)$ and $(\mathrm{lc}(f)/a \mathrm{lc}(f_j)) = b^{-1}$. Note that $\mathrm{lt}(\mathbf{p}) = a t \mathrm{lt}(\mathbf{p}_j)$; combining the definition of the

order, we have $b^{-1} \cdot a\mathrm{lt}(\mathbf{p}_j) \prec \mathrm{lt}(\mathbf{p})$. Therefore, $(\mathbf{p}, f)$ can be top-reduced by $at(\mathbf{p}_j, f_j)$, and this contradicts the choice of $(\mathbf{p}, f)$.

Thus, such pairs like $(\mathbf{p}, f)$ cannot exist in M at all; hence, all pairs of M can be top-reduced by G_s, and G_s is a strong Gröbner basis for M.

The pair $p = (\mathbf{p}, f)$ is covered by G_s when there exists at least one pair such as $p_1 = (\mathbf{p}_1, f_1)$ in G_s, such that $\mathrm{lt}(\mathbf{p}_1) | \mathrm{lt}(\mathbf{p})$ and $t\mathrm{lm}(f_1) < \mathrm{lm}(f)$, where $t = \mathrm{lm}(\mathbf{p})/\mathrm{lm}(\mathbf{p}_1)$.

Theorem 2. *Suppose G_s is a special subset of M, whose particularity is reflected in for any term $T \in R^m$, there always exists a pair $(\mathbf{p}_i, f_i) \in G_s$ and monomial $t \in R$ and $a \in V$, such that $T = at\mathrm{lt}(\mathbf{p}_i)$. Then, G_s is a strong Gröbner basis for M if the factors of the J pairs formed by G_s can always be eventually super top-reduced by the set G_s.*

Proof. Assume that $t_1(\mathbf{p}_1, f_1), t_2(\mathbf{p}_2, f_2)$ are the two J pairs formed by G, and $t_1(\mathbf{p}_1, f_1)$ is a factor of $t_2(\mathbf{p}_2, f_2)$. Then, $t_1(\mathbf{p}_1, f_1)$ can be eventually super top-reducible by G_s, that is, after doing a series of regular top-reduction to $t_1(\mathbf{p}_1, f_1)$, say it to get (p', f') where $\mathrm{lt}(p') = t_1\ \mathrm{lm}(\mathbf{p}_1), \mathrm{lt}(f') < t_1\mathrm{lt}(f_1)$; besides, (p', f') can be super top-reduced by G_s, say $(\mathbf{p}_i, f_i)$; then, $a_i\mathrm{lt}(\mathbf{p}_i) = t_1\mathrm{lt}\ (\mathbf{p}_1), a_i t_i \mathrm{lt}(f_i) < t_1\mathrm{lt}(f_1)$, where $a_i t_i = (\mathrm{lt}(f')/\mathrm{lt}(f_i))$. Clearly, $t_2(\mathbf{p}_2, f_2)$ can be covered by $(\mathbf{p}_i, f_i)$. It is also correct for the rest of J pairs and their factors. By Theorem 1, we have that G_s is a strong Gröbner basis for M.

According to Theorems 1 and 2, we can discard the J pairs which can be covered by G_s without doing any regular nor super top-reductions. As a consequence, there are four criteria for discarding redundant J pairs. □

Corollary 1. *(Covered criterion) For any J pair $(\mathbf{p}, f)$ of G_s, it can be discard if $(\mathbf{p}, f)$ is covered by G_s.*

Corollary 2. *(Syzygy criterion) If a J pair $(\mathbf{p}, f)$ can be top-reduced by a syzygy, then it can be discarded.*

Corollary 3. *(Signature criterion) As for the J pairs with the same signature, we only need to keep the one whose f-part is minimal.*

Corollary 4. *(Factor criterion) As for the J pair which has a factor, we just need to keep the factor.*

4. Algorithm and Example

According to the theorems and corollaries in Section 3, we can get an algorithm for computing Gröbner bases for the polynomial ideals over valuation domains. We call the algorithm as VID − GVW. The main idea of VID − GVW is analogue to the GVW algorithm of principal ideal domain [23]. First, we form J pairs by the initial pairs $(\mathbf{e}_1, g_1), \ldots, (\mathbf{e}_m, g_m)$. By Theorems 1 and 2, we just store the J pairs with different signatures. We only consider the J pairs that we store, choose any one of them, denoted as $(\mathbf{p}, f)$, and then check whether it satisfies Corollary 1, that is, whether it is covered by G_s. If so, discard it. Otherwise, delete all the J pairs whose factor is $(\mathbf{p}, f)$ and perform regular top-reductions to it repeatedly until it cannot be regular top-reduced any more, say to get (p', f') finally. If $f' = 0$, then p' is a syzygy in $\mathbf{H}$. We add it to $\mathbf{H}$ and delete the pairs whose signature is divisible by $\mathrm{lt}(p')$. Otherwise, (p', f') adds to the set JP, and we will not stop the process until the set of J pair is empty. In the while-loop, all the J pairs formed from G_s will be top-reduced by the set G_s. We describe the algorithm in more detail and accurately with Figure 1. $\mathbf{H}$ is used to store the leading terms of syzygies; the Gröbner basis we get is a list of pairs $(T_1, f_1), (T_2, f_2), \ldots, (T_k, f_k)$, where $f_j \neq 0$ for $1 \leq j \leq k$. We store this list as

$$\begin{aligned} P &= [T_1, T_2, \ldots, T_k], \\ F &= [f_1, f_2, \ldots, f_k]. \end{aligned} \tag{33}$$

So the whole list $(T_1, f_1), (T_2, f_2), \ldots, (T_k, f_k)$ is represented by $[P, F]$.

Theorem 3. *Assume the term orders in R are compatible with which in R^m; then, the algorithm shown in Figure 1 will terminate after a finite number of steps and get a strong Gröbner basis for M.*

Proof. The correctness of our algorithm is obviously according to Theorems 1 and 2. As for the termination of the algorithm, we can refer to the Theorem 1 in [21].

Next, we present our algorithm and use a concrete example to demonstrate the improvement clearly. □

Example 1. Let $V[x, y] = \mathbb{Z}_{(3^3)}[x \cdot y]$, and we consider the Gröbner basis of the ideal $I = \langle f_1, f_2, f_3 \rangle \in V[x, y]$, where

$$\begin{aligned} f_1 &= x^2 y - x, \\ f_2 &= xy^2 - xy, \\ f_3 &= y^3, \end{aligned} \tag{34}$$

and $\mathbb{Z}_{(3^3)}$ is a discrete valuation ring, for each prime p (here is 3); set v_p: $Q^* \longrightarrow \mathbb{Z}$ is a function given by $v_p(p^k a/b) = k$ if a, b are integers relatively prime to p.

The term order we set on $V[x, y]$ is the lexicographical ordering, which is defined by $y \prec x$. Besides, the term order $\prec$ on $V^m[x, y]$ is $x^\alpha \mathbf{e}_i \prec x^\beta \mathbf{e}_j$ if $i < j$ or $i = j$ and $x^\alpha \prec x^\beta$.

First, let $[P, F] = \{(\mathbf{e}_1, f_1), (\mathbf{e}_2, f_2), (\mathbf{e}_3, f_3)\}$; then, $H = \{xy^2\mathbf{e}_1, y^3\mathbf{e}_1, y^3\mathbf{e}_2\}$, which is a set that stands for leading term of principle syzygy.

Choose $(\mathbf{e}_1, f_1), (\mathbf{e}_2, f_2)$ from $[P, F]$; then, $\mathrm{lcm}(\mathrm{lm}(f_1), \mathrm{lm}(f_2)) = x^2y^2$. Hence, $t_1 = y, t_2 = x$, and J pair of $(\mathbf{e}_1, f_1)$ and $(\mathbf{e}_2, f_2)$ is $y(\mathbf{e}_1, f_1)$, whose signature cannot be reduced by H, and we store it.

Doing the same process, the J-pairs set of other pairs in $[P, F]$ is $\{y^2(\mathbf{e}_1, f_1), y(\mathbf{e}_2, f_2)\}$, but $y(\mathbf{e}_1, f_1)$ is a factor of $y^2(\mathbf{e}_1, f_1)$, so delate. Then,

$$\mathrm{JP} = \{y(\mathbf{e}_1, f_1), y(\mathbf{e}_2, f_2)\}. \tag{35}$$

Selecting $y(\mathbf{e}_2, f_2)$ from JP, we get $((y+1)\mathbf{e}_2 - x\mathbf{e}_3, -xy)$ after a series of regular top-reduction, add it to $[P, F]$, and recalculate the J pairs and

The GVW algorithm to valuation domain (VID-GVW)	
Input:	$f_1, \ldots, f_m \in R = V[x_1, \ldots, x_n]$, the term orders for R^m and R
Output:	A Gröbner basis of $I = \langle f_1, \ldots, f_m \rangle$
Variables:	P : the set of terms $T_i = \mathrm{lt}(p_i)$, denoting the signatures of $(p_i, f_i) \in M$;
	F : the set of polynomials for $f \in I$;
	H : the set for $lt(p)$ where $p \in R^m$ refers to a syzygy obtained at present,
	JP : the set of J pairs.
	Deal : the pair dealt in this step.
Initial:	$P = [e_1, \ldots, e_m]$, and $F = [f_1, \ldots, f_m]$.
	$H :=$ { the syzygies' leading terms $f_j e_i - f_j e_j$ for $1 \le i < j \le m$}★,
	JP := { the J Pairs formed by $(e_1, f_1), \ldots, (e_m, f_m)$}♠
WHILE	JP $\neq \Phi$ Do
	Deal := Take any pair (T, f_1) from JP
	JP : = JP − $\{(T, f_1)\}$ − the J pair whose factor is $\{(T, f_1)\}$.
	IF (Deal can be covered by $[P, F]$) THEN
	Next;
	ELSE
	(T, f) := the Deal's remainder obtained by a series regular top-reduced by $[P, F]$.
	IF (T, f) is super top-reduced by $[P, F]$. THEN
	Next;
	IF $f = 0$ THEN
	$H := H \cup \{T\}$
	JP := JP — the pairs whose signature can be divisible by T.
	ELSE
	$H := H \cup$ {the syzygies' leading terms of $fP_i - f_i P$ } for $1 \le i \le \lvert P \rvert$ (and delete any redundant ones).
	JP := JP ∪ { J pairs between deal and $[P, F]$}♠
	$P := P \cup \{T\}$ and $F := F \cup \{f\}$
Return:	V .
★ :	Deposit the syzygies in H with the minimal leading terms.
♠ :	Deposit the J pairs whose signatures can not be reduced by H and for the J pair with the distinct signature, we only keep the one whose f-part is minimal.

Figure 1: VID-GVW.

$$H = \left\{xy^2\mathbf{e}_1, y^3\mathbf{e}_1, y^3\mathbf{e}_2\right\}. \tag{36}$$

Using the syzygy criterion and factor criterion, we obtain

$$\mathrm{JP} = \left\{y\left(\mathbf{e}_1, f_1\right), y\left((y+1)\mathbf{e}_2 - x\mathbf{e}_3, -xy\right)\right\}. \tag{37}$$

Select $y((y+1)\mathbf{e}_2 - x\mathbf{e}_3, -xy)$ from JP; it can be regular top-reduced to $(y(y+1)\mathbf{e}_2 - x\mathbf{e}_3, 0))$ and add $y^2\mathbf{e}_2$ to H.

Select $y(\mathbf{e}_1, f_1)$ from JP; it can be regular top-reduced to $(y\mathbf{e}_1, 0))$ and add $y\mathbf{e}_1$ to H.

Now, JP = ∅. So $\{x^2y - x, xy^2 - xy, y^3, -xy\}$ is a strong Gröbner basis for I in $V[x, y]$.

5. Conclusions

In this study, we have generalized the GVW algorithm and presented an algorithm to compute Gröbner bases for polynomial ideals over valuation domains. We have also given an example to illustrate our method. All of these could provide useful information for engineers to solve large linear

systems and networked dynamical systems. Valuation domains are special kinds of valuation rings which may have zero divisors. In the future, we want to consider new algorithms for Gröbner bases of ideals over general valuation rings. And we also hope to establish a dynamical Gröbner basis algorithm which combined with the algorithm in the study.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (11871207).

References

[1] B. Buchberger, "An algorithm for finding a basis for the redidue class ring of a zero dimensiomal polynomial," Ph. D. thesis, Universität Innsbruck, Institut für Mathematik, Innsbruck, Austria, 1965.

[2] D. Li, J. Liu, and L. Zheng, "On the equivalence of multivariate polynomial matrices," *Multidimensional Systems and Signal Processing*, vol. 28, no. 1, pp. 225–235, 2017.

[3] J. Zhang and H. Su, "Formation-containment control for multi-agent systems with sampled data and time delays," *Neurocomputing*, vol. 424, pp. 125–131, 2021.

[4] Y. Liu and H. Su, "Some necessary and sufficient conditions for containment of second-order multi-agent systems with sampled position data," *Neurocomputing*, vol. 378, pp. 228–237, 2020.

[5] Y. Liu and H. Su, "Containment control of second-order multi-agent systems via intermittent sampled position data communication," *Applied Mathematics and Computation*, vol. 362, Article ID 124522, 2019.

[6] C. Xu, H. Xu, H. Su, and C. Liu, "Disturbance-observer based consensus of linear multi-agent systems with exogenous disturbance under intermittent communication," *Neurocomputing*, vol. 404, pp. 26–33, 2020.

[7] C. Xu, B. Li, and L. Yang, "Semi-global containment of discrete-time high-order multi-agent systems with input saturation via intermittent control," *IET Control Theory & Applications*, vol. 14, no. 16, pp. 2303–2309, 2020.

[8] C. Zhao, S. Zhong, X. Zhang, Q. Zhong, and K. Shi, "Novel results on nonfragile sampled-data exponential synchronization for delayed complex dynamical networks," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 10, pp. 4022–4042, 2020.

[9] C. Zhao, S. Zhong, Q. Zhong, and K. Shi, "Synchronization of Markovian complex networks with input mode delay and Markovian directed communication via distributed dynamic event-triggered control," *Nonlinear Analysis: Hybrid Systems*, vol. 36, Article ID 100883, 2020.

[10] X. Cai, J. Wang, S. Zhong, K. Shi, and Y. Tang, "Fuzzy quantized sampled-data control for extended dissipative analysis of T-S fuzzy system and its application to WPGSs," *Journal of the Franklin Institute*, vol. 358, no. 2, pp. 1350–1375, 2021.

[11] K. Shi, J. Wang, S. Zhong, Y. Tang, and J. Cheng, "Non-fragile memory filtering of T-S fuzzy delayed neural networks based on switched fuzzy sampled-data control," *Fuzzy Sets and Systems*, vol. 394, pp. 40–64, 2020.

[12] K. Shi, J. Wang, Y. Tang, and S. Zhong, "Reliable asynchronous sampled-data filtering of T-S fuzzy uncertain delayed neural networks with stochastic switched topologies," *Fuzzy Sets and Systems*, vol. 381, pp. 1–25, 2020.

[13] X. Cai, S. Zhong, J. Wang, and K. Shi, "Robust H∞ control for uncertain delayed T-S fuzzy systems with stochastic packet dropouts," *Applied Mathematics and Computation*, vol. 385, Article ID 125432, 2020.

[14] B. Buchberger, "A criterion for detecting unnecessary reductions in the construction of Gröbner bases," in *EUROSAM'79: Proceedings of the International Symposiumon on Symbolic and Algebraic Computation*, pp. 3–21, Marseille, France, June 1979.

[15] B. B.. Gröbner-Bases, *An Algorithmic Method in Polynomial Ideal Theory*, Reidel Publishing Company, Dordrecht, Netherlands, 1985.

[16] N. Courtois, A. Klimov, J. Patarin, and A. Shamir, "Efficient algorithms for solving overdefined systems of multivariate polynomial equations," *Advances in Cryptology-EUROCRYPT 2000*, Springer-Verlag, Berlin, Germany, pp. 392–407, 2000.

[17] C. Eder and J. Perry, "F5C: a variant of Faugere's F5 algorithm with reduced Gröbner bases," *Journal of Symbolic Computation*, vol. 45, no. 12, pp. 1442–1458, 2009.

[18] Y. Sun and D. Wang, "F5B: A new proof of the F5 algorithm," 2010, http://arxiv.org/abs/1004.0084.

[19] J. C. Faugere, "A new effcient algorithm for computing Gröbner bases without reduction to zero (F5)," in *ISSAC 02: Proceedings of the 2002 International Symposium on Symbolic and algebraic computation*, pp. 75–83, Lille, France, July 2002.

[20] S. Gao, Y. Guan, and F. Volny, "A new incremental algorithm for computing Gröbner bases," in *ISSAC 10: Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation*, pp. 13–19, München, Germany, July 2010.

[21] S. Gao, F. Volny, and M. Wang, "A new framework for computing Gröbner bases," *Mathematics of Computation*, vol. 85, no. 297, pp. 449–465, 2015.

[22] D. Li, J. Liu, and L. Zheng, "A zero-dimensional valuation ring is 1-Gröbner," *Journal of Algebra*, vol. 484, pp. 334–343, 2017.

[23] D. Li, J. Liu, W. Liu, and L. Zheng, "GVW algorithm over principal ideal domains," *Journal of Systems Science and Complexity*, vol. 26, no. 4, pp. 619–633, 2013.

[24] D. M. Li and J. W. Liu, "A Gröbner basis algorithm for ideals over zero-dimensional valuation rings," *Journal of Systems Science and Complexity*, vol. 33, pp. 1–14, 2020.

Hindawi
Complexity
Volume 2021, Article ID 9983999, 11 pages
https://doi.org/10.1155/2021/9983999

Research Article

Bipartite Consensus of Linear Discrete-Time Multiagent Systems with Exogenous Disturbances under Competitive Networks

Yi Yuan,[1] Shamrie Sainin Mohd ,[1] and Yanhui Zhu[2]

[1]*Faculty of Computing and Informatics, Universiti Malaysia Sabah, 88400, Kota Kinabalu Sabah, Malaysia*
[2]*School of Computer Science, Hunan University of Technology, Zhuzhou, 412008, China*

Correspondence should be addressed to Shamrie Sainin Mohd; shamrie@ums.edu.my

Received 30 March 2021; Accepted 15 May 2021; Published 26 May 2021

Academic Editor: Hou-Sheng Su

This paper investigates the bipartite consensus of linear discrete-time multiagent systems (MASs) with exogenous disturbances. A discrete-time disturbance-observer- (DTDO-) based technology is involved for attenuating the exogenous disturbances. And both the state feedback and observer-based output feedback bipartite consensus protocols are proposed by using the DTDO method. It turned out that bipartite consensus can be realized under the given protocols if the topology is connected and structurally balanced. Finally, numerical simulations are presented to illustrate the theoretical findings.

1. Introduction

In the past two decades, the coordination of MASs has attracted much attention for its wide applications [1–4]. In coordination issues, consensus plays a very important role, which means that the final states of all agents can asymptotically reach a common value. Many works about consensus have been reported in the past few years, including consensus of MASs with different dynamics [5–10] and consensus of MASs via different control methods [11–16].

Many existing works mainly focus on the consensus of cooperative networks. However, in many real networks, there exist competitive relations between the nodes. Bipartite consensus was firstly investigated for MASs with antagonistic links [17], in which the nodes were divided into two parts, one will asymptotically track the leader and the other part will asymptotically converge to the reverse state of the leader. The topology of the discussed network was assumed as structurally balanced. And gauge transformation was used for solving the problems of stability analysis. Then, many efforts were devoted to bipartite consensus of MASs with antagonistic links or competitive topologies. Bipartite consensus was investigated for MASs with time-varying delay [18]. Bipartite edge consensus was studied for MASs with edge dynamics under the corresponding line graph spanned by the nodal graph [19]. Qin et al. [20] investigated global bipartite consensus of MASs with input saturation, and Hu et al. [21] solved bipartite consensus problems of MASs with communication noise. The event-triggered bipartite consensus was investigated for first-order MASs [22]. Under different topologies, event-triggered adaptive bipartite output consensus of heterogenous linear MASs was studied [23]. Considering the rate of convergency, finite-time bipartite consensus was investigated [24].

However, the above papers mainly investigated bipartite consensus of continuous-time MASs without disturbances. Disturbance often exists and is the main resource of poor performance of the controlled systems. In real networks, the subsystem in the network may be sufferred by exogenous disturbances. Therefore, research studies about multiagent systems with exogenous disturbances are very important and significant. By using a novel backstepping method, robust global coordination was investigated for MASs with input saturation and disturbances [25]. In [26], the disturbance-

observer- (DO-) based control method was proposed for stabilizing nonlinear systems with exogenous disturbances. Using the continuous DO method proposed in [26], Yang et al. [27] solved the consensus problems of second-order MASs with exogenous disturbances. Containment control of continuous-time MASs with exogenous disturbances was investigated in virtue of the DO technique [28]. Intermittent consensus of MASs with exogenous disturbances was investigated by using both state feedback control and output feedback control protocols [29]. And the DO-based method is used on many other control plants [30, 31].

The existing relative works mainly focus on the consensus of MASs with continuous dynamics and cooperative topology. But there exist few results about bipartite consensus for discrete-time MASs with competitive topology and exogenous disturbances. Motivated by the above literatures, this paper investigates the bipartite consensus of discrete-time linear MASs with exogenous disturbances. The main contributions are as follows. (i) Discrete-time MASs are discussed in this paper, which are more challenging because the analysis of stability is more complex than the continuous-time MASs. (ii) Competitive network is considered in this paper, which can be used to describe more real network. And many results about the bipartite consensus of competitive networks can be applied into the consensus of cooperative networks. (iii) The DTDO-based method is used for attenuating the exogenous disturbances. Both the state feedback and output feedback control protocols are proposed in this paper.

The rest of the paper is organized as follows. Section 2 states the model considered in the paper and gives some basic lemmas and assumptions. In Section 3, discrete-time DO-based state feedback containment protocol is proposed. In Section 4, discrete-time DO-based output feedback containment protocol is given. Numerical examples are included to demonstrate the proposed protocol in Section 5. Finally, Section 6 gives a conclusion for this paper.

2. Preliminaries and Model Description

A network can be described by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$, which includes a set of nodes $\mathcal{V} = \{1, 2, \ldots, N\}$, a set of edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$, and an adjacent matrix $\mathcal{A} = [a_{ij}]$. For an undirected graph, $a_{ij} = a_{ji}$, for $i, j = 1, 2, \ldots, N$. $a_{ij} \neq 0 \Leftrightarrow (j, i) \in \mathcal{E}$. $\mathcal{N}_i = j | (j, i) \cdot \mathcal{E}$ is a neighbor set of ith node. A path between node v_{i_1} and node v_{i_n} is formed by an array of edges $(v_{i_1}, v_{i_2}), (v_{i_2}, v_{i_3}), \ldots, (v_{i_{n-1}}, v_{i_n})$. An undirected graph is connected if there exists a path between any two nodes. $L = [l_{ij}]_{N \times N}$ is the Laplacian matrix, where $l_{ii} = \sum_{j=1, j \neq i}^{N} |a_{ij}|$, $l_{ij} = -a_{ij}$.

The dynamics of the ith follower are described as

$$\begin{aligned} x_i(k+1) &= Ax_i(k) + B\big(u_i(k) + \gamma_i(k)\big), \\ y_i(k) &= Cx_i(k), \\ & i = 1, 2, \ldots, N. \end{aligned} \tag{1}$$

And the dynamics of the leader are described as

$$\begin{aligned} x_0(k+1) &= Ax_0(k), \\ y_0(k) &= Cx_0(k), \end{aligned} \tag{2}$$

where $x_i(k) \in R^n$, $u_i(k) \in R^p$, $\gamma_i(k) \in R^p$, and $y_i(k) \in R^m$ denote the state, control input, exogenous disturbance, and output of the ith follower, respectively, $x_0(k) \in R^n$ and $y_0(k) \in R^m$ denote the state and output of the leader, and $A \in R^{n \times n}, B \in R^{n \times p}$, and $C \in R^{m \times n}$ are constant matrices. It is assumed that the disturbance $\gamma_i(k)$ is generated by the following exogenous system:

$$\begin{aligned} \omega_i(k+1) &= S\omega_i(k), \\ \gamma_i(k) &= F\omega_i(k), \end{aligned} \tag{3}$$

where $\omega_i \in R^l$ is the state of the exogenous system and $S \in R^{l \times l}$ and $F \in R^{p \times l}$ are constant matrices.

The following assumptions and lemmas are necessary for the main results of this paper.

Definition 1 (see [24]). A signed graph $\mathcal{G}$ is said to be structurally balanced if the following hold:

(1) It admits a bipartition of nodes as $\mathcal{V}_1$ and $\mathcal{V}_2$, where $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$, $\mathcal{V}_1 \cup \mathcal{V}_2 = \mathcal{V}$.

(2) The elements of $\mathcal{V}_q, q \in \{1, 2\}$ have the following relation:

$$a_{ij} \geq 0, \forall i, j \in \mathcal{V}_q; a_{ij} \leq 0, \forall i \in \mathcal{V}_q, j \in \mathcal{V}_r, q \neq r. \tag{4}$$

Otherwise, the graph $\mathcal{G}$ is structurally unbalanced.

Definition 2. The leader-following bipartite consensus of system (1) with leader (2) is said to be achieved, if there exists a protocol $u_i(k)$ such that

$$\lim_{t \longrightarrow \infty} \left\| x_i(k) - d_i x_0(k) \right\| = 0, \quad i = 1, 2, \ldots, N, \tag{5}$$

for any initial condition $x_l(1), l = 0, 1, \ldots, N$, where $d_i = 1$, if $i \in \mathcal{V}_1$; $d_i = -1$, if $i \in \mathcal{V}_2$.

Remark 1. According to gauge transformation [17], denoting $D = \text{diag}\{d_1, d_2, \ldots, d_N\}$, one has that $DL\,D$ is semipositive defined under the assumption that G is connected.

Assumption 1. Suppose the undirected signed graph $\mathcal{G}$ is connected and structurally balanced.

Assumption 2. The matrix pair (A, B) is stabilizable.

Assumption 3. The matrix pair (A, C) is detectable.

Lemma 1 (see [32]). *For any matrix $\Lambda = diag(h, 0, \ldots, 0)$ with $h > 0$, all eigenvalues of the matrix $D\tilde{L}D = DL\,D + \Lambda$ are positive if Assumption 1 holds, where $\tilde{L} = L + \Lambda$.*

Lemma 2 (see [33]). *Under Assumption 2, there exists a unique positive definite matrix P, satisfying the algebraic Riccati equation*

$$A^T PA - P - A^T PB\left(B^T PB + I\right)^{-1} B^T PA + I = 0. \quad (6)$$

Lemma 3 (see [34]). *Suppose that A, B, C are matrices with appropriate dimensions; then, the following inequalities are equivalent:*

(1) $\begin{pmatrix} A & B \\ B^T & C \end{pmatrix} < 0.$

(2) $A < 0$ *and* $C - B^T A^{-1} B < 0.$

(3) $C < 0$ *and* $A - BC^{-1} B^T < 0.$

3. DO-Based State Feedback Bipartite Consensus

In this section, based on the DTDO method, bipartite consensus of MAS with disturbances is solved by using relative state information.

A disturbance observer is designed as follows:

$$\begin{aligned} v_i(k+1) &= (S + HBF)\left(v_i(k) - Hx_i(k)\right) + H\left(Ax_i(k) + Bu_i(k)\right), \\ \widehat{\omega}_i(k) &= v_i(k) - Hx_i(k), \\ \widehat{\gamma}_i(k) &= F\widehat{\omega}_i(k), \end{aligned} \quad (7)$$

where $v_i(k) \in R^{l\times l}$ is the internal state variable of the observer, $\widehat{\omega}_i(k) \in R^l$ and $\widehat{\gamma}_i(k) \in R^p$ are the estimated values of $\omega_i(k)$ and $\gamma_i(k)$, respectively, and $H \in R^{l\times n}$ is the gain matrix of the observer.

Remark 2. The agents in the network cannot get the information of the disturbances, which leads to that the agents have to estimate the value of the exogenous disturbances. A discrete disturbance observer (7) is proposed for estimating the disturbances.

According to (1) and (7), one has

$$\begin{aligned} \widehat{\omega}_i(k+1) &= v_i(k+1) - Hx_i(k+1) \\ &= (S + HBF)\left(v_i(k) - Hx_i(k)\right) + H\left(Ax_i(k)\right) \\ &\quad \left(+Bu_i(k)\right) - H\left(Ax_i(k) + B\left(u_i(k)\right) + \gamma_i(k)\right) \\ &= (S + HBF)\widehat{\omega}_i(k) - HBF\omega_i(k). \end{aligned} \quad (8)$$

Then, denoting the state error of exogenous system as $e_i(k)$, one has

$$\begin{aligned} e_i(k+1) &= \omega_i(k+1) - \widehat{\omega}_i(k+1) \\ &= S\omega_i(k) - (S + HBF)\widehat{\omega}_i(k) + HBF\omega_i(k) \\ &= (S + HBF)\left(\omega_i(k) - \widehat{\omega}_i(k)\right) \\ &= (S + HBF)e_i(k). \end{aligned} \quad (9)$$

Consider the following distributed bipartite consensus protocol for discrete-time MAS (1):

$$u_i(k) = K\left[\sum_{j\in N_i} \left|a_{ij}\right|\left(x_i(k) - \operatorname{sgn}\left(a_{ij}\right)x_j(k)\right) + h_i\left(x_i(k) - d_i x_0(k)\right)\right] - \widehat{\gamma}_i(k), \quad (10)$$

where K is the gain matrix to be determined.

Substituting (10) into (1), one has that by (7) and (9),

$$x_i(k+1) = Ax_i(k) + BFe_i(k) + BK\left[\sum_{j\in N_i} \left|a_{ij}\right|\left(x_i(k)\left(-\operatorname{sgn}\left(a_{ij}\right)x_j(k)\right) + h_i\left(x_i(k) - d_i x_0(k)\right)\right)\right]. \quad (11)$$

Theorem 1. *Suppose Assumptions 1 and 2 hold. The bipartite consensus of MAS (1) with leader (2) will be achieved by error system (9) with disturbance observer (7) under bipartite consensus protocol (10) if*

(i) Suppose there exists at least one follower pinned by the leader.

(ii) $S + HBF$ *is Schur stable.*

(iii) $K = -(1/\lambda_1(D\tilde{L}D))(B^T PB + I)^{-1} B^T PA$, *where* $P > 0$ *is the unique solution of algebraic Riccati equation (6),* $\lambda_1(D\tilde{L}D)$ *is the minimum eigenvalue of* $D\tilde{L}D$, *and* $D\tilde{L}D = DL\ D + \Lambda, \Lambda = diag(h_1, h_2, \ldots, h_N)$.

Proof. Let $\tilde{x}_i(k) = x_i(k) - d_i x_0(k)$. Because $d_i a_{ij} d_j = a_{ij}\operatorname{sgn}(a_{ij}) = |a_{ij}|$, one has

$$\begin{aligned} \tilde{x}_i(k+1) &= x_i(k+1) - d_i x_0(k+1) \\ &= A\tilde{x}_i(k) + BFe_i(k) + BK\left[\sum_{j\in N_i} \left|a_{ij}\right|\left(\tilde{x}_i(k) - \operatorname{sgn}\left(a_{ij}\right)\tilde{x}_j(k)\right) + h_i\tilde{x}_i(k)\right], \\ e_i(k+1) &= (S + HBF)e_i(k). \end{aligned} \quad (12)$$

Denote $\tilde{x}(k) = (\tilde{x}_1^T(k), \tilde{x}_2^T(k), \ldots, \tilde{x}_N^T(k))^T$, $e(k) = (e_1^T(k), e_2^T(k), \ldots, e_N^T(k))^T$. Then, (12) can be rewritten as follows:

$$\begin{aligned}\tilde{x}(k+1) &= \left(I_N \otimes A + \tilde{L} \otimes BK\right)\tilde{x}(k) + \left(I_N \otimes BF\right)e(k),\\ e(k+1) &= \left(I_N \otimes (S + HBF)\right)e(k).\end{aligned} \tag{13}$$

Consider the following Lyapunov function candidate

$$\begin{aligned}V(k) &= \eta^T(k)\begin{pmatrix} I_N \otimes P & 0 \\ 0 & \alpha I_N \otimes Q_1 \end{pmatrix}\eta(k)\\ &= \tilde{x}^T(k)\left(I_N \otimes P\right)\tilde{x}(k) + \alpha e^T(k)\left(I_N \otimes Q_1\right)e(k),\end{aligned} \tag{14}$$

where $\eta(k) = (\tilde{x}^T(k), e^T(k))^T$, $Q_1 > 0$ is a matrix designed later, and α is a large enough positive constant.

Let $\overline{x}(k) = (D \otimes I_n)\tilde{x}(k)$ and $\overline{\eta}(k) = (\overline{x}^T(k), e^T(k))^T$. Then,

$$\begin{aligned}V(k+1) - V(k) &= \overline{x}^T(k)\left(I_N \otimes A^T PA + 2\, D\tilde{L}D \otimes A^T PBK\right)\\ &\quad + \left(\left(DL\, D\right)^2 \otimes K^T B^T PBK - I_N \otimes P\right)\overline{x}(k)\\ &\quad + \overline{x}^T(k)\left(D \otimes A^T PBF + D\tilde{L} \otimes K^T B^T PBF\right)e(k)\\ &\quad + e^T(k)\left(D \otimes F^T B^T PA + \tilde{L}D \otimes F^T B^T PBK\right)\overline{x}(k)\\ &\quad + e^T(k)\left(I_N \otimes \alpha(S + HBF)^T Q_1 (S + HBF) - \alpha Q_1 + F^T B^T PBF\right)e(k)\\ &= \overline{\eta}^T(k)\begin{pmatrix} \Xi_1 & *_1 \\ *_1^T & \Xi_2 \end{pmatrix}\overline{\eta}(k),\end{aligned} \tag{15}$$

where

$$\begin{aligned}\Xi_1 &= I_N \otimes A^T PA + 2\, D\tilde{L}D \otimes A^T PBK + \left(DL\, D\right)^2 \otimes K^T B^T PBK - I_N \otimes P,\\ \Xi_2 &= I_N \otimes \left(\alpha(S + HBF)^T Q_1 (S + HBF) - \alpha Q_1 + F^T B^T PBF\right),\\ *_1 &= D \otimes A^T PBF + D\tilde{L} \otimes K^T B^T PBF.\end{aligned} \tag{16}$$

Since $D\tilde{L}D$ is a symmetric matrix, there exists an orthodox matrix Υ such that $\Upsilon^T(D\tilde{L}D)\Upsilon = \Lambda$, where $\Lambda = \mathrm{diag}(\lambda_1(D\tilde{L}D), \lambda_2(D\tilde{L}D), \ldots, \lambda_N(D\tilde{L}D))$. Moreover, by Assumption 1, (i), and Lemma 1, one has a sequence of eigenvalues $0 < \lambda_1(D\tilde{L}D) \le \lambda_2(D\tilde{L}D) \le \ldots \le \lambda_N(D\tilde{L}D)$. Then, let $\check{x}(k) = (\Upsilon^T \otimes I_n)\overline{x}(k)$; by (iii) and Lemma 2, one has

$$\begin{aligned}\overline{x}^T(k)\Xi_1\overline{x}(k) &= \hat{x}^T(k)\left(I_N \otimes A^T PA + 2\Lambda \otimes A^T PBK + \Lambda^2 \otimes K^T B^T PBK - I_N \otimes P\right)\hat{x}(k)\\ &\le -\check{x}^T(k)\check{x}(k) = -\overline{x}^T(k)\overline{x}(k).\end{aligned} \tag{17}$$

Meanwhile, by (ii), there exists a positive definite matrix $Q_1 > 0$ such that the following discrete Lyapunov matrix equation holds:

$$(S + HBF)^T Q_1 (S + HBF) - Q_1 + I_l = 0. \tag{18}$$

According to (17) and (18), one has

$$\begin{aligned} &V(k+1) - V(k) \\ &\le \overline{\eta}^T(k) \begin{pmatrix} -I_{Nn} & *_1 \\ *_1^T & -\Psi_1 \end{pmatrix} \overline{\eta}(k) \\ &= \eta^T(k) \begin{pmatrix} -I_{Nn} & *_2 \\ *_2^T & -\Psi_1 \end{pmatrix} \eta(k) \\ &= \eta T(k) \Phi_1 \eta(k), \end{aligned} \tag{19}$$

where $\Phi_1 = \begin{pmatrix} -I_{Nn} & *_2 \\ *_2^T & -\Psi_1 \end{pmatrix}$, $\Psi_1 = I_N \otimes (\alpha I_l - F^T B^T P B F)$ and $*_2 = (D \otimes I_n) *_1$. By Lemma 3, one has $\Phi_1 < 0$ by choosing sufficiently large constant α. Then,

$$\begin{aligned} &V(k+1) - V(k) \\ &\le \lambda_{\max}(\Phi_1) \eta^T(k) \eta(k) \\ &= \lambda_{\max}(\Phi_1) \sum_{i=1}^{N} \left(\widetilde{x}_i^T(k) \widetilde{x}_i(k) + e_i^T(k) e_i(k) \right) \\ &\le 0, \end{aligned} \tag{20}$$

where $V(k+1) - V(k) = 0$ if and only if $\widetilde{x}_1 = \widetilde{x}_2 = \cdots = \widetilde{x}_N = 0$ and $e_1 = e_2 = \cdots = e_N = 0$. Thus, $\lim_{k \longrightarrow \infty} \widetilde{x}_i(k) = 0$, $\lim_{k \longrightarrow \infty} e_i(k) = 0$, $i = 1, 2, \ldots, N$, which implies that Theorem 1 holds. □

4. DO-Based Output Feedback Bipartite Consensus

In this section, a DO-based output feedback protocol is proposed for disturbed linear MASs. State observer and disturbance observers are given, respectively, and the conditions are obtained.

The state observer of the ith follower is designed as

$$\begin{aligned} \hat{x}_i(k+1) &= A\hat{x}_i(k) + B\left(u_i(k) + \gamma_i(k)\right) \\ &\quad + E\left(C\hat{x}_i(k) - y_i(k)\right), \quad i = 1, 2, \ldots, N, \end{aligned} \tag{21}$$

and the state observer of the leader is designed as

$$\hat{x}_0(k+1) = A\hat{x}_0(k) + E\left(C\hat{x}_0(k) - y_0(k)\right), \tag{22}$$

where $\hat{x}_i(k)$ is the observed state of the ith follower, $\hat{x}_0(k)$ is the observed state of the leader, and E is the gain matrix to be determined.

The discrete-time disturbance observer based on output information is proposed as

$$\begin{aligned} v_i(k+1) &= (S + HBF)\left(v_i(k) - H\hat{x}_i(k)\right) \\ &\quad + H\left(A\hat{x}_i(k) + Bu_i(k) + E\left(C\hat{x}_i(k) - y_i(k)\right)\right), \\ \widehat{\omega}_i(k) &= v_i(k) - H\hat{x}_i(k), \\ \widehat{d}_i(k) &= F\widehat{\omega}_i(k), \end{aligned} \tag{23}$$

where $v_i \in R^l$ is the internal state variable of the observer, $\widehat{\omega}_i$ and $\widehat{\gamma}_i$ are the estimated values of ω_i and γ_i, respectively, and $H \in R^{l \times n}$ is the gain matrix of the observer.

Remark 3. For the case that the state of each agent cannot be obtained, the state observer can be used for estimating the state. Moreover, the disturbances exist in the subsystems. One has to design corresponding controller to attenuate the disturbances. Discrete output-based disturbance observer (23) is proposed.

When the state cannot be obtained, state observer can be used for estimating the state. Therefore, the bipartite consensus protocol can be designed as

$$u_i(k) = K\left[\sum_{j \in N_i} \left|a_{ij}\right| \left(\hat{x}_i(k) - \operatorname{sgn}\left(a_{ij}\right) \hat{x}_j(k)\right) + h_i\left(\hat{x}_i(k) - d_i \hat{x}_0(k)\right)\right] - \widehat{\gamma}_i(k), \tag{24}$$

where K is the gain to be designed.

Substituting (24) into (1), one has

$$\begin{aligned} x_i(k+1) &= Ax_i(k) + BF\left(\omega_i(k) - \widehat{\omega}_i(k)\right) \\ &\quad + BK\left[\sum_{j \in N_i} \left|a_{ij}\right| \left(\hat{x}_i(k) - \operatorname{sgn}\left(a_{ij}\right) \hat{x}_j(k)\right) + h_i\left(\hat{x}_i(k) - d_i \hat{x}_0(k)\right)\right], \end{aligned} \tag{25}$$

and then, we give the following result.

Theorem 2. *Suppose Assumptions 1–3 hold. The bipartite consensus of MAS (1) with leader (2) will be achieved with state observers (21) and (22) and disturbance observer (23) under bipartite consensus protocol (24) if*

(i) Suppose there exists at least one follower pinned by the leader.

(ii) $S + HBF$ is Schur stable.

(iii) $K = -(1/\lambda_1(D\tilde{L}D))(B^TPB + I)^{-1}B^TPA$, where $P > 0$ is the unique solution of algebraic Riccati equation (6), $\lambda_1(D\tilde{L}D)$ is the minimum eigenvalue of $D\tilde{L}D$, and $D\tilde{L}D = DL\ D + \overline{H}, \overline{H} = diag(h_1, h_2, \ldots, h_N)$.

Proof. Let $\tilde{x}_i(k) = x_i(k) - d_i x_0(k)$, $\tilde{\hat{x}}_i(k) = \hat{x}_i(k) - d_i\hat{x}_0(k)$, $\delta_i(k) = \tilde{x}_i(k) - \tilde{\hat{x}}_i(k)$, and $e_i(k) = \omega_i(k) - \hat{\omega}_i(k)$. By $d_i a_{ij} d_j = a_{ij}\mathrm{sgn}(a_{ij}) = |a_{ij}|$, one has

$$\begin{aligned}
&\left|a_{ij}\right|\left(\hat{x}_i(k) - \mathrm{sgn}\left(a_{ij}\right)\hat{x}_j(k)\right)\\
&\quad = \left|a_{ij}\right|\left(\tilde{\hat{x}}_i(k) + d_i\hat{x}_0(k) - \mathrm{sgn}\left(a_{ij}\right)\hat{x}_j(k)\right)\\
&\quad = \left|a_{ij}\right|\left(\tilde{\hat{x}}_i(k) - \mathrm{sgn}\left(a_{ij}\right)\tilde{\hat{x}}_j(k)\right)\\
&\quad = \left|a_{ij}\right|\left(\tilde{x}_i(k) - \mathrm{sgn}\left(a_{ij}\right)\tilde{x}_j(k)\right)\\
&\quad \left(+\mathrm{sgn}\left(a_{ij}\right)\delta_j(k) - \delta_j(k)\right),
\end{aligned} \tag{26}$$

and then, the error system can be written as follows:

$$\begin{aligned}
\tilde{x}_i(k+1) &= A\tilde{x}_i(k) + BFe_i(k) + BK\left[\sum_{j\in N_i}\left|a_{ij}\right|\left(x_i(k) - \mathrm{sgn}\left(a_{ij}\right)\tilde{x}_j(k) + \mathrm{sgn}\left(a_{ij}\right)\delta_j(k) - \delta_j(k)\right)\right.\\
&\quad \left. + h_i\left(\tilde{x}_i(k) - \delta_i(k)\right)\right],\\
\tilde{\hat{x}}_i(k+1) &= A\tilde{\hat{x}}_i(k) + BF\left(e_i(k) - EC\delta_i(k)\right)\\
&\quad + BK\left[\sum_{j\in N_i}\left|a_{ij}\right|\left(\tilde{x}_i(k) - \mathrm{sgn}\left(a_{ij}\right)\tilde{x}_j(k) + \mathrm{sgn}\left(a_{ij}\right)\delta_j(k) - \delta_j(k)\right)\right.\\
&\quad \left. + h_i\left(\tilde{x}_i(k) - \delta_i(k)\right)\right],\\
\delta_i(k+1) &= (A + EC)\delta_i(k),\\
e_i(k+1) &= (S + HBF)e_i(k).
\end{aligned} \tag{27}$$

Denote $\tilde{x}(k) = (\tilde{x}_1^T(k), \tilde{x}_2^T(k), \ldots, \tilde{x}_N^T(k))^T$, $\delta(k) = (\delta_1^T(k), \delta_2^T(k), \ldots, \delta_N^T(k))^T$, $e(k) = (e_1^T(k), e_2^T(k), \ldots, e_N^T(k))^T$, and (27) can be rewritten as follows:

$$\begin{aligned}
\tilde{x}(k+1) &= \left(I_N \otimes A + L \otimes BK\right)\tilde{x}(k) + \left(I_N \otimes BF\right)e(k)\\
&\quad - (L \otimes BK)\delta(k),\\
\delta(k+1) &= \left(I_N \otimes (A + EC)\right)\delta(k),\\
e(k+1) &= \left(I_N \otimes (S + HBF)\right)e(k).
\end{aligned} \tag{28}$$

Consider the following Lyapunov candidate function:

$$\begin{aligned}
V(k) &= \zeta^T(k)\begin{pmatrix} I_N \otimes P & 0 & 0\\ 0 & \beta I_N \otimes Q_2 & 0\\ 0 & 0 & \beta I_N \otimes Q_3 \end{pmatrix}\zeta(k)\\
&= \tilde{x}^T(k)(I_N \otimes P)\tilde{x}(k) + \beta\delta^T(k)(I_N \otimes Q_2)\delta(k)\\
&\quad + \beta e^T(k)(I_N \otimes Q_3)e(k),
\end{aligned} \tag{29}$$

where $\zeta(k) = (\tilde{x}^T(k), \delta^T(k), e^T(k))^T$, Q_2 and Q_3 are positive definite matrices which were designed later, and $\beta > 0$ is a sufficiently large constant.

Let $\overline{x}(k) = (D \otimes I_n)\tilde{x}(k)$, $\overline{\zeta}(k) = (\overline{x}^T(k), \delta^T(k), e^T(k))^T$, and then

$$
\begin{aligned}
&V(k+1)-V(k)\\
&\quad=\overline{x}^{T}(k)\Big(I_N\otimes A^{T}PA+2\ D\tilde{L}D\otimes A^{T}PBK\\
&\quad+(DL\ D)^{2}\otimes K^{T}B^{T}PBK-I_N\otimes P\Big)\overline{x}(k)\\
&\quad-\overline{x}^{T}(k)\Big(D\tilde{L}\otimes A^{T}PBK+D\tilde{L}^{2}\otimes K^{T}B^{T}PBK\Big)\delta(k)\\
&\quad-\delta^{T}(k)\Big(\tilde{L}D\otimes K^{T}B^{T}PA+\tilde{L}^{2}D\otimes K^{T}B^{T}PBK\Big)\overline{x}(k)\\
&\quad+\overline{x}^{T}(k)\Big(D\otimes A^{T}PBF+D\tilde{L}\otimes K^{T}B^{T}PBF\Big)e(k)\\
&\quad+e^{T}(k)\Big(D\otimes F^{T}B^{T}PA+\tilde{L}D\otimes F^{T}B^{T}PBK\Big)\overline{x}(k)\\
&\quad-\delta^{T}(k)\Big(\tilde{L}\otimes K^{T}B^{T}PBF\Big)e(k)-e^{T}(k)\Big(\tilde{L}\otimes F^{T}B^{T}P\times BK\Big)\\
&\quad\cdot\delta(k)+\delta^{T}(k)\Big(\beta I_N\otimes\Big((A+EC)^{T}Q_2(A+EC)\\
&\quad-Q_2\Big)+\tilde{L}^{2}\otimes K^{T}B^{T}PBK\Big)\delta(k)+e^{T}(k)\big(I_N\otimes(\beta(S\\
&\quad+HBF)^{T}Q_3(S+HBF)-\beta Q_3+F^{T}B^{T}PBF\big)\big)e(k)\\
&\quad=\overline{\zeta}^{T}(k)\begin{pmatrix}\Xi_3 & *_3 & *_4\\ *_3^{T} & \Xi_4 & *_5\\ *_4^{T} & *_5^{T} & \Xi_5\end{pmatrix}\overline{\zeta}(k),
\end{aligned}
\tag{30}
$$

where

$$
\begin{aligned}
\Xi_3&=I_N\otimes A^{T}PA+2\ D\tilde{L}D\otimes A^{T}PBK\\
&\quad+(DL\ D)^{2}\otimes K^{T}B^{T}PBK-I_N\otimes P,\\
\Xi_4&=\beta I_N\otimes\Big((A+EC)^{T}Q_2(A+EC)-Q_2\Big)+\tilde{L}^{2}\otimes K^{T}B^{T}PBK,\\
\Xi_5&=I_N\otimes\Big(\beta(S+HBF)^{T}Q_3(S+HBF)-\beta Q_3+F^{T}B^{T}PBF\Big),\\
*_3&=D\tilde{L}\otimes A^{T}PBK+D\tilde{L}^{2}\otimes K^{T}B^{T}PBK,\\
*_4&=D\otimes A^{T}PBF+D\tilde{L}\otimes K^{T}B^{T}PBF,\\
*_5&=\tilde{L}\otimes K^{T}B^{T}PBF.
\end{aligned}
\tag{31}
$$

And, under Assumption 3, the gain matrix E can be selected such that $A+EC$ is Schur.

Similar to (17), there exists an orthodox matrix Υ such that $\Upsilon^{T}(D\tilde{L}D)\Upsilon=\Lambda$, where $\Lambda=\mathrm{diag}(\lambda_1(D\tilde{L}D),\lambda_2(D\tilde{L}D),\ldots,\lambda_N(D\tilde{L}D))$. Let $\check{x}(k)=(\Upsilon^{T}\otimes I_n)\overline{x}(k)$. By (iii) and Lemma 2, one has

$$
\overline{x}^{T}(k)\Xi_3\overline{x}(k)\le-\check{x}^{T}(k)\check{x}(k)=-\overline{x}^{T}(k)\overline{x}(k),
\tag{32}
$$

and since matrices $A+EC$ and $S+HBF$ are Schur, there exist, respectively, positive definite matrices Q_2 and Q_3 such that the discrete Lyapunov matrix equations hold as follows:

$$
(A+EC)^{T}Q_2(A+EC)-Q_2+I_n=0,
\tag{33}
$$

$$
(S+HBF)^{T}Q_3(S+HBF)-Q_3+I_l=0.
\tag{34}
$$

According to (32)–(34), one has

$$
\begin{aligned}
&V(k+1)-V(k)\\
&\le\overline{\zeta}^{T}(k)\begin{pmatrix}-I_{Nn} & *_3 & *_4\\ *_3^{T} & -\Psi_2 & *_5\\ *_4^{T} & *_5^{T} & -\Psi_3\end{pmatrix}\overline{\zeta}(k)\\
&=\zeta^{T}(k)\begin{pmatrix}-I_{Nn} & *_6 & *_7\\ *_6^{T} & -\Psi_2 & *_5\\ *_7^{T} & *_5^{T} & -\Psi_3\end{pmatrix}\zeta(k)\\
&=\zeta^{T}(k)\Phi_2\zeta(k),
\end{aligned}
\tag{35}
$$

where

$$
\begin{aligned}
\Phi_2&=\begin{pmatrix}-I_{Nn} & *_6 & *_7\\ *_6^{T} & -\Psi_2 & *_5\\ *_7^{T} & *_5^{T} & -\Psi_3\end{pmatrix},\\
\Psi_2&=\beta I_{Nn}-\tilde{L}^{2}\otimes K^{T}B^{T}PBK,\\
\Psi_3&=I_N\otimes\Big(\beta I_l-F^{T}B^{T}PBF\Big),\\
*_6&=(D\otimes I_n)*_3,\\
*_7&=(D\otimes I_n)*_4.
\end{aligned}
\tag{36}
$$

By choosing sufficiently large constant β, according to Lemma 3, one has $\Phi_2<0$. Then,

$$
\begin{aligned}
&V(k+1)-V(k)\\
&\le\lambda_{\max}(\Phi_2)\zeta^{T}(k)\zeta(k)\\
&=\lambda_{\max}(\Phi_2)\sum_{i=1}^{N}\Big(\tilde{x}_i^{T}(k)\tilde{x}_i(k)+\delta_i^{T}(k)\delta_i(k)+e_i^{T}(k)e_i(k)\Big)\\
&\le0,
\end{aligned}
\tag{37}
$$

where $V(k+1)-V(k)=0$ if and only if $\tilde{x}_1=\tilde{x}_2=\cdots=\tilde{x}_N=0$, $\delta_1(k)=\delta_2(k)=\cdots=\delta_N(k)=0$, and $e_1=e_2=\cdots=e_N=0$. Thus, $\lim_{k\longrightarrow\infty}\tilde{x}_i(k)=0$, $\lim_{k\longrightarrow\infty}\delta_i(k)=0$, $\lim_{k\longrightarrow\infty}e_i(k)=0$. Moreover, $\lim_{k\longrightarrow\infty}\hat{x}_i(k)=0$, $i=1,2,\ldots,N$. Thus, Theorem 2 holds. □

5. Simulations

In this section, we give two simulation examples to illustrate the theoretical results of Sections 3 and 4. In Figure 1, consider seven agents composed of six followers with one leader in a network with competitive interaction. Moreover, we choose $n=2$ and $m=3$ as the dimensions of the state and output of the seven agents, respectively.

Example 1. (the case of state feedback) In Figure 1, note that the signed topology is connected and the bipartition $\mathcal{V}_1=\{1,2,3\}$, $\mathcal{V}_2=\{4,5,6\}$ satisfies Definition 1. Therefore, Assumption 1 is satisfied. Moreover, the matrix D can be selected as $D=\mathrm{diag}(1,1,1,-1,-1,-1)$. Meanwhile, the matrices $\tilde{L}$ and $D\tilde{L}D$ can be calculated as follows:

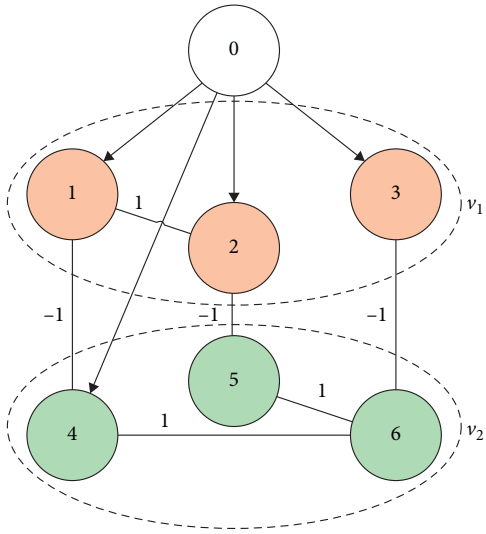


Figure 1: A network with competitive interactions containing 7 agents.

$$\widetilde{L} = \begin{pmatrix} 3 & -1 & 0 & 1 & 0 & 0 \\ -1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 3 & 0 & -1 \\ 0 & 1 & 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & -1 & -1 & 3 \end{pmatrix},$$
$$D\widetilde{L}D = \begin{pmatrix} 3 & -1 & 0 & -1 & 0 & 0 \\ -1 & 3 & 0 & 0 & -1 & 0 \\ 0 & 0 & 2 & 0 & 0 & -1 \\ -1 & 0 & 0 & 3 & 0 & -1 \\ 0 & -1 & 0 & 0 & 2 & -1 \\ 0 & 0 & -1 & -1 & -1 & 3 \end{pmatrix}, \tag{38}$$

and one can obtain $\lambda_1(D\widetilde{L}D) = 0.5395$.

Then, choose the following system matrices:

$$A = \begin{pmatrix} 0.9 & -2 \\ 0 & -1.01 \end{pmatrix},$$
$$B = \begin{pmatrix} 0.02 \\ -0.05 \end{pmatrix},$$
$$F = \begin{pmatrix} -0.03 & 1.8 \end{pmatrix}, \tag{39}$$
$$S = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix},$$

where matrices A and B satisfy Assumption 2. And the gain matrix H can be selected as follows:

$$H = \begin{pmatrix} -0.2 & 0 \\ 0 & -0.5 \end{pmatrix}, \tag{40}$$

and thus one has that the matrix $S + HBF$ is Schur. By algebraic Riccati equation (6), one has

$$P = \begin{pmatrix} 4.8332 & -4.7634 \\ -4.7634 & 40.0959 \end{pmatrix}, \tag{41}$$

and thus the gain matrix K of bipartite consensus protocol (10) can be obtained:

$$K = \begin{pmatrix} -0.5025 & -2.4200 \end{pmatrix}. \tag{42}$$

For the MAS (1) with leader (2) and exogenous system (3), the initial values of $x_i(k)$, $x_0(k)$, and $\omega_i(k)$ are given as follows:

$$\begin{aligned} x_1(1) &= (9, 10)^T, \\ x_2(1) &= (-16, -4)^T, \\ x_3(1) &= (12, -15)^T, \\ x_4(1) &= (-11, -9)^T, \\ x_5(1) &= (-9, 14)^T, \\ x_6(1) &= (13, 7)^T, \\ x_0(1) &= (1.5, -1.5)^T, \\ \omega_1(1) &= (2, 2)^T, \\ \omega_2(1) &= (-1, -1.2)^T, \\ \omega_3(1) &= (-3, -3)^T, \\ \omega_4(1) &= (5, 4)^T, \\ \omega_5(1) &= (-2, -2)^T, \\ \omega_6(1) &= (1, 1)^T, \end{aligned} \tag{43}$$

where $i = 1, 2, \ldots, 6$. In Figures 2 and 3, the trajectories of $x_i(k)$, $i = 1, 2, \ldots, 6$, are displayed. And one can obtain that the bipartite consensus of MAS (1) with leader (2) can be achieved under bipartite consensus protocol (10) by Figures 2 and 3. Thus, the effectiveness of Theorem 1 is verified.

Example 2. (the case of output feedback) In this case, consider the same system matrices A, B, F, S and the gain matrix H as Example 1, and one has that Assumption 1 can be satisfied and the solution $P > 0$ of algebraic Riccati equation (6) and the same gain matrix K as Example 1 can be calculated. Moreover, we can obtain that the matrix $S + HBF$ is Schur. Then, choose the output matrix C as follows:

$$C = \begin{pmatrix} -0.5 & 0 \\ 0 & 0.5 \\ 0 & 0 \end{pmatrix}, \tag{44}$$

and thus the matrix pair (A, C) is detectable satisfying Assumption 3. Meanwhile, the matrix E can be selected as follows:

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \tag{45}$$

and then one has that $A + EC$ is Schur.

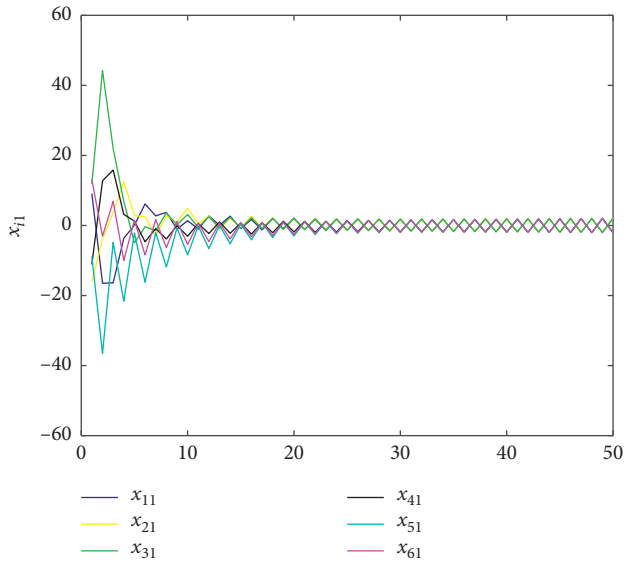


FIGURE 2: The trajectories of $x_{i1}, i = 1, 2, \ldots, 6$, under state feedback protocol.

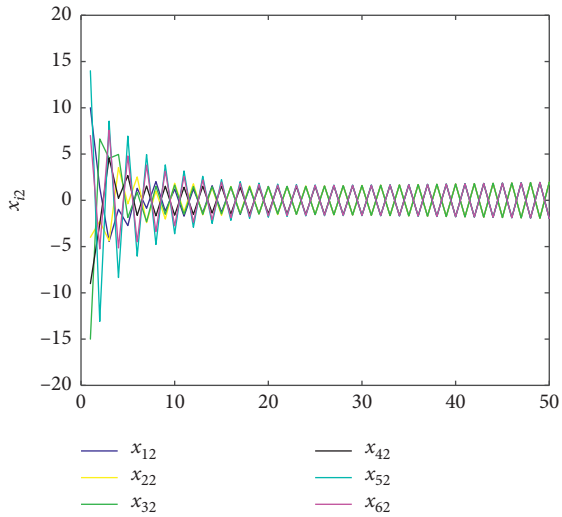


FIGURE 3: The trajectories of $x_{i2}, i = 1, 2, \ldots, 6$, under state feedback protocol.

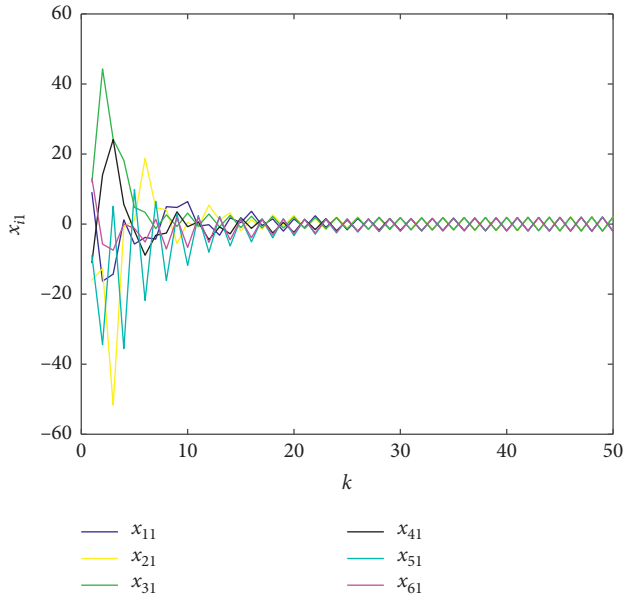


FIGURE 4: The trajectories of $x_{i1}, i = 1, 2, \ldots, 6$, under output feedback protocol.

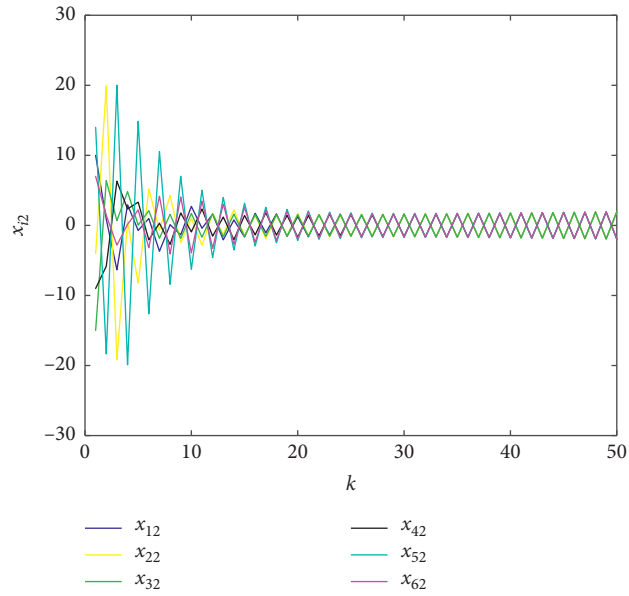


FIGURE 5: The trajectories of $x_{i2}, i = 1, 2, \ldots, 6$, under output feedback protocol.

For MAS (1) with leader (2) and exogenous system (3), we choose the same initial values of $x_i(k)$, $x_0(k)$, and $\omega_i(k)$ as Example 1 and the observed values of $\hat{x}_i(k)$, $\hat{x}_0(k)$ as follows: $\hat{x}_1(1) = (14, 13)^T$, $\hat{x}_2(1) = (18, 17)^T$, $\hat{x}_3(1) = (-10, -11)^T$, $\hat{x}_4(1) = (-14, -12)^T$, $\hat{x}_5(1) = (-16, -17)^T$, $\hat{x}_6(1) = (11, 14)^T$, $\hat{x}_0(1) = (-0.1, -0.1)^T$. In Figures 4 and 5, the trajectories of $x_i(k)$ are presented, and one can note that the bipartite consensus can be achieved for MAS (1) with leader (2) via bipartite consensus protocol (24). Thus, the effectiveness of Theorem 2 is verified.

6. Conclusions

In this paper, bipartite consensus is investigated for discrete-time MASs with exogenous disturbances. With the help of DTDO proposed in this paper, both the state feedback and the output feedback protocols are given, in which the gains can be determined by solving some discrete-time algebraic Riccati equations. Then, using stability theory, some sufficient conditions are obtained. Finally, numerical simulations are presented to illustrate the theoretical findings.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This study was partially supported by the Natural Science Foundation of Hunan Province (2020JJ6089) and the Key Project of the Department of Education in Hunan Province (19A133).

References

[1] H. Zhang, X. Zhou, Z. Wang, H. Yan, and J. Sun, "Adaptive consensus-based distributed target tracking with dynamic cluster in sensor networks," *IEEE Transactions on Cybernetics*, vol. 49, no. 5, pp. 1580–1591, 2019.

[2] D. Wang, Z. Wang, C. Wen, and W. Wang, "Second-order continuous-time algorithm for optimal resource allocation in power systems," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 2, pp. 626–637, 2019.

[3] H. Yan, X. Zhou, H. Zhang, F. Yang, and Z.-G. Wu, "A novel sliding mode estimation for microgrid control with communication time delays," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1509–1520, 2019.

[4] D. Ding, Z. Wang, and Q.-L. Han, "A set-membership approach to event-triggered filtering for general nonlinear systems over sensor networks," *IEEE Transactions on Automatic Control*, vol. 65, no. 4, pp. 1792–1799, 2020.

[5] A. Jadbabaie, J. Jie Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, pp. 988–1001, 2003.

[6] W. Ren and R. Beard, "Consensus seeking in multi-agent systems under dynamically changing interaction topologies," *IEEE Transactions on Automatic Control*, vol. 50, no. 5, pp. 655–661, 2005.

[7] Z. Li and J. Chen, "Robust consensus of linear feedback protocols over uncertain network graphs," *IEEE Transactions on Automatic Control*, vol. 62, no. 8, pp. 4251–4258, 2017.

[8] H. Su, H. Wu, and J. Lam, "Positive edge-consensus for nodal networks via output feedback," *IEEE Transactions on Automatic Control*, vol. 64, no. 3, pp. 1244–1249, 2019.

[9] X. Wang, G.-P. Jiang, H. Su, and Z. Zeng, "Consensus-based distributed reduced-order observer design for LTI systems," *IEEE Transactions on Cybernetics*, pp. 1–11, 2020.

[10] X. Wang, X. Wang, H. Su, and J. Lam, "Coordination control for uncertain networked systems using interval observers," *IEEE Transactions on Cybernetics*, vol. 50, no. 9, pp. 4008–4019, 2020.

[11] C. Xu, Y. Zhao, B. Qin, and H. Zhang, "Adaptive synchronization of coupled harmonic oscillators under switching topology," *Journal of the Franklin Institute*, vol. 356, no. 2, pp. 1067–1087, 2019.

[12] Y. Liu and H. Su, "Some necessary and sufficient conditions for containment of second-order multi-agent systems with sampled position data," *Neurocomputing*, vol. 378, pp. 228–237, 2020.

[13] Y. Liu and H. Su, "Containment control of second-order multi-agent systems via intermittent sampled position data communication," *Applied Mathematics and Computation*, vol. 362, Article ID 124522, 2019.

[14] B. Cheng and Z. Li, "Fully distributed event-triggered protocols for linear multiagent networks," *IEEE Transactions on Automatic Control*, vol. 64, no. 4, pp. 1655–1662, 2019.

[15] S. Chen, H. Pei, Q. Lai, and H. Yan, "Multitarget tracking control for coupled heterogeneous inertial agents systems based on flocking behavior," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 12, pp. 2605–2611, 2019.

[16] X. Wang and H. Su, "Completely model-free RL-based consensus of continuous-time multi-agent systems," *Applied Mathematics and Computation*, vol. 382, Article ID 125312, 2020.

[17] C. Altafini, "Consensus problems on networks with antagonistic interactions," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 935–946, 2013.
[18] L. Tian, Z. Ji, T. Hou, and K. Liu, "Bipartite consensus on coopetition networks with time-varying delays," *IEEE Access*, vol. 6, no. 1, pp. 10169–10178, 2018.
[19] L. Tian, Z. J. Ji, T. Hou, and H. Xu, "Bipartite consensus of edge dynamics on coopetition multi-agent systems," *Science China: Information Sciences*, vol. 62, no. 12, Article ID 229201, 2019.
[20] J. Qin, W. Fu, W. Zheng, and H. Gao, "On the bipartite consensus for generic linear multiagent systems with input saturation," *IEEE Transactions on Cybernetics*, vol. 47, no. 8, pp. 1948–1958, 2016.
[21] J. Hu, Y. Wu, T. Li, and B. K. Ghosh, "Consensus control of general linear multiagent systems with antagonistic interactions and communication noises," *IEEE Transactions on Automatic Control*, vol. 64, no. 5, pp. 2122–2127, 2019.
[22] J. Li, X. Chen, F. Hao, and J. Xie, "Event-triggered bipartite consensus for multi-agent systems with antagonistic interactions," *International Journal of Control, Automation and Systems*, vol. 17, pp. 1–13, 2019.
[23] H. Zhang, Y. Cai, Y. Wang, and H. Su, "Adaptive bipartite event-triggered output consensus of heterogeneous linear multiagent systems under fixed and switching topologies," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 11, pp. 4816–4830, 2020.
[24] D. Meng, Y. Jia, and J. Du, "Finite-time consensus for multiagent systems with cooperative and antagonistic interactions," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 4, pp. 762–770, 2016.
[25] X. Wang, G.-P. Jiang, H. Su, and X. Wang, "Robust global coordination of networked systems with input saturation and external disturbances," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–13, 2020.
[26] L. Guo and W.-H. Chen, "Disturbance attenuation and rejection for systems with nonlinearity via DOBC approach," *International Journal of Robust and Nonlinear Control*, vol. 15, no. 3, pp. 109–125, 2005.
[27] H. Yang, Z. Zhang, and S. Zhang, "Consensus of second-order multi-agent systems with exogenous disturbances," *International Journal of Robust and Nonlinear Control*, vol. 21, no. 9, pp. 945–956, 2011.
[28] C. Xu, Y. Zheng, H. Su, and H.-B. Zeng, "Containment for linear multi-agent systems with exogenous disturbances," *Neurocomputing*, vol. 160, pp. 206–212, 2015.
[29] C. Xu, H. Xu, H. Su, and C. Liu, "Disturbance-observer based consensus of linear multi-agent systems with exogenous disturbance under intermittent communication," *Neurocomputing*, vol. 404, pp. 26–33, 2020.
[30] Y. Wang, Z. Wang, and L. Kong, "Robust synchronization of class chaotic systems using novel time-varying gain disturbance observer-based sliding mode control," *Complexity*, vol. 2021, Article ID 8845553, 14 pages, 2021.
[31] M. Chen, S. Xiong, and Q. Wu, "Tracking flight control of quadrotor based on disturbance observer," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 3, pp. 1414–1423, 2021.
[32] H. Xu, C. Liu, Y. Lv, and J. Zhou, "Adaptive bipartite consensus of second-order multi-agent systems with bounded disturbances," *IEEE Access*, vol. 8, pp. 186441–186447, 2020.
[33] Z. Lin, *Low Gain Feedback, Lecture Notes in Control and Information Sciences*, Springer, London, UK, 1998.
[34] S. Boyd, L. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM, Philadelphia, PA, USA, 1994.

Hindawi
Complexity
Volume 2021, Article ID 5528089, 8 pages
https://doi.org/10.1155/2021/5528089

WILEY | Hindawi

Research Article

Hypo-EP Matrices of Adjointable Operators on Hilbert C^*-Modules

Xiaopeng Li, Junjie Huang, and Alatancang Chen

School of Mathematical Sciences, Inner Mongolia University, Hohhot 010021, China

Correspondence should be addressed to Junjie Huang; huangjunjie@rocketmail.com

Received 4 February 2021; Accepted 27 April 2021; Published 25 May 2021

Academic Editor: Hou-Sheng Su

This paper introduces and studies hypo-EP matrices of adjointable operators on Hilbert C^*-modules, based on the generalized Schur complement. The necessary and sufficient conditions for some modular operator matrices to be hypo-EP are given, and some special circumstances are also analyzed. Furthermore, an application of the EP operator in operator equations is given.

1. Introduction and Preliminaries

The EP matrix, as an extension of the normal matrix, was proposed by Schwerdtfeger; a square matrix T over the complex field $\mathbb{C}$ is said to be an EP matrix if T and T^* share the same range [1, 2]. The notion of EP matrices was extended by Campbell and Meyer to operators with closed range on a Hilbert space in [3]. Let H be a complex Hilbert space and $\mathscr{B}(H)$ the collection of all bounded linear operators on H. Let $T \in \mathscr{B}(H)$. Recall that T is called an EP operator if its range, $\mathscr{R}(T)$, is closed, and $\mathscr{R}(T) = \mathscr{R}(T^*)$ [3]. It is well known that $\mathscr{R}(T)$ is closed if and only if the Moore–Penrose inverse $T^\dagger$ of T exists and that T is an EP operator if and only if $T^\dagger T = TT^\dagger$. Sharifi [4] provided a generalization of the result for EP operators on Hilbert C^*-modules. This has been studied by many other authors, see, e.g., [5–8] and references therein. More generally, T is said to be a hypo-EP operator if $T^\dagger T \geq TT^\dagger$ [9]. In fact, T is a hypo-EP operator if and only if $\mathscr{R}(T)$ is closed and $\mathscr{R}(T) \subseteq \mathscr{R}(T^*)$. It is also shown that T is a hypo-EP operator if and only if $T^\dagger T^2 T^\dagger = TT^\dagger$. The hypo-EP operator is our focus of attention in this paper, and it has been studied in [10, 11]. The EP operator can be applied to the solution of operator equations, see Section 3 of this article. The properties of hypo-EP and EP operators can find applications also in reverse order law [12] and core partial order [13] and will be useful in some other applied fields [14, 15]. In this note, we investigate the hypo-EP operators on Hilbert C^*-modules, and then we formulate some results of hypo-EP matrices of adjointable operators on Hilbert C^*-modules. As an application, the solvability conditions, and the general expression for the EP solution to the operator equations are given.

Since the finite-dimensional spaces, Hilbert spaces, and C^*-algebras can all be regarded as Hilbert C^*-modules, one can study hypo-EP modular operators in a unified way in the framework of Hilbert C^*-modules. Let us briefly recall some basic knowledge about Hilbert C^*-modules and adjointable operators. Throughout this paper, $\mathbb{A}$ is a C^*-algebra. A Hilbert $\mathbb{A}$-module $\mathscr{H}$ is a right $\mathbb{A}$-module equipped with an $\mathbb{A}$-valued inner product $\langle \cdot, \cdot \rangle : \mathscr{H} \times \mathscr{H} \longrightarrow \mathbb{A}$ such that $\mathscr{H}$ is complete with respect to the induced norm $\| \cdot \| = \| \langle \cdot, \cdot \rangle \|^{(1/2)}$. Suppose that $\mathscr{H}$ and $\mathscr{K}$ are Hilbert $\mathbb{A}$-modules, and let $\mathscr{L}(\mathscr{H}, \mathscr{K})$ be the set of all maps $T : \mathscr{H} \longrightarrow \mathscr{K}$ for which there is a map $T^* : \mathscr{K} \longrightarrow \mathscr{H}$ such that $\langle Tx, y \rangle = \langle x, T^* y \rangle$ for $x \in \mathscr{H}$ and $y \in \mathscr{K}$. It is well known that an arbitrary element T of $\mathscr{L}(\mathscr{H}, \mathscr{K})$ must be a bounded linear operator, which is also $\mathbb{A}$-linear in the sense of $T(xa) = (Tx)a$ for any $x \in \mathscr{H}$ and $a \in \mathbb{A}$. We call $\mathscr{L}(\mathscr{H}, \mathscr{K})$ the set of adjointable operators from $\mathscr{H}$ to $\mathscr{K}$. We use $\mathscr{L}(\mathscr{H})$ to denote the C^*-algebra $\mathscr{L}(\mathscr{H}, \mathscr{H})$. Let $\mathscr{L}(\mathscr{H})_{sa}$ be the set of Hermitian elements of $\mathscr{L}(\mathscr{H})$. For $T \in \mathscr{L}(\mathscr{H}, \mathscr{K})$, the range and the null space of T are denoted by $\mathscr{R}(T)$ and $\mathscr{N}(T)$, respectively. An operator

$T \in \mathscr{L}(\mathscr{H}, \mathscr{K})$ is said to be regular if there is an operator $T^- \in \mathscr{L}(\mathscr{H}, \mathscr{K})$ satisfying $TT^-T = T$; T^- is called an inner inverse (or $\{1\}$-inverse) of T. It is easy to prove that T is regular if and only if $\mathscr{R}(T)$ is closed. The $\{1\}$-inverse of T is not unique in general.

In this paper, we use the generalized inverse to the generalized Schur complement as defined in [16]. Suppose $M \in \mathscr{L}(\mathscr{H} \oplus \mathscr{K})$ is a modular operator matrix of the form

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{1}$$

where $A \in \mathscr{L}(\mathscr{H}), B \in \mathscr{L}(\mathscr{K}, \mathscr{H}), C \in \mathscr{L}(\mathscr{H}, \mathscr{K}), D \in \mathscr{L}(\mathscr{K})$. Then, the generalized Schur complement of A in M is

$$\frac{M}{A} = D - CA^- B, \tag{2}$$

where A^- is an inner inverse of A. Similarly, the generalized Schur complement of D in M is

$$\frac{M}{D} = A - BD^- C, \tag{3}$$

where D^- is an inner inverse of D. The formulas (2) and (3) have previously appeared in papers dealing with generalized inverses of partitioned matrices (cf. [17–19]).

Definition 1 (see [20]). Let $T \in \mathscr{L}(\mathscr{H}, \mathscr{K})$. The Moore–Penrose inverse $T^\dagger$ of T (if exists) is an element in $\mathscr{L}(\mathscr{K}, \mathscr{H})$ which satisfies

(a) $TT^\dagger T = T$

(b) $T^\dagger TT^\dagger = T^\dagger$

(c) $(TT^\dagger)^* = TT^\dagger$

(d) $(T^\dagger T)^* = T^\dagger T$

These equations imply that $T^\dagger$ will be uniquely determined if it exists, and $T^\dagger T$ and $TT^\dagger$ are both orthogonal projections. Moreover, $\mathscr{R}(T^\dagger) = \mathscr{R}(T^\dagger T)$, $\mathscr{R}(T) = \mathscr{R}(TT^\dagger)$, $\mathscr{N}(T) = \mathscr{N}(T^\dagger T)$, and $\mathscr{N}(T^\dagger) = \mathscr{N}(TT^\dagger)$. Clearly, the Moore–Penrose inverse $T^\dagger$ of T exists if and only if $\mathscr{R}(T)$ is closed; T is Moore–Penrose invertible if and only if T^* is Moore–Penrose invertible, and in this case, $(T^*)^\dagger = (T^\dagger)^*$. Obviously, the Moore–Penrose inverse $T^\dagger$ of T is one of inner inverses of T.

Similar to [21], Lemma 2.2.4, and [22], Lemma 2.2, we have the following conclusions on Hilbert C^*-modules.

Lemma 1. *Let $A \in \mathscr{L}(\mathscr{H}), B \in \mathscr{L}(\mathscr{K}, \mathscr{H})$ and $C \in \mathscr{L}(\mathscr{H}, \mathscr{K})$. If A has an inner inverse A^-, then*

(i) $\mathscr{N}(A) \subseteq \mathscr{N}(C)$ if and only if $C = CA^- A$

(ii) $\mathscr{N}(A^*) \subseteq \mathscr{N}(B^*)$ if and only if $B = AA^- B$

Lemma 2. *Let M be a modular operator matrix of form (1) with $\mathscr{N}(A) \subseteq \mathscr{N}(C)$ and $\mathscr{N}(A^*) \subseteq \mathscr{N}(B^*)$. If A has an inner inverse A^-, then M is regular if and only if M/A is regular, where $M/A = D - CA^- B$. In this case, the inner inverse of M is given by*

$$M^- = \begin{pmatrix} A^- + A^- B\left(\frac{M}{A}\right)^- CA^- & -A^- B\left(\frac{M}{A}\right)^- \\ -\left(\frac{M}{A}\right)^- CA^- & \left(\frac{M}{A}\right)^- \end{pmatrix}. \tag{4}$$

From Lemma 2, we can obtain the following corollary.

Corollary 1. *Let M be a modular operator matrix of form (1) with $\mathscr{N}(A) \subseteq \mathscr{N}(C)$, $\mathscr{N}(A^*) \subseteq \mathscr{N}(B^*)$, $\mathscr{N}(M/A) \subseteq \mathscr{N}(B)$, and $\mathscr{N}((M/A)^*) \subseteq \mathscr{N}(C^*)$. If $\mathscr{R}(A)$ and $\mathscr{R}(M/A)$ are closed, then the Moore–Penrose inverse $M^\dagger$ of M can be expressed as*

$$M^\dagger = \begin{pmatrix} A^\dagger + A^\dagger B\left(\frac{M}{A}\right)^\dagger CA^\dagger & -A^\dagger B\left(\frac{M}{A}\right)^\dagger \\ -\left(\frac{M}{A}\right)^\dagger CA^\dagger & \left(\frac{M}{A}\right)^\dagger \end{pmatrix}. \tag{5}$$

Remark 1. The preceding result given in [17], Theorem 1, was proved for finite matrices.

Lemma 3 (see [23]). *Let $M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \in \mathscr{L}(\mathscr{H} \oplus \mathscr{K})$, where $A \in \mathscr{L}(\mathscr{H}), B \in \mathscr{L}(\mathscr{K}, \mathscr{H})$, and $D \in \mathscr{L}(\mathscr{K})$. If $\mathscr{R}(A)$ and $\mathscr{R}(D)$ are closed, then $M^\dagger = \begin{pmatrix} A^\dagger & -A^\dagger BD^\dagger \\ 0 & D^\dagger \end{pmatrix}$ if and only if $\mathscr{N}(D) \subseteq \mathscr{N}(B)$ and $\mathscr{N}(A^*) \subseteq \mathscr{N}(B^*)$.*

Proof. The proof is similar to that in [22], Corollary 12, for Hilbert space operators. □

Definition 2 (see [4]). Let $\mathscr{H}$ be a Hilbert $\mathbb{A}$-module. An operator $T \in \mathscr{L}(\mathscr{H})$ is called EP if $\overline{\mathscr{R}(T)} = \overline{\mathscr{R}(T^*)}$.

Definition 3. Let $\mathscr{H}$ be a Hilbert $\mathbb{A}$-module. An operator $T \in \mathscr{L}(\mathscr{H})$ is called hypo-EP if $\overline{\mathscr{R}(T)} \subseteq \overline{\mathscr{R}(T^*)}$.

Obviously, the range of an EP or a hypo-EP operator on Hilbert C^*-modules is not necessarily closed, and we further have the following properties.

Proposition 1 (see [4]). *Let $\mathscr{H}$ be a Hilbert $\mathbb{A}$-module and $T \in \mathscr{L}(\mathscr{H})$ with closed range. Then, the following conditions are equivalent:*

(i) T is an EP operator

(ii) $\mathscr{N}(T) = \mathscr{N}(T^*)$

(iii) T is Moore–Penrose invertible and $T^\dagger T = TT^\dagger$

Proposition 2. *Let $\mathscr{H}$ be a Hilbert $\mathbb{A}$-module and $T \in \mathscr{L}(\mathscr{H})$ with closed range. Then, the following conditions are equivalent:*

(i) T is a hypo-EP operator

(ii) $\mathscr{N}(T) \subseteq \mathscr{N}(T^*)$

(iii) T is Moore–Penrose invertible and $T^{\dagger}T^{2}T^{\dagger} = TT^{\dagger}$

Remark 2. The class of all hypo-EP operators contains the class of all EP operators on Hilbert $\mathbb{A}$-modules. Meanwhile, the EP operator with closed range is an extension of the invertible operator and the normal operator with closed range. In the case of finite dimensional situation, EP and hypo-EP are the same.

2. Main Results and Proofs

First, using generalized Schur complements, we study the hypo-EP property of matrices of adjointable operators on Hilbert C^{*}-modules.

Theorem 1. *Let M be a modular operator matrix of the form (1) with $\mathcal{N}(A) \subseteq \mathcal{N}(C)$, $\mathcal{N}(A^{*}) \subseteq \mathcal{N}(B^{*})$, $\mathcal{N}(M/A) \subseteq \mathcal{N}(B)$, and $\mathcal{N}((M/A)^{*}) \subseteq \mathcal{N}(C^{*})$. Suppose that $\mathcal{R}(A)$ and $\mathcal{R}(M/A)$ are closed. Then, the following conditions are equivalent:*

(i) M is a hypo-EP operator matrix with closed range

(ii) A and M/A are hypo-EP operators

Proof. Let M be a hypo-EP operator matrix with closed range. Since $\mathcal{R}(A)$ and $\mathcal{R}(M/A)$ are closed, let us consider the operator matrices

$$L = \begin{pmatrix} I & 0 \\ CA^{-} & I \end{pmatrix},$$

$$R = \begin{pmatrix} I & B\left(\frac{M}{A}\right)^{-} \\ 0 & I \end{pmatrix}, \tag{6}$$

$$P = \begin{pmatrix} A & 0 \\ 0 & \frac{M}{A} \end{pmatrix},$$

where $M/A = D - CA^{-}B$. Obviously, L and R are invertible. By using Lemma 1 and by assumptions $\mathcal{N}(A) \subseteq \mathcal{N}(C)$ and $\mathcal{N}(M/A) \subseteq \mathcal{N}(B)$, it is clear that M can be factorized as $M = LRP$. Hence, $\mathcal{N}(P) = \mathcal{N}(M)$. Since M is a hypo-EP operator matrix with closed range, $\mathcal{N}(P) = \mathcal{N}(M) \subseteq \mathcal{N}(M^{*})$. By using Lemma 1 again, it is immediate that

$$M^{*} = M^{*}P^{-}P \tag{7}$$

holds for every inner inverse P^{-} of P. In particular, for

$$P^{-} = \begin{pmatrix} A^{-} & 0 \\ 0 & \left(\frac{M}{A}\right)^{-} \end{pmatrix}, \tag{8}$$

we have from relation (7) that

$$M^{*} = \begin{pmatrix} A^{*} & C^{*} \\ B^{*} & D^{*} \end{pmatrix} = \begin{pmatrix} A^{*} & C^{*} \\ B^{*} & D^{*} \end{pmatrix} \begin{pmatrix} A^{-}A & 0 \\ 0 & \left(\frac{M}{A}\right)^{-}\left(\frac{M}{A}\right) \end{pmatrix},$$

$$= \begin{pmatrix} A^{*}A^{-}A & C^{*}\left(\frac{M}{A}\right)^{-}\left(\frac{M}{A}\right) \\ B^{*}A^{-}A & D^{*}\left(\frac{M}{A}\right)^{-}\left(\frac{M}{A}\right) \end{pmatrix}. \tag{9}$$

Then, $A^{*} = A^{*}A^{-}A$ implies $\mathcal{N}(A) \subseteq \mathcal{N}(A^{*})$. Hence, A is a hypo-EP operator. Since $C^{*} = C^{*}(M/A)^{-}(M/A)$, substituting $D = (M/A) + CA^{-}B$ into

$$D^{*} = D^{*}(M/A)^{-}(M/A) \tag{10}$$

yields $(M/A)^{*} = (M/A)^{*}(M/A)^{-}(M/A)$. This implies $\mathcal{N}(M/A) \subseteq \mathcal{N}((M/A)^{*})$. Thus, M/A is a hypo-EP operator.

Conversely, according to the assumptions $\mathcal{N}(A) \subseteq \mathcal{N}(C)$, $\mathcal{N}(M/A) \subseteq \mathcal{N}(B)$, $\mathcal{N}(A^{*}) \subseteq \mathcal{N}(B^{*})$, and $\mathcal{N}((M/A)^{*}) \subseteq \mathcal{N}(C^{*})$, the Moore–Penrose inverse $M^{\dagger}$ of M exists, and $M^{\dagger}$ is given by

$$M^{\dagger} = \begin{pmatrix} A^{\dagger} + A^{\dagger}B\left(\frac{M}{A}\right)^{\dagger}CA^{\dagger} & -A^{\dagger}B\left(\frac{M}{A}\right)^{\dagger} \\ -\left(\frac{M}{A}\right)^{\dagger}CA^{\dagger} & \left(\frac{M}{A}\right)^{\dagger} \end{pmatrix}, \tag{11}$$

by Corollary 1. Using $\mathcal{N}(A^{*}) \subseteq \mathcal{N}(B^{*})$ and $\mathcal{N}((M/A)^{*}) \subseteq \mathcal{N}(C^{*})$, by Lemma 1, $MM^{\dagger}$ is described as

$$MM^{\dagger} = \begin{pmatrix} AA^{\dagger} & 0 \\ 0 & \left(\frac{M}{A}\right)\left(\frac{M}{A}\right)^{\dagger} \end{pmatrix}. \tag{12}$$

Similarly, by using $\mathcal{N}(A) \subseteq \mathcal{N}(C)$, $\mathcal{N}(M/A) \subseteq \mathcal{N}(B)$, and Lemma 1, it is given that

$$M^{\dagger}M = \begin{pmatrix} A^{\dagger}A & 0 \\ 0 & \left(\frac{M}{A}\right)^{\dagger}\left(\frac{M}{A}\right) \end{pmatrix}. \tag{13}$$

Then,

$$M^{\dagger}M^{2}M^{\dagger} = \left(M^{\dagger}M\right)\left(MM^{\dagger}\right) = \begin{pmatrix} A^{\dagger}A^{2}A^{\dagger} & 0 \\ 0 & \left(\frac{M}{A}\right)^{\dagger}\left(\frac{M}{A}\right)^{2}\left(\frac{M}{A}\right)^{\dagger} \end{pmatrix}. \tag{14}$$

Since A and M/A are hypo-EP operators with closed range,

$$A^{\dagger}A^{2}A^{\dagger} = AA^{\dagger},$$
$$\left(\frac{M}{A}\right)^{\dagger}\left(\frac{M}{A}\right)^{2}\left(\frac{M}{A}\right)^{\dagger} = \left(\frac{M}{A}\right)\left(\frac{M}{A}\right)^{\dagger}. \tag{15}$$

Thus, $M^{\dagger}M^{2}M^{\dagger} = MM^{\dagger}$. Therefore, M is a hypo-EP operator matrix with closed range.

The following conclusion is a natural extension of [10], Theorem 3.1, on Hilbert C^{*}-modules. □

Corollary 2. *Let $A, X \in \mathscr{L}(\mathscr{H})$ and $M = \begin{pmatrix} A & AX \\ X^{*}A & X^{*}AX \end{pmatrix} \in \mathscr{L}(\mathscr{H} \oplus \mathscr{H})$. If $\mathscr{R}(A)$ is closed, then M is a hypo-EP operator matrix with closed range if and only if A is a hypo-EP operator.*

If using the generalized Schur complement $M/D = A - BD^{-}C$ of D in M, similar to Theorem 1, one can get the following results.

Theorem 2. *Let M be a modular operator matrix of form (1) with $\mathscr{N}(D) \subseteq \mathscr{N}(B)$, $\mathscr{N}(D^{*}) \subseteq \mathscr{N}(C^{*})$, $\mathscr{N}(M/D) \subseteq \mathscr{N}(C)$, and $\mathscr{N}((M/D)^{*}) \subseteq \mathscr{N}(B^{*})$. Suppose that $\mathscr{R}(D)$ and $\mathscr{R}(M/D)$ are closed. Then, the following conditions are equivalent:*

(i) M is a hypo-EP operator matrix with closed range

(ii) D and M/D are hypo-EP operators

Corollary 3. *Let $D, X \in \mathscr{L}(\mathscr{H})$ and $M = \begin{pmatrix} X^{*}DX & X^{*}D \\ DX & D \end{pmatrix} \in \mathscr{L}(\mathscr{H} \oplus \mathscr{H})$. If $\mathscr{R}(D)$ is closed, then M is a hypo-EP operator matrix with closed range if and only if D is a hypo-EP operator.*

Next, using the properties of generalized inverses, we study upper triangular hypo-EP matrices of adjointable operators on Hilbert C^{*}-modules.

Theorem 3. *Let $M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \in \mathscr{L}(\mathscr{H} \oplus \mathscr{K})$ with $\mathscr{N}(D) \subseteq \mathscr{N}(B)$ and $\mathscr{N}(A^{*}) \subseteq \mathscr{N}(B^{*})$, where $A \in \mathscr{L}(\mathscr{H}), B \in \mathscr{L}(\mathscr{K}, \mathscr{H})$, and $D \in \mathscr{L}(\mathscr{K})$. If $\mathscr{R}(A)$ and $\mathscr{R}(D)$ are closed, then M is a hypo-EP operator matrix with closed range if and only if A and D are hypo-EP operators.*

Proof. Let M be a hypo-EP operator matrix with closed range. We write

$$L := \begin{pmatrix} I & BD^{\dagger} \\ 0 & I \end{pmatrix},$$
$$P := \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}. \tag{16}$$

Obviously, L is invertible. By Lemma 1 and assumption $\mathscr{N}(D) \subseteq \mathscr{N}(B)$, it is clear that M can be decomposed as $M = LP$. Hence, $\mathscr{N}(M) = \mathscr{N}(P)$. Since M is a hypo-EP operator matrix with closed range, $\mathscr{N}(P) = \mathscr{N}(M) \subseteq \mathscr{N}(M^{*})$. By Lemma 1, it is immediate that $M^{*} = M^{*}P^{\dagger}P$, where $P^{\dagger}$ is given by

$$P^{\dagger} = \begin{pmatrix} A^{\dagger} & 0 \\ 0 & D^{\dagger} \end{pmatrix}. \tag{17}$$

This gives

$$M^{*} = \begin{pmatrix} A^{*} & 0 \\ B^{*} & D^{*} \end{pmatrix} = \begin{pmatrix} A^{*} & 0 \\ B^{*} & D^{*} \end{pmatrix}\begin{pmatrix} A^{\dagger}A & 0 \\ 0 & D^{\dagger}D \end{pmatrix}, \quad = \begin{pmatrix} A^{*}A^{\dagger}A & 0 \\ B^{*}A^{\dagger}A & D^{*}D^{\dagger}D \end{pmatrix}. \tag{18}$$

Hence, $A^{*} = A^{*}A^{\dagger}A$ implies $\mathscr{N}(A) \subseteq \mathscr{N}(A^{*})$. Thus, A is a hypo-EP operator. From $D^{*} = D^{*}D^{\dagger}D$, it follows that $\mathscr{N}(D) \subseteq \mathscr{N}(D^{*})$. Therefore, D is a hypo-EP operator.

Conversely, suppose A and D are hypo-EP operators. Since $\mathscr{R}(A)$ and $\mathscr{R}(D)$ are closed and $\mathscr{N}(D) \subseteq \mathscr{N}(B)$ and $\mathscr{N}(A^{*}) \subseteq \mathscr{N}(B^{*})$, by Lemma 3, the Moore–Penrose inverse $M^{\dagger}$ of M exists and

$$M^{\dagger} = \begin{pmatrix} A^{\dagger} & -A^{\dagger}BD^{\dagger} \\ 0 & D^{\dagger} \end{pmatrix}. \tag{19}$$

Since $\mathscr{N}(A^{*}) \subseteq \mathscr{N}(B^{*})$, by Lemma 1, $MM^{\dagger}$ is described as

$$MM^{\dagger} = \begin{pmatrix} AA^{\dagger} & 0 \\ 0 & DD^{\dagger} \end{pmatrix}. \tag{20}$$

Similarly, by Lemma 1, $\mathscr{N}(D) \subseteq \mathscr{N}(B)$ leads to

$$M^{\dagger}M = \begin{pmatrix} A^{\dagger}A & 0 \\ 0 & D^{\dagger}D \end{pmatrix}. \tag{21}$$

Then,

$$M^{\dagger}M^{2}M^{\dagger} = \left(M^{\dagger}M\right)\left(MM^{\dagger}\right) = \begin{pmatrix} A^{\dagger}A^{2}A^{\dagger} & 0 \\ 0 & D^{\dagger}D^{2}D^{\dagger} \end{pmatrix}. \tag{22}$$

Since A and D are hypo-EP operators with closed range, $A^{\dagger}A^{2}A^{\dagger} = AA^{\dagger}$ and $D^{\dagger}D^{2}D^{\dagger} = DD^{\dagger}$. Thus, $M^{\dagger}M^{2}M^{\dagger} = MM^{\dagger}$. Therefore M is a hypo-EP operator matrix with closed range. □

Corollary 4. *Let $M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \in \mathscr{L}(\mathscr{H} \oplus \mathscr{K})$ with $\mathscr{N}(D) \subseteq \mathscr{N}(B)$ and $\mathscr{N}(A^{*}) \subseteq \mathscr{N}(B^{*})$, where $A \in \mathscr{L}(\mathscr{H})$, $B \in \mathscr{L}(\mathscr{K}, \mathscr{H})$, and $D \in \mathscr{L}(\mathscr{K})$. If $\mathscr{R}(A)$ and $\mathscr{R}(D)$ are closed, then M is an EP operator matrix with closed range if and only if A and D are EP operators.*

Proof. Let M be an EP operator matrix with closed range. In view of Theorem 3, to prove the necessity, it is enough to show $\mathscr{N}(A^{*}) \subseteq \mathscr{N}(A)$ and $\mathscr{N}(D^{*}) \subseteq \mathscr{N}(D)$. Since M is an EP operator matrix with closed range, by the proof of Theorem 3, $\mathscr{N}(P) = \mathscr{N}(M) = \mathscr{N}(M^{*})$. Applying Lemma 1, we have $P^{*} = MM^{\dagger}P^{*}$, i.e.,

$$P^* = \begin{pmatrix} A^* & 0 \\ 0 & D^* \end{pmatrix} = \begin{pmatrix} AA^\dagger & 0 \\ 0 & DD^\dagger \end{pmatrix} \begin{pmatrix} A^* & 0 \\ 0 & D^* \end{pmatrix}, \\ = \begin{pmatrix} AA^\dagger A^* & 0 \\ 0 & DD^\dagger D^* \end{pmatrix}. \tag{23}$$

Hence, $A^* = AA^\dagger A^*$ and $D^* = DD^\dagger D^*$ imply $\mathcal{N}(A^*) \subseteq \mathcal{N}(A)$ and $\mathcal{N}(D^*) \subseteq \mathcal{N}(D)$, respectively.

Conversely, let A and D be EP operators. Since $\mathcal{R}(A)$and $\mathcal{R}(D)$ are closed, $\mathcal{N}(D) \subseteq \mathcal{N}(B)$, and $\mathcal{N}(A^*) \subseteq \mathcal{N}(B^*)$, we get $M^\dagger = \begin{pmatrix} A^\dagger & -A^\dagger B D^\dagger \\ 0 & D^\dagger \end{pmatrix}$. Thus,

$$M^\dagger M - MM^\dagger = \begin{pmatrix} A^\dagger A - AA^\dagger & 0 \\ 0 & D^\dagger D - DD^\dagger \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{24}$$

Therefore, M is an EP operator matrix with closed range. □

Corollary 5. *Let $A, X, D \in \mathcal{L}(\mathcal{H})$ and $M = \begin{pmatrix} A & AXD \\ 0 & D \end{pmatrix} \in \mathcal{L}(\mathcal{H} \oplus \mathcal{H})$. If $\mathcal{R}(A)$ and $\mathcal{R}(D)$ are closed, then M is a hypo-EP operator matrix with closed range if and only if A and D are hypo-EP operators.*

Corollary 6. *Let $A \in \mathcal{L}(\mathcal{H})$ and $M = \begin{pmatrix} A & A \\ 0 & A \end{pmatrix} \in \mathcal{L}(\mathcal{H} \oplus \mathcal{H})$. If $\mathcal{R}(A)$ is closed, then M is a hypo-EP operator matrix with closed range if and only if A is a hypo-EP operator.*

Remark 3. The Hilbert space version of the preceding four conclusions is given by [10], and the conditions of closed range can be naturally omitted there. Moreover, the alternative proofs of the conclusions in Hilbert space setting can be found in section 3 of [10]. In addition, these results originated from the research of the EP property of block matrices, according to Hartwig [24].

Finally, the following are devoted to investigating the hypo-EP property of antitriangular block matrices of adjointable operators on Hilbert C*-modules.

Lemma 4. *Let $M = \begin{pmatrix} A & B \\ C & 0 \end{pmatrix} \in \mathcal{L}(\mathcal{H} \oplus \mathcal{H})$. If $\mathcal{R}(B)$ and $\mathcal{R}(C)$ are closed, then $M^\dagger = \begin{pmatrix} 0 & C^\dagger \\ B^\dagger & -B^\dagger AC^\dagger \end{pmatrix}$ if and only if $\mathcal{N}(C) \subseteq \mathcal{N}(A)$ and $\mathcal{N}(B^*) \subseteq \mathcal{N}(A^*)$.*

Proof. Sufficiency: since $\mathcal{R}(B)$ and $\mathcal{R}(C)$ are closed, B and C are Moore–Penrose invertible. From $\mathcal{N}(C) \subseteq \mathcal{N}(A)$ and $\mathcal{N}(B^*) \subseteq \mathcal{N}(A^*)$, it follows that $AC^\dagger C = A$ and $BB^\dagger A = A$. We write $X := \begin{pmatrix} 0 & C^\dagger \\ B^\dagger & -B^\dagger AC^\dagger \end{pmatrix}$. A direct calculation shows that

$$MXM = M, \\ XMX = X, \\ (MX)^* = MX, \\ (XM)^* = XM. \tag{25}$$

By Definition 1, $M^\dagger = X$ as desired.

Necessity: since

$$MM^\dagger = \begin{pmatrix} BB^\dagger & AC^\dagger - BB^\dagger AC^\dagger \\ 0 & CC^\dagger \end{pmatrix}, \\ M^\dagger M = \begin{pmatrix} C^\dagger C & 0 \\ B^\dagger A - B^\dagger AC^\dagger C & B^\dagger B \end{pmatrix} \tag{26}$$

are self-adjoint, we have $AC^\dagger - BB^\dagger AC^\dagger = 0$ and $B^\dagger A - B^\dagger AC^\dagger C = 0$. From $MM^\dagger M = M$, we get $BB^\dagger A = A = AC^\dagger C$. Therefore, $\mathcal{N}(C) \subseteq \mathcal{N}(A)$ and $\mathcal{N}(B^*) \subseteq \mathcal{N}(A^*)$. □

Lemma 5. *Let $M = \begin{pmatrix} A & B \\ C & 0 \end{pmatrix} \in \mathcal{L}(\mathcal{H} \oplus \mathcal{H})$ with $\mathcal{N}(C) \subseteq \mathcal{N}(A)$ and $\mathcal{N}(B^*) \subseteq \mathcal{N}(A^*)$. If B and C are hypo-EP operators with closed ranges, then M is a hypo-EP operator matrix with closed range.*

Proof. Since $\mathcal{R}(B)$ and $\mathcal{R}(C)$ are closed, $\mathcal{N}(C) \subseteq \mathcal{N}(A)$, and $\mathcal{N}(B^*) \subseteq \mathcal{N}(A^*)$, by Lemma 4, the Moore–Penrose inverse $M^\dagger$ of M is given by

$$M^\dagger = \begin{pmatrix} 0 & C^\dagger \\ B^\dagger & -B^\dagger AC^\dagger \end{pmatrix}. \tag{27}$$

Using $\mathcal{N}(C) \subseteq \mathcal{N}(A)$ and $\mathcal{N}(B^*) \subseteq \mathcal{N}(A^*)$, by Lemma 1, we have

$$MM^\dagger = \begin{pmatrix} BB^\dagger & 0 \\ 0 & CC^\dagger \end{pmatrix}, \\ M^\dagger M = \begin{pmatrix} C^\dagger C & 0 \\ 0 & B^\dagger B \end{pmatrix}. \tag{28}$$

By Definition 1, we have $\mathcal{R}(MM^\dagger) = \mathcal{R}(M)$ and $\mathcal{R}(M^\dagger M) = \mathcal{R}(M^\dagger) = \mathcal{R}(M^*)$. Since B and C are hypo-EP operators with closed ranges, $\mathcal{R}(BB^\dagger) = \mathcal{R}(B) \subseteq \mathcal{R}(B^*) = \mathcal{R}(B^\dagger B)$ and $\mathcal{R}(CC^\dagger) = \mathcal{R}(C) \subseteq \mathcal{R}(C^*) = \mathcal{R}(C^\dagger C)$. Then,

$$\mathcal{R}(M) = \mathcal{R}\left(MM^\dagger\right) = \mathcal{R}\left(BB^\dagger\right) \oplus \mathcal{R}\left(CC^\dagger\right) \subseteq \mathcal{R}\left(B^\dagger B\right) \oplus \mathcal{R}\left(C^\dagger C\right) = \mathcal{R}\left(M^\dagger M\right) = \mathcal{R}\left(M^*\right). \tag{29}$$

Therefore, M is a hypo-EP operator with closed range. □

Corollary 7. *Let* $M = \begin{pmatrix} A & B \\ C & 0 \end{pmatrix} \in \mathscr{L}(\mathscr{H} \oplus \mathscr{H})$ *with* $\mathscr{N}(C) \subseteq \mathscr{N}(A)$ *and* $\mathscr{N}(B^*) \subseteq \mathscr{N}(A^*)$. *If* B *and* C *are EP operators with closed ranges, then* M *is an EP operator matrix with closed range.*

Proof. According to the assumption, as with Lemma 5, we have equation (28). Since B and C are EP operators with closed ranges, $\mathscr{R}(BB^\dagger) = \mathscr{R}(B) = \mathscr{R}(B^*) = \mathscr{R}(B^\dagger B)$ and $\mathscr{R}(CC^\dagger) = \mathscr{R}(C) = \mathscr{R}(C^*) = \mathscr{R}(C^\dagger C)$. Then,

$$\mathscr{R}(M) = \mathscr{R}\left(MM^\dagger\right) = \mathscr{R}\left(BB^\dagger\right) \oplus \mathscr{R}\left(CC^\dagger\right) = \mathscr{R}\left(B^\dagger B\right) \oplus \mathscr{R}\left(C^\dagger C\right) = \mathscr{R}\left(M^\dagger M\right) = \mathscr{R}\left(M^*\right). \tag{30}$$

Thus, M is an EP operator matrix with closed range. □

Theorem 4. *Let* $M = \begin{pmatrix} A & B \\ C & 0 \end{pmatrix} \in \mathscr{L}(\mathscr{H} \oplus \mathscr{H})$ *with* $\mathscr{N}(B) = \mathscr{N}(C) \subseteq \mathscr{N}(A)$ *and* $\mathscr{N}(B^*) \subseteq \mathscr{N}(A^*)$. *If* $\mathscr{R}(B)$ *and* $\mathscr{R}(C)$*are closed, then* M *is a hypo-EP operator matrix with closed range if and only if* B *and* C *are hypo-EP operators.*

Proof. The sufficiency is clear by Lemma 5. Now, we suppose that M is a hypo-EP operator matrix with closed range. We write

$$\begin{aligned} L &\coloneqq \begin{pmatrix} I & AC^\dagger \\ 0 & I \end{pmatrix}, \\ P &\coloneqq \begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix}. \end{aligned} \tag{31}$$

In the similar way as in the proof of Theorem 3,

we have $M = LP$, and hence, $\mathscr{N}(P) = \mathscr{N}(M) \subseteq \mathscr{N}(M^*)$, since M is a hypo-EP with closed range. This means $M^* = M^*P^\dagger P$ by Lemma 1, i.e.,

$$\begin{pmatrix} A^* & C^* \\ B^* & 0 \end{pmatrix} = M^* = M^*P^\dagger P = \begin{pmatrix} A^*C^\dagger C & C^*B^\dagger B \\ B^*C^\dagger C & 0 \end{pmatrix}. \tag{32}$$

Hence, $C^* = C^*B^\dagger B$, which together with $\mathscr{N}(C) = \mathscr{N}(B)$, implies $\mathscr{N}(C) \subseteq \mathscr{N}(C^*)$. Thus, C is a hypo-EP operator. Similarly, it follows from $B^* = B^*C^\dagger C$ and $\mathscr{N}(B) = \mathscr{N}(C)$ that $\mathscr{N}(B) \subseteq \mathscr{N}(B^*)$, and therefore, B is a hypo-EP operator. □

Corollary 8. *Let* $B, X, C \in \mathscr{L}(\mathscr{H})$ *and* $M = \begin{pmatrix} BXC & B \\ C & 0 \end{pmatrix} \in \mathscr{L}(\mathscr{H} \oplus \mathscr{H})$ *with* $\mathscr{N}(B) = \mathscr{N}(C)$. *If* $\mathscr{R}(B)$ *and* $\mathscr{R}(C)$ *are closed, then* M *is a hypo-EP operator matrix with closed range if and only if* B *and* C *are hypo-EP operators.*

Corollary 9. *Let* $M = \begin{pmatrix} A & B \\ C & 0 \end{pmatrix} \in \mathscr{L}(\mathscr{H} \oplus \mathscr{H})$ *with* $\mathscr{N}(B) = \mathscr{N}(C) \subseteq \mathscr{N}(A)$ *and* $\mathscr{N}(B^*) \subseteq \mathscr{N}(A^*)$. *If* $\mathscr{R}(B)$ *and* $\mathscr{R}(C)$*are closed, then* M *is an EP operator matrix with closed range if and only if* B *and* C *are EP operators.*

Proof. By Corollary 7, we only need to show the necessity, which can be easily verified according to the proofs of Corollary 4 and Theorem 4. □

Corollary 10. *Let* $B, X, C \in \mathscr{L}(\mathscr{H})$ *and* $M = \begin{pmatrix} BXC & B \\ C & 0 \end{pmatrix} \in \mathscr{L}(\mathscr{H} \oplus \mathscr{H})$ *with* $\mathscr{N}(B) = \mathscr{N}(C)$. *If* $\mathscr{R}(B)$ *and* $\mathscr{R}(C)$ *are closed, then* M *is an EP operator matrix with closed range if and only if* B *and* C *are EP operators.*

Remark 4. In Hilbert space case, the conditions of closed range in Theorem 4 and Corollary 9 can be naturally omitted in Theorem 3.8 and Theorem 3.9 of [10], and the alternative proofs of Theorem 4 and Corollary 9 can be, respectively, found in Theorem 3.8 and Theorem 3.9 of [10].

3. The Application of EP Operators

In this section, let $\mathscr{H}, \mathscr{K}$, and $\mathscr{G}$ be Hilbert spaces. We establish the solvability conditions and the general expression for the EP solution to the operator equations

$$\begin{aligned} AX &= C, \\ XB &= D, \end{aligned} \tag{33}$$

where $A, C \in \mathscr{B}(\mathscr{H}, \mathscr{K})$, $B, D \in \mathscr{B}(\mathscr{G}, \mathscr{H})$, and $X \in \mathscr{B}(\mathscr{H})$.

Lemma 6 (see [25]). *Let* $T \in \mathscr{B}(\mathscr{H})$ *with closed range. Then, the operator* T *is EP if and only if there exist Hilbert spaces* $\mathscr{H}_1$ *and* $\mathscr{H}_2$, $U \in \mathscr{B}(\mathscr{H}_1 \oplus \mathscr{H}_2, \mathscr{H})$ *unitary, and* $T_1 \in \mathscr{B}(\mathscr{H}_1)$ *isomorphism such that*

$$T = U(T_1 \oplus 0)U^*, \tag{34}$$

where $\mathscr{H}_1 \oplus \mathscr{H}_2 = \mathscr{H}$.

Lemma 7 (see [22]). *Let* $A, C \in \mathscr{B}(\mathscr{H}, \mathscr{K})$ *and* $B, D \in \mathscr{B}(\mathscr{G}, \mathscr{H})$. *Suppose that* A *and* B *have closed ranges. Then, equation (33) has a common solution* $X \in \mathscr{B}(\mathscr{H})$ *if and only if*

$$\begin{aligned} &\mathscr{N}(A^*) \subseteq \mathscr{N}(C^*), \\ &\mathscr{N}(B) \subseteq \mathscr{N}(D), \\ &AD = CB. \end{aligned} \tag{35}$$

In this case, the general common solution is given by

$$X = A^{-}C + DB^{-} - A^{-}ADB^{-} + \left(I_{\mathscr{H}} - A^{-}A\right)Y\left(I_{\mathscr{H}} - BB^{-}\right), \tag{36}$$

where $Y \in \mathscr{B}(\mathscr{H})$ is arbitrary.

Now, we consider the EP solution to equation (33). By the Lemma 6, for the unitary operator $U \in \mathscr{B}(\mathscr{H}_1 \oplus \mathscr{H}_2, \mathscr{H})$, the solution has the following factorization:

$$X = U\begin{pmatrix} X_1 & 0 \\ 0 & 0 \end{pmatrix}U^{*}. \tag{37}$$

Let $\mathscr{R}(A)$, $\mathscr{R}(B)$ be closed, and

$$\begin{aligned} AU &= \begin{pmatrix} A_1 & A_2 \end{pmatrix}, \\ CU &= \begin{pmatrix} C_1 & C_2 \end{pmatrix}, \\ U^{*}B &= \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \\ U^{*}D &= \begin{pmatrix} D_1 \\ D_2 \end{pmatrix}, \end{aligned} \tag{38}$$

where $A_1, C_1 \in \mathscr{B}(\mathscr{H}_1, \mathscr{K})$, $A_2, C_2 \in \mathscr{B}(\mathscr{H}_2, \mathscr{K})$, $B_1, D_1 \in \mathscr{B}(\mathscr{G}_1, \mathscr{H})$, $B_2, D_2 \in \mathscr{B}(\mathscr{G}_2, \mathscr{H})$, and $\mathscr{R}(A_1)$, $\mathscr{R}(B_1)$ are closed. Then, equation (33) has an EP solution if and only if operator equations

$$\begin{aligned} A_1X_1 &= C_1, \\ X_1B_1 &= D_1, \\ C_2 &= 0, \\ D_2 &= 0 \end{aligned} \tag{39}$$

have a common solution. By Lemma 7, we have the following theorem.

Theorem 5. *Let $A, C \in \mathscr{B}(\mathscr{H}, \mathscr{K})$ and $B, D \in \mathscr{B}(\mathscr{G}, \mathscr{H})$, and let $\mathscr{R}(A)$, $\mathscr{R}(B)$ be closed. Suppose that $U \in \mathscr{B}(\mathscr{H}_1 \oplus \mathscr{H}_2, \mathscr{H})$ is unitary such that*

$$\begin{aligned} AU &= \begin{pmatrix} A_1 & A_2 \end{pmatrix}, \\ CU &= \begin{pmatrix} C_1 & C_2 \end{pmatrix}, \\ U^{*}B &= \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \\ U^{*}D &= \begin{pmatrix} D_1 \\ D_2 \end{pmatrix}, \end{aligned} \tag{40}$$

where $A_1, C_1 \in \mathscr{B}(\mathscr{H}_1, \mathscr{K})$, $A_2, C_2 \in \mathscr{B}(\mathscr{H}_2, \mathscr{K})$, $B_1, D_1 \in \mathscr{B}(\mathscr{G}_1, \mathscr{H})$, $B_2, D_2 \in \mathscr{B}(\mathscr{G}_2, \mathscr{H})$, and $\mathscr{R}(A_1)$ and $\mathscr{R}(B_1)$ are closed. Then, equation (33) has an EP solution $X \in \mathscr{B}(\mathscr{H})$ if and only if

$$\begin{aligned} \mathscr{N}(A_1^{*}) &\subseteq \mathscr{N}(C_1^{*}), \\ \mathscr{N}(B_1) &\subseteq \mathscr{N}(D_1), \\ A_1D_1 &= C_1B_1, \\ C_2 &= D_2 = 0. \end{aligned} \tag{41}$$

In this case, the general EP solution of (33) is given by

$$X = U\begin{pmatrix} A_1^{-}C_1 + D_1B_1^{-} - A_1^{-}A_1D_1B_1^{-} + \left(I_{\mathscr{H}_1} - A_1^{-}A_1\right)Y_1\left(I_{\mathscr{H}_1} - B_1B_1^{-}\right) & 0 \\ 0 & 0 \end{pmatrix}U^{*}, \tag{42}$$

where $Y_1 \in \mathscr{B}(\mathscr{H}_1)$ is arbitrary.

4. Concluding Remarks

In this work, we have characterized hypo-EP and EP matrices of adjointable operators on Hilbert C^{*}-modules, based on the generalized Schur complement, and an application of EP operator in operator equations is presented. In addition, the properties of hypo-EP and EP operators may have potential applications in some fields involving mathematics and its applications. In our opinion, it is worth establishing the hypo-EP and EP matrices of bounded linear operators on Krein space and of the adjointable operators on Krein C^{*}-modules.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the NNSF of China (Nos. 11961052 and 11761029) and the NSF of Inner Mongolia (No. 2017MS0118).

References

[1] I. J. Katz and M. H. Pearl, "On EPr and normal EPr matrices," *Journal of Research of the Notional Bureau of Standards–B. Mathematics and Mathematical Physics*, vol. 70, pp. 46–77, 1966.

[2] H. Schwerdtfeger, *Introduction to Linear Algebra and the Theory of Matrices*, P. Noordhoff, Groningen, Netherlands, 1950.

[3] S. L. Campbell and C. D. Meyer, "EP operators and generalized inverses," *Canadian Mathematical Bulletin*, vol. 18, no. 3, pp. 327–333, 1975.
[4] K. Sharifi, "EP modular operators and their products," *Journal of Mathematical Analysis and Applications*, vol. 419, no. 2, pp. 870–877, 2014.
[5] E. Boasso, "Factorizations of EP banach space operators and EP banach algebra elements," *Journal of Mathematical Analysis and Applications*, vol. 379, no. 1, pp. 245–255, 2011.
[6] D. S. Djordjević, "Product of EP operators on Hilbert spaces," *Proceedings of the American Mathematical Society*, vol. 129, pp. 1727–1731, 2000.
[7] R. E. Hartwig and I. J. Katz, "On products of EP matrices," *Linear Algebra and its Applications*, vol. 252, no. 1-3, pp. 339–345, 1997.
[8] D. Mosic, D. S. Djordjevic, and J. J. Koliha, "EP elements in rings," *Linear Algebra and its Applications*, vol. 431, pp. 527–535, 2009.
[9] M. Itoh, "On some EP operators," *Nihonkai Mathematical Journal*, vol. 16, pp. 49–56, 2005.
[10] A. B. Patel and M. P. Shekhawat, "Hypo-EP operators," *Indian Journal of Pure and Applied Mathematics*, vol. 47, no. 1, pp. 73–84, 2016.
[11] A. Vinoth and P. Sam Johnson, "On sum and restriction of hypo-EP operators," *Functional Analysis, Approximation and Computation*, vol. 9, no. 1, pp. 37–41, 2017.
[12] M. M. Karizaki and D. S. Djordjević, "Commuting C^* modular operators," *Aequationes Mathematicae*, vol. 90, pp. 1103C–1114, 2016.
[13] D. S. Rakiá, N. Č. Dinčić, and D. S. Djordjević, "Core inverse and core partial order of Hilbert space operators," *Applied Mathematics and Computation*, vol. 244, pp. 283–302, 2014.
[14] Y. Liu and H. Su, "Some necessary and sufficient conditions for containment of second-order multi-agent systems with sampled position data," *Neurocomputing*, vol. 378, pp. 228–237, 2020.
[15] Y. Liu and H. Su, "Containment control of second-order multi-agent systems via intermittent sampled position data communication," *Applied Mathematics and Computation*, vol. 362, p. 124522, 2019.
[16] F. Z. Zhang, *The Schur Complement and its Applications*, Springer-Verlag, Berlin, Germany, 2005.
[17] F. Burns, D. Carlson, E. Haynsworth, and T. Markham, "Generalized inverse formulas using the schur complement," *SIAM Journal on Applied Mathematics*, vol. 26, no. 2, pp. 254–259, 1974.
[18] A. Meenakshi, "On schur complements in an EP matrix," *Periodica Mathematica Hungarica*, vol. 16, no. 3, pp. 193–200, 1985.
[19] C. A. Rohde, "Generalized inverses of partitioned matrices," *Journal of the Society for Industrial and Applied Mathematics*, vol. 13, no. 4, pp. 1033–1035, 1965.
[20] M. M. Karizaki, M. Hassani, M. Amyari, and M. Khosravi, "Operator matrix of moore-penrose inverse operators on Hilbert C^*-modules," *Colloquium Mathematicum*, vol. 140, pp. 171–182, 2015.
[21] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and its Applications*, Wiley, Hoboken, NJ, USA, 1971.
[22] A. Dajić and J. J. Koliha, "Positive solutions to the equations AX=C and XB=D for Hilbert space operators," *Journal of Mathematical Analysis and Applications*, vol. 333, pp. 567–576, 2007.
[23] C. Y. Deng and H. K. Du, "Representations of the moore-penrose inverse of 2×2 block operator valued matrices," *Journal of the Korean Mathematical Society*, vol. 46, no. 6, pp. 1139–1150, 2009.
[24] R. E. Hartwig, "EP perturbations," *Sankhyā: The Indian Journal of Statistics, Series A*, vol. 56, pp. 347–357, 1994.
[25] D. Drivaliaris, S. Karanasios, and D. Pappas, "Factorizations of EP operators," *Linear Algebra and Its Applications*, vol. 429, no. 7, pp. 1555–1567, 2008.

Hindawi
Complexity
Volume 2021, Article ID 9971855, 9 pages
https://doi.org/10.1155/2021/9971855

WILEY | Hindawi

Research Article

Intermittent Sampled Data Control for Time-Varying Formation-Containment of the Multiagent System with/without Time Delay

Ming Chi,[1,2] Xu-Long Wang,[1,2] Yangming Dou,[1,2] and Zhi-Wei Liu[1,2]

[1]*School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430074, China*
[2]*Key Laboratory of Image Processing and Intelligent Control, Ministry of Education, Huazhong University of Science and Technology, Wuhan 430074, China*

Correspondence should be addressed to Zhi-Wei Liu; zwliu@hust.edu.cn

Received 12 March 2021; Accepted 30 April 2021; Published 7 May 2021

Academic Editor: Xiao Ling Wang

Time-varying formation-containment problems for a second-order multiagent system (SOMAS) are studied via pulse-modulated intermittent control (PMIC) in this paper. A distributed control framework utilizing the neighbors' positions and velocities is designed so that leaders in the multiagent system form a formation, and followers move to the convex hull formed by each leader. Different from the traditional formation-containment problems, this paper applies the PMIC framework, which is more common and more in line with the actual control scenarios. Based on the knowledge of matrix theory, algebraic graph theory, and stability theory, some sufficient conditions are given for the time-varying formation-containment problem of the second-order multiagent system. Some numerical simulations are proposed to verify the effectiveness of the results presented in this paper.

1. Introduction

Many scholars start to pay attention to the multiagent systems (MASs) with the rapid development of complex network system theory. In recent decades, many major breakthroughs have been made in this field [1], and those results are also widely used in various fields of production and life, such as UAV cruise system [2], smart grid [3, 4], economic dispatching [5], and multiple underactuated surface vessels [6]. The most studied collaboration problems of MASs include swarm, consensus, formation, and distributed optimization. In very recent years, the clustering behavior of MASs has also attracted widespread interest, including but not limited to consensus [7–9], tracking, and formation [10].

In these cooperative control problems, many important advances have been made in the field of formation control and containment control. Huang et al. [11] studied the containment control problem of MASs via intermittent control-based sampled data information. Wang et al. [12] investigated the containment control problem of first-order MASs in the noisy communication environments. Rahimi et al. [13] studied time-varying formation control of collaborative heterogeneous MASs. However, most of the existing works are carried out on formation control and containment control separately. In many applications, both these cooperative behaviors often require simultaneous implementation. For example, in the coordination of multiple tanker airplanes and multiple UAVs, the tanker airplanes will form a specific formation in advance and wait for the UAVs to reach the area they surround. In order to solve this kind of problem, Dong et al. designed a continuous control strategy [14] for second-order MASs with multiple leaders and multiple followers, which can ensure the formation control of the leaders and containment of the followers. It is worth noting that the above work [14] adopted a continuous-time control which may be difficult to implement in some cases due to sampled measurement. The traditional zero-order sampling control adopts the same amplitude control input in the whole sampling interval,

which can also lead to application difficulties. For instance, in a driverless system, it is difficult for the sensors on the vehicle to work all the time, allowing for fuel economy and other reasons.

Aiming at solving the aforementioned problems, the formation-containment problem of the second-order multiagent system (SOMAS) is studied in this paper, and the control framework of pulse-modulated intermittent control (PMIC) is adopted. The main contributions are given as follows. (i) The formation and the containment are achieved simultaneously in second-order MASs, where the leaders form a formation and the followers are contained in this formation. (ii) The PIMC is a framework that can unify impulsive control [8, 15] and zero-order sampling control [16]. It can be applied to a wider range of real-world scenarios. (iii) In this paper, some sufficient conditions are given for the parameters of the control strategy under with/without time delay cases.

The remainder sections of this paper are described as follows. Section 2 lists the basic preliminary knowledge and a model of the problem to be studied. Section 3 introduces the formation-containment analysis without time delay. Section 4 obtains and provides the results when MASs contain time delay. Section 5 gives several simulations to verify the theorems are correct. Section 6 draws the conclusion.

2. Preliminaries

2.1. Graph Theory and Some Lemmas. Let $G = (V, E, A)$ be a weighted digraph, where $V = \{1, \ldots, N\}$ is a vertex set, $E \subseteq V \times V$ is a link set, and $A = [a_{ij}] \in \mathbb{R}^{N\times N}$ is a nonnegative weighted adjacency matrix. The information flow from vertex j to vertex i is represented by a directed link $e_{ij} = (j, i)$. The elements of matrix A are described as follows: $a_{ij} > 0$ if $e_{ij} \in E$, and $a_{ij} = 0$, otherwise. Furthermore, $a_{ii} = 0$ for all $i \in V$, and $a_{ij} = a_{ji}$ in an undirected topology. Denote $N_i = \{v_j \in V: (v_j, v_i) \in E\}$ as the set of neighbors of node v_i. The Laplacian matrix L is described in terms of $L = D - A$, where $D = \text{diag}\{D_1, D_2, \ldots, D_N\}$ is the in-degree matrix with $D_i = \sum_{j=1}^{N} a_{ij}$. A sequence of edges with different nodes of V as $\{(v_i, v_{i1}), (v_i, v_{i2}), \ldots, (v_i, v_j)\}$ can be obtained if there is a directed path from the ith agent to the jth agent.

In a MAS, a SOMAS with N agents is considered, and there are M followers and $N - M$ leaders. Assume that followers can receive messages from the leaders or followers, while the leaders can only receive messages from the leaders. Let $F = \{1, 2, \ldots, M\}$ and $E = \{M + 1, M + 2, \ldots, N\}$ denote the sets of followers and leaders, respectively. And the Laplacian matrix L_N is described as

$$L_N = \begin{bmatrix} L_1 & L_2 \\ 0 & L_3 \end{bmatrix}, \tag{1}$$

where $L_1 \in \mathbb{R}^{M\times M}$, $L_2 \in \mathbb{R}^{M\times(N-M)}$, and $L_3 \in \mathbb{R}\ (N - M) \times (N - M)$.

Lemma 1 (see [17]). *A complex characteristic polynomial $R(z) = z^2 + sz + r$ is Hurwitz stable if and only if $Re(s) > 0$ and $Re(s)Im(s)Im(r) + Re^2(s)Re(r) - Im^2(r) > 0$.*

Lemma 2 (see [18]). *If directed graph G contains a spanning tree, then the Laplacian matrix $L \in \mathbb{R}^{N\times N}$ of G has a simple zero eigenvalue with $\overrightarrow{1}_N$ as the associated eigenvector, and all the other $N - 1$ eigenvalues have positive real parts.*

Assumption 1. Each follower of graph G has at least one directed path from one leader.

Assumption 2. Let G_E be the graph associated with the leaders in the MAS, and G_E contains a spanning tree.

Lemma 3 (see [19]). *By Assumptions 1 and 2, it is obtained that the eigenvalues of L_1 have positive real parts, each row of $-L_1^{-1}L_2$ has a sum equal, and each entry of $-L_1^{-1}L_2$ is nonnegative.*

2.2. Model Formulation and Some Definitions. The control input of the ith agent is denoted by $u_i \in \mathbb{R}^n$, position by $p_i \in \mathbb{R}^n$, and velocity by $v_i \in \mathbb{R}^n$, respectively. Consider a SOMAS as

$$\begin{aligned} \dot{p}_i(t) &= v_i(t), \\ \dot{v}_i(t) &= u_i(t). \end{aligned} \tag{2}$$

In the following, for the sake of description, we assume that $n = 1$. However, more cases such as $n > 1$ can be derived by using the Kronecker product.

Let, $B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, and $x_i(t) = [p_i(t), v_i(t)]^T$ $(i = 1, 2, \ldots, N)$. Then, the dynamic equation of the ith agent can be described in a neat form as follows:

$$\dot{x}_i(t) = Ax_i(t) + Bu_i(t). \tag{3}$$

Definition 1. With the bounded initial state of each agent being chosen arbitrarily, the leaders in MAS (3) are said to realize time-varying formation if there exists a vector function $c(t) \in \mathbb{R}^2$ such that

$$\lim_{t\longrightarrow\infty} \left(x_i(t) - h_i(t) - c(t)\right) = 0 \quad (i \in E), \tag{4}$$

where $h_i(t) = [h_{ip}(t), h_{iv}(t)]^T$ is a piecewise continuously differentiable vector. Then, we can denote the time-varying formation as $h_E(t) = [h_{M+1}^T(t), h_{M+2}^T(t), \ldots, h_N^T(t)]^T \in \mathbb{R}^{2(N-M)}$.

Definition 2. Similarly, with the bounded initial state of each agent being chosen arbitrarily, MAS (3) is said to achieve containment if there exist nonnegative constants $\tau_{k,j}$ $(k \in F, j \in E)$ such that $\sum_{j=M+1}^{N} \tau_{k,j} = 1$ and

$$\lim_{t\longrightarrow\infty}\left(x_k(t)-\sum_{j=M+1}^{N}\tau_{k,j}x_j(t)\right)=0. \tag{5}$$

3. Formation-Containment Analysis without Time Delay

In this section, we mainly study how to design the PMIC protocol to make MAS (3) achieve the time-varying formation-containment and propose some sufficient and necessary conditions for parameters. We will investigate the time-varying formation-containment problem in two steps. The first step is to transform the formation-containment problem into a stability problem. The second step is to solve the stability problem according to the related theory.

3.1. Problem Transformation. Consider the following PMIC protocols:

$$\begin{aligned}u_i(t)&=a(t-t_k)K_1\sum_{j\in N_i}a_{ij}\big((x_i(t_k)-h_i(t_k))\\&\quad-\big(x_j(t_k)-h_j(t_k)\big)\big),\quad i\in E,\\u_i(t)&=a(t-t_k)K_2\sum_{j\in N_i}a_{ij}\big(x_i(t_k)-x_j(t_k)\big),\quad i\in F,\end{aligned} \tag{6}$$

where $K_1=[k_{11},k_{12}]\in\mathbb{R}^{1\times2}$ and $K_2=[k_{21},k_{22}]\in\mathbb{R}^{1\times2}$ are feedback gain matrices and $t_k<t<t_{k+1}$. The pulse function $a(t)$ is described as

$$a(t)=\begin{cases}\widehat{a}(t), & t\in(0,d],\\0, & t\notin(0,d],\end{cases} \tag{7}$$

where $\widehat{a}(t)$ is a piecewise continuous function. Let $T=t_{k+1}-t_k$ and $d<T$ be the control duration within a complete sampling cycle. $(t_k+d,t_{k+1}]$ is the rest interval, and $(t_k,t_k+d]$ is the control interval [20].

Under the control framework (6), MAS (3) can be described in a compact form as follows:

$$\begin{aligned}\dot{x}_E(t)&=(I_{N-M}\otimes A)x_E(t)+a(t-t_k)\big((L_3\otimes BK_1)x_E(t_k)\\&\quad-(L_3\otimes BK_1)h_E(t_k)\big),\\\dot{x}_F(t)&=(I_M\otimes A)x_F(t)+a(t-t_k)\big((L_1\otimes BK_2)x_F(t_k)\\&\quad+(L_2\otimes BK_2)x_E(t_k)\big),\end{aligned} \tag{8}$$

where $x_E(t)=[x_{M+1}^T(t),x_{M+2}^T(t),\ldots,x_N^T(t)]^T$ and $x_F(t)=[x_1^T(t),x_2^T(t),\ldots,x_M^T(t)]^T$.

Let $\theta_i(t)=x_i(t)-h_i(t)(i\in E)$ and $\theta_E(t)=[\theta_{M+1}^T(t),\theta_{M+2}^T(t),\ldots,\theta_N^T(t)]^T$. Then, system (8) can be written as

$$\begin{aligned}\dot{\theta}_E(t)&=(I_{N-M}\otimes A)\theta_E(t)+a(t-t_k)(L_3\otimes BK_1)\theta_E(t_k)\\&\quad+(I_{N-M}\otimes A)h_E(t)-(I_{N-M}\otimes I_2)\dot{h}_E(t).\end{aligned} \tag{9}$$

The eigenvalues of L_3 relating to G_E are denoted by $\lambda_i(i\in E)$, where $\lambda_{M+1}=0$ with the eigenvector as $p_{M+1}=(\mathbf{1}_{N-M}/\sqrt{N-M})$, and $0<\mathrm{Re}(\lambda_{M+2})\le\cdots\le\mathrm{Re}(\lambda_N)$. Let $P_E^{-1}L_3P_E=J_E$, where J_E is the Jordan canonical form of L_3, $P_E=[p_{M+1},p_{M+2},\ldots,p_N]$, and $P_E^{-1}=[\widetilde{p}_{M+1},\widetilde{p}_{M+2},\ldots,\widetilde{p}_N]^H$. By Lemma 2, one has $J_E=\mathrm{diag}\{0,\overline{J}_E\}$, where $\overline{J}_E$ consists of the Jordan blocks relating to $\lambda_i(i=M+2,M+3,\ldots,N)$. Let $\widetilde{\theta}_E(t)=(P_E^{-1}\otimes I_2)\theta_E(t)$; then, system (9) can be rewritten as

$$\begin{aligned}\dot{\widetilde{\theta}}_E(t)&=(I_{N-M}\otimes A)\widetilde{\theta}_E(t)+a(t-t_k)\big(P_E^{-1}L_3P_E\otimes BK_1\big)\widetilde{\theta}_E(t_k)\\&\quad+\big(P_E^{-1}\otimes A\big)h_E(t)-\big(P_E^{-1}\otimes I_2\big)\dot{h}_E(t).\end{aligned} \tag{10}$$

Let $\widetilde{P}_E=[\widetilde{p}_{M+2},\widetilde{p}_{M+3},\ldots,\widetilde{p}_N]^H$, $\alpha_E(t)=(\widetilde{p}_{M+1}^H\otimes I_2)\theta_E(t)$, and $\beta_E(t)=(\widetilde{P}_E\otimes I_2)\theta_E(t)$. Then, system (10) can be divided as

$$\begin{aligned}\dot{\alpha}_E(t)&=A\alpha_E(t)+\big(\widetilde{p}_{M+1}^H\otimes A\big)h_E(t)-\big(\widetilde{p}_{M+1}^H\otimes I_2\big)\dot{h}_E(t),\\\dot{\beta}_E(t)&=(I_{N-M-1}\otimes A)\beta_E(t)+a(t-t_k)(\overline{J}_E\otimes BK_1)\beta_E(t_k)\\&\quad+(\widetilde{P}_E\otimes A)h_E(t)-(\widetilde{P}_E\otimes I_2)\dot{h}_E(t).\end{aligned} \tag{11}$$

Let $\delta_i=\sum_{j\in N_i}w_{ij}(x_i(t)-x_j(t)),i\in F$, and $\delta_F(t)=[\delta_1(t),\delta_2(t),\ldots,\delta_M(t)]$. Then, the following lemma is used to transform the formation-containment problem.

Lemma 4. *MAS (3) under the PMIC framework (6) can achieve time-varying formation-containment if*

$$\begin{cases}\lim_{t\longrightarrow\infty}\beta_E(t)=0_{2(N-M-1)},\\\lim_{t\longrightarrow\infty}\delta_F(t)=0_{2M}.\end{cases} \tag{12}$$

Proof. It is able to be proved by a similar way in [14]. □

Remark 1. Lemma 4 is proved because the following results in [14] are worked out:

$$\begin{aligned}\lim_{t\longrightarrow\infty}\left(x_E(t)-h_E(t)-\left(\frac{\mathbf{1}_{N-M}}{\sqrt{N-M}}\otimes I_2\right)\alpha_E(t)\right)&=0_{2(N-M)},\\\lim_{t\longrightarrow\infty}\big(x_F(t)-\big(-L_1^{-1}L_2\otimes I_2\big)x_E(t)\big)&=0_{2M}.\end{aligned} \tag{13}$$

Considering Lemma 3, we can conclude that equation (13) satisfies Definitions 1 and2, respectively. In other words, Lemma 4 can be proved by the above two equations.

3.2. Control Design. By means of Lemma 4, the time-varying formation-containment problem can be transformed into the convergence analysis of $\beta_E(t)$ and $\delta_F(t)$. This section presents the conditions for the parameters in control protocol (6) when $\beta_E(t)$ and $\delta_F(t)$ converge to 0.

Assumption 3. For $\forall i\in\{M+1,M+2,\ldots,N\}$,

$$\lim_{t\longrightarrow\infty}\left(A\left(h_i(t)-h_j(t)\right)-\left(\dot{h}_i(t)-\dot{h}_j(t)\right)\right)=0,\quad j\in N_i. \tag{14}$$

Lemma 5. *Let*

$$\dot{\widetilde{\beta}}_E(t)=\left(I_{N-M-1}\otimes A\right)\widetilde{\beta}_E(t)+a\left(t-t_k\right)\left(\overline{J}_E\otimes BK_1\right)\widetilde{\beta}_E(t_k). \tag{15}$$

Under Assumption 3, $\lim_{t\longrightarrow\infty}\beta_E(t)=0_{2(N-M-1)}$ is satisfied if

$$\lim_{t\longrightarrow\infty}\widetilde{\beta}_E(t)=0_{2(N-M-1)}. \tag{16}$$

Proof. If Assumption 3 holds, one has

$$\lim_{t\longrightarrow\infty}\left(\left(L_3\otimes A\right)h_E(t)-\left(L_3\otimes I_2\right)\dot{h}_E(t)\right)=0. \tag{17}$$

Substitute $L_3=P_EJ_EP_E^{-1}$ into (17), and then premultiply both sides of (17) by $P_E^{-1}\otimes I_2$. One can obtain that

$$\lim_{t\longrightarrow\infty}\left(\left(\overline{J}_E\widetilde{P}_E\otimes A\right)h_E(t)-\left(\overline{J}_E\widetilde{P}\otimes I_2\right)\dot{h}_E(t)\right)=0. \tag{18}$$

Under Assumption 2 and Lemma 2, $\overline{J}_E$ is nonsingular, obviously. By premultiplying both sides of (18) by $\overline{J}_E^{-1}\otimes I_2$, we have

$$\lim_{t\longrightarrow\infty}\left(\left(\widetilde{P}_E\otimes A\right)h_E(t)-\left(\widetilde{P}\otimes I_2\right)\dot{h}_E(t)\right)=0. \tag{19}$$

Considering equation (19), Lemma 5 is set up. □

Lemma 6. *$\forall i\in\{M+2,M+3,\ldots,N\}$, equation (16) holds if the following inequalities are satisfied:*

$$\frac{k_{12}}{k_{11}}>\frac{d_2}{d_1}, \tag{20}$$

$$\left(\frac{2k_{12}}{k_{11}T}-\frac{2d_2}{d_1T}\right)^2\left(\frac{2d_2}{d_1T}-\frac{2k_{12}}{k_{11}T}-\frac{4\mathrm{Re}(\lambda_i)}{k_{11}d_1|\lambda_i|^2T}-1\right)-\frac{16\mathrm{Im}^2(\lambda_i)}{k_{11}^2d_1^2|\lambda_i|^4T^2}>0, \tag{21}$$

where $d_1=\int_0^d\widehat{a}(s)ds$ and $d_2=\int_0^d s\widehat{a}(s)ds$.

Proof. Equation (16) holds if and only if $\widetilde{\beta}_E$ in (15) is asymptotically stable. The solution of (15) can be written as

$$\widetilde{\beta}_E(t)=e^{C\left(t-t_0\right)}\widetilde{\beta}_E(t_0)+e^{Ct}\int_{t_0}^{t}e^{-Cs}a\left(s-t_k\right)D\widetilde{\beta}_E(t_k)\mathrm{d}s, \tag{22}$$

where $C=I_{N-M-1}\otimes A$ and $D=\overline{J}_E\otimes BK_1$. Let $t_k=t_0$ and $t_k+d=t$ in (22); then, we have

$$\widetilde{\beta}_E\left(t_k+d\right)=e^{Cd}\widetilde{\beta}_E(t_k)+e^{C\left(t_k+d\right)}\int_{t_k}^{t_k+d}e^{-Cs}a\left(s-t_k\right)D\widetilde{\beta}_E(t_k)\mathrm{d}s. \tag{23}$$

According to $C^2=0a$ and the theory of matrix function, one can obtain that

$$\widetilde{\beta}_E\left(t_k+d\right)=\left(I_{2(N-M-1)}+dC+d_1D+\left(dd_1-d_2\right)CD\right)\widetilde{\beta}_E(t_k). \tag{24}$$

Notice that $\widetilde{\beta}_E(t_{k+1})=\widetilde{\beta}_E(t_k+T)$; then, the system can be discretely expressed as

$$\widetilde{\beta}_E\left(t_{k+1}\right)=\left(I_{2(N-M-1)}+TC+d_1D+\left(Td_1-d_2\right)CD\right)\widetilde{\beta}_E(t_k). \tag{25}$$

Let $\Gamma=I_{2(N-M-1)}+TC+d_1D+(Td_1-d_2)CD$ and $\det(sI_{2(N-M-1)}-\Gamma)=0$. It yields

$$s^2-\left[2+k_{12}d_1\lambda_i+k_{11}\lambda_i\left(Td_1-d_2\right)\right]s+1+k_{12}d_1\lambda_i-k_{11}d_2\lambda_i=0. \tag{26}$$

Considering $\lambda_i\neq0$, then it is focused on the conditions that the eigenvalues of Γ are encircled by the unit circle. By applying a bilinear transformation, $z=((s+1)/(s-1))$, an updated polynomial can be found as

$$R(z)=z^2+\frac{2\left(k_{12}d_1\lambda_i-k_{11}\lambda_id_2\right)}{k_{11}\lambda_id_1T}z+\frac{2k_{11}\lambda_id_2-2k_{12}d_1\lambda_i-k_{11}d_1\lambda_iT-4}{k_{11}\lambda_id_1T}. \tag{27}$$

$\|s\|<1$ in (26) can be obtained if and only if $R(z)$ in (27) is Hurwitz stable. Based on Lemma 1, if (27) is stable, then both equations (20) and (21) hold. □

Lemma 7. *$\lim_{t\longrightarrow\infty}\delta_F(t)=0_{2M}$ holds if and only if $\forall i\in\{1,2,\ldots,M\}$,*

$$\frac{k_{22}}{k_{21}}>\frac{d_2}{d_1}, \tag{28}$$

$$\left(\frac{2k_{22}}{k_{21}T}-\frac{2d_2}{d_1T}\right)^2\left(\frac{2d_2}{d_1T}-\frac{2k_{22}}{k_{21}T}-\frac{4\mathrm{Re}(\lambda_i)}{k_{21}d_1|\lambda_i|^2T}-1\right)-\frac{16\mathrm{Im}^2(\lambda_i))}{k_{21}^2d_1^2|\lambda_i|^4T^2}>0, \tag{29}$$

where $d_1=\int_0^d\widehat{a}(s)ds$ and $d_2=\int_0^d s\widehat{a}(s)ds$.

Proof. It is obtained that

$$\delta_F(t)=\left(L_2\otimes I_2\right)x_E(t)+\left(L_1\otimes I_2\right)x_F(t). \tag{30}$$

Substituting (8) into (30), we can obtain that

$$\dot{\delta}_F(t)=\left(I_M\otimes A\right)\delta_F(t)+a\left(t-t_k\right)\left(L_1\otimes BK_2\right)\delta_F(t_k)\ a\left(t-t_k\right)\left(L_2L_3\otimes BK_1\right)\left(x_E(t_k)-h_E(t_k)\right). \tag{31}$$

By Lemmas 4 and 5, MAS (8) can achieve formation $h_E(t)$. Considering Definition 1, when the leaders' formation $h_E(t)$ is achieved, then

$$\lim_{t\longrightarrow\infty}\left(L_2L_3\otimes BK_1\right)\left(x_E\left(t_k\right)-h_E\left(t_k\right)\right)=\lim_{t\longrightarrow\infty}\left(L_2L_3\otimes BK_1\right)\left(\mathbf{1}\otimes c\left(t_k\right)\right). \tag{32}$$

Since $L_3\mathbf{1}=0$, one has

$$\lim_{t\longrightarrow\infty}\left(L_2L_3\otimes BK_1\right)\left(x_E\left(t_k\right)-h_E\left(t_k\right)\right)=0. \tag{33}$$

Then, consider the following system:

$$\dot{\delta}_F\left(t\right)=\left(I_M\otimes A\right)\delta_F\left(t\right)+a\left(t-t_k\right)\left(L_1\otimes BK_2\right)\delta_F\left(t_k\right). \tag{34}$$

$\lim_{t\longrightarrow\infty}\delta_F\left(t\right)=0_{2M}$ means that system (34) is asymptotically stable. Similar to the analysis of Lemma 5, the conditions that make system (34) asymptotically stable are worked out with (28) and (29). □

Theorem 1. *MAS (3) under Assumption 3 and PMIC framework (6) realizes formation-containment if the following conditions simultaneously hold:*

(i) $\forall i\in\{1,2,\ldots,M\}$,

$$\frac{k_{22}}{k_{21}}>\frac{d_2}{d_1},$$
$$\left(\frac{2k_{22}}{k_{21}T}-\frac{2d_2}{d_1T}\right)^2\left(\frac{2d_2}{d_1T}-\frac{2k_{22}}{k_{21}T}-\frac{4\mathrm{Re}\left(\lambda_i\right)}{k_{21}d_1\left|\lambda_i\right|^2T}-1\right)-\frac{16\mathrm{Im}^2\left(\lambda_i\right)}{k_{21}^2d_1^2\left|\lambda_i\right|^4T^2}>0. \tag{35}$$

(ii) $\forall i\in\{M+2,M+3,\ldots,N\}$,

$$\frac{k_{12}}{k_{11}}>\frac{d_2}{d_1},$$
$$\left(\frac{2k_{12}}{k_{11}T}-\frac{2d_2}{d_1T}\right)^2\left(\frac{2d_2}{d_1T}-\frac{2k_{12}}{k_{11}T}-\frac{4\mathrm{Re}\left(\lambda_i\right)}{k_{11}d_1\left|\lambda_i\right|^2T}-1\right)-\frac{16\mathrm{Im}^2\left(\lambda_i\right)}{k_{11}^2d_1^2\left|\lambda_i\right|^4T^2}>0, \tag{36}$$

where $d_1=\int_0^d\widehat{a}\left(s\right)ds$ *and* $d_2=\int_0^d s\widehat{a}\left(s\right)ds$.

Proof. It is proved naturally by Lemmas 4, 6, and 7. □

Remark 2. When selecting parameters in the control rules, we generally give K_1, K_2, d_1, and d_2 that meet the conditions first. According to the values of these parameters and conditions (35) and (36) in Theorem 1, we can solve parameter T as follows:

(i) $\forall i\in\{1,2,\ldots,M\}$,

$$T<\min\frac{M_1^3\left|\lambda_i\right|^4+4\mathrm{Re}\left(\lambda_i\right)\left|\lambda_i\right|^2M_1^2}{16k_{21}d_1\mathrm{Im}^2\left(\lambda_i\right)+k_{21}d_1M_1^2\left|\lambda_i\right|^4}, \tag{37}$$

where $M_1=2k_{21}d_2-2k_{22}d_1$.

(ii) $\forall i\in\{M+2,M+3,\ldots,N\}$,

$$T<\min\frac{M_1^3\left|\lambda_i\right|^4+4\mathrm{Re}\left(\lambda_i\right)\left|\lambda_i\right|^2M_1^2}{16k_{11}d_1\mathrm{Im}^2\left(\lambda_i\right)+k_{11}d_1M_1^2\left|\lambda_i\right|^4}, \tag{38}$$

where $M_2=2k_{11}d_2-2k_{12}d_1$.

4. Formation-Containment Analysis with Time Delay

We know that the time delay cannot be ignored due to its widely existence in the real world. Then, a control protocol with considering time delay is proposed as follows:

$$\begin{aligned}u_i\left(t\right)&=a\left(t-t_k\right)K_1\sum_{j\in N_i}w_{ij}\Big(\left(x_i\left(t_k-\tau\right)-h_i\left(t_k\right)\right)-\left(x_j\left(t_k-\tau\right)-h_j\left(t_k\right)\right)\Big),\quad i\in E,\\ u_i\left(t\right)&=a\left(t-t_k\right)K_2\sum_{j\in N_i}w_{ij}\left(x_i\left(t_k-\tau\right)-x_j\left(t_k-\tau\right)\right),\quad i\in F.\end{aligned} \tag{39}$$

The control protocol (39) is substituted into MAS (3), and we can obtain

$$\begin{aligned}\dot{x}_E(t) &= (I_{N-M}\otimes A)x_E(t) + a(t-t_k)\left((L_3\otimes BK_1)x_E(t_k-\tau) - (L_3\otimes BK_1)h_E(t_k)\right),\\ \dot{x}_F(t) &= (I_M\otimes A)x_F(t) + a(t-t_k)\left((L_1\otimes BK_2)x_F(t_k-\tau) + (L_2\otimes BK_2)x_E(t_k-\tau)\right).\end{aligned} \tag{40}$$

Considering Assumption 3, when there is a time delay in the system, similar to Section 3, we have

$$\begin{aligned}\dot{\widetilde{\beta}}_E(t) &= (I_{N-M-1}\otimes A)\widetilde{\beta}_E(t) + a(t-t_k)(\overline{J}_E\otimes BK_1)\widetilde{\beta}_E(t_k-\tau),\\ \dot{\delta}_F(t) &= (I_M\otimes A)\delta_F(t) + a(t-t_k)(L_1\otimes BK_2)\delta_F(t_k-\tau).\end{aligned} \tag{41}$$

It is not hard to see that Lemma 4 also applies to cases with time delay. Therefore, we still transform the formation-containment problem of the SOMAS into the stability problem of the system.

Remark 3. Considering that the time delay of the actual system is not long and the intermittent control has the advantage of selectively adjusting the control time interval d, we assume that the time delay $\tau < \min\{d, T-d\}$.

Lemma 8. *$lim_{t\longrightarrow\infty}\widetilde{\beta}_E(t) = 0_{2(N-M-1)}$ holds with time delay if $\forall i \in \{M+2, M+3, \ldots, N\}$,*

$$-\frac{2\tau}{T} - \frac{2d_2}{d_1T} + \frac{2k_{12}}{k_{11}T} > 0, \tag{42}$$

$$\left(-\frac{2\tau}{T} - \frac{2d_2}{d_1T} + \frac{2k_{12}}{k_{11}T}\right)^2\left[-\frac{4\mathrm{Re}(\lambda_i)}{|\lambda_i|^2k_{11}d_1T} - 1 + \frac{2\tau}{T} + \frac{2d_2}{d_1T} - \frac{2k_{12}}{k_{11}T}\right] - \frac{16\mathrm{Im}^2(\lambda_i)}{|\lambda_i|^4k_{11}^2d_1^2T^2} > 0, \tag{43}$$

where $d_1 = \int_0^d \widehat{a}(s)ds$, $d_2 = \int_0^d s\widehat{a}(s)ds$, and τ is the time delay.

Proof. Similarly as Lemma 6, differential equation (41) is solved. Since $\tau < \min\{d, T-d\}$, when $t \in [t_k, t_{k+1})$, we can obtain that

$$\widetilde{\beta}_i(t) = M(t-t_k)\widetilde{\beta}_i(t_k) + N(t-t_k)\widetilde{\beta}_i(t_k-\tau), \tag{44}$$

where $M(\chi) = \begin{bmatrix}1 & t\\ 0 & 1\end{bmatrix}$ and $N(\chi) = \begin{bmatrix}(td_1-d_2)\lambda_ik_{11} & (td_1-d_2)\lambda_ik_{12}\\ \lambda_ik_{11}d_1 & \lambda_ik_{12}d_1\end{bmatrix}$. To solve the stability problem, we need to consider the iterative problem of the system. When $t = t_k$, one has

$$\begin{aligned}\widetilde{\beta}_i(t_k) &= M(T)\widetilde{\beta}_i(t_{k-1}) + N(T)\widetilde{\beta}_i(t_{k-1}-\tau),\\ \widetilde{\beta}_i(t_k-\tau) &= M(T-\tau)\widetilde{\beta}_i(t_{k-1}) + N(T-\tau)\widetilde{\beta}_i(t_{k-1}-\tau).\end{aligned} \tag{45}$$

Let $\omega_i(t) = \begin{bmatrix}\widetilde{\beta}_i(t)\\ \widetilde{\beta}_i(t-\tau)\end{bmatrix}$; combined with (43), it is obtained that

$$\omega_i(t) = \begin{bmatrix}\widetilde{\beta}_i(t)\\ \widetilde{\beta}_i(t-\tau)\end{bmatrix} = \Gamma_1\Gamma_2^k\begin{bmatrix}\widetilde{\beta}_i(t_0)\\ \widetilde{\beta}_i(t_0-\tau)\end{bmatrix}, \tag{46}$$

where $\Gamma_1 = \begin{bmatrix}M(t-t_k) & N(t-t_k)\\ M(t-t_k-\tau) & N(t-t_k-\tau)\end{bmatrix}$ and $\Gamma_2 = \begin{bmatrix}M(T) & N(T)\\ M(T-\tau) & N(T-\tau)\end{bmatrix}$. It is easy to get that $\omega_i(t) \longrightarrow 0$ if the eigenvalues of Γ_2 meet $\|s\| < 1$. Let $|sI_4 - \Gamma_2| = 0$. It yields

$$s^4 + a_1s^3 + a_0s^2 = 0, \tag{47}$$

where $a_1 = -\lambda_ik_{11}d_1T + \lambda_ik_{11}d_1\tau + \lambda_ik_{11}d_2 - \lambda_ik_{12}d_1 - 2$ and $a_0 = -\lambda_ik_{11}d_1\tau - \lambda_ik_{11}d_2 + \lambda_ik_{12}d_1 + 1$. We can obtain that two eigenvalues of Γ_2 satisfy $s_1 = s_2 = 0$. Then, the remaining proof process has been omitted, which is similar to the proof of Lemma 6. □

Lemma 9. *$lim_{t\longrightarrow\infty}\delta_F(t) = 0_{2M}$ holds with time delay if $\forall i \in \{1, 2, \ldots, M\}$,*

$$-\frac{2\tau}{T} - \frac{2d_2}{d_1 T} + \frac{2k_{22}}{k_{21}T} > 0, \tag{48}$$

$$\left(-\frac{2\tau}{T} - \frac{2d_2}{d_1 T} + \frac{2k_{22}}{k_{21}T}\right)^2 \left[-\frac{4\mathrm{Re}(\lambda_i)}{|\lambda_i|^2 k_{21} d_1 T} - 1 + \frac{2\tau}{T} + \frac{2d_2}{d_1 T} - \frac{2k_{22}}{k_{21}T}\right] - \frac{16\mathrm{Im}^2(\lambda_i)}{|\lambda_i|^4 k_{21}^2 d_1^2 T^2} > 0, \tag{49}$$

where $d_1 = \int_0^d \widehat{a}(s)ds$, $d_2 = \int_0^d s\widehat{a}(s)ds$, and τ is the time delay.

Proof. This proof is analogous to Lemma 8 and is therefore omitted. □

Theorem 2. *MAS (3) under Assumption 3 and control framework (39) realizes formation-containment if conditions (42), (43), (48), and (49) hold.*

Proof. It is not hard to prove by applying Lemmas 4, 8, and 9. □

5. Simulation

In this section, a two-dimensional formation-containment case of the MAS will be shown by numerical simulations. The simulations illustrate the effectiveness of the results in this paper. In this two-dimensional formation-containment case, suppose there are three leaders and two followers in the MAS. The Laplace matrix of the directed topology is given as

$$L = \begin{bmatrix} 2 & -\frac{2}{3} & -\frac{1}{3} & -1 & 0 \\ -1 & 2 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}. \tag{50}$$

Note that $p_i(t) \in \mathbb{R}^2$, $v_i(t) \in \mathbb{R}^2$, and $x_i(t) = [p_{iX}(t), v_{iX}(t), p_{iY}(t), v_{iY}(t)]^T$, $i \in \{1, 2, \ldots, 5\}$. Then, matrices A and B in dynamic equation (3) should be extended as

$$A = I_2 \otimes \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad B = I_2 \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{51}$$

The formation function for leaders is given by

$$h_i(t) = \begin{bmatrix} \sin\left(t + \frac{2(i-3)\pi}{3}\right) \\ \cos\left(t + \frac{2(i-3)\pi}{3}\right) \\ \cos\left(t + \frac{2(i-3)\pi}{3}\right) \\ -\sin\left(t + \frac{2(i-3)\pi}{3}\right) \end{bmatrix} \quad (i = 3, 4, 5). \tag{52}$$

The eigenvalues of L are $\lambda_1 = 2.8165$, $\lambda_2 = 1.1835$, $\lambda_3 = 0$, $\lambda_4 = 1.5 + 0.866i$, and $\lambda_5 = 1.5 - 0.866i$. The function $\widehat{a}(t) = (t/d)$ is set, and the control duration $d = 0.2$ is chosen; then, $d_1 = 0.1$ and $d_2 = 0.01333$.

When we do not care about time delays, it follows from (20) and (28) that $k_{12}/k_{11} > 0.1333$ and $k_{22}/k_{21} > 0.1333$. The feedback gains are chosen as $K_1 = I_2 \otimes [-1, -2]$ and $K_2 = I_2 \otimes [-1, -2]$. According to inequalities (21) and (29), $T < 1.5396$ is obtained, and $T = 0.6$ is chosen. When time delays are taken into account, we assume the time delay $\tau = 0.1$. According to Lemmas 8 and 9, we also choose the feedback gains as $K_1 = I_2 \otimes [-1, -2]$ and $K_2 = I_2 \otimes [-1, -2]$ and set $T = 0.6$.

The simulation results are shown in figures. Among them, Figures 1 and 2 are the trajectories of the position and velocity of each agent changing without considering the time delay, in which the initial state is marked with a circle and the final state is marked with a triangular row. Figures 3 and 4 show the trajectories when time delay is taken into account. The same initial state is indicated by a circle, and the final state is indicated by a triangular row. The initial states of all agents are randomly selected.

During the simulation, the five agents begin with arbitrary initial states. As time goes on, the three leaders reach a time-varying circular formation, and two followers move into the convex hull formed by leaders. At the same time, it can be seen that the system with time delay converges more slowly with the same sampling period.

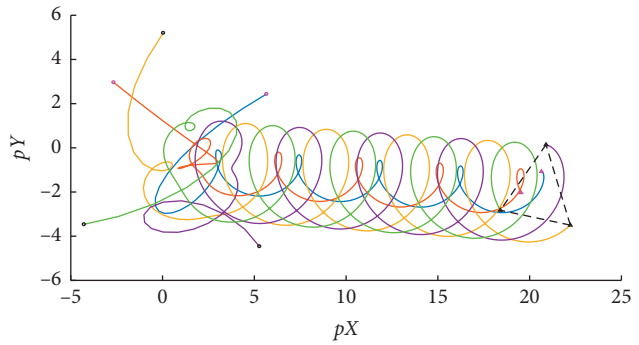


Figure 1: The trajectory of agents' positions over time without time delay.

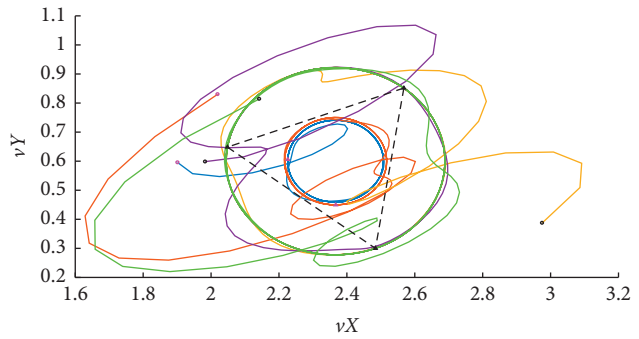


Figure 2: The trajectory of agents' velocities over time without time delay.

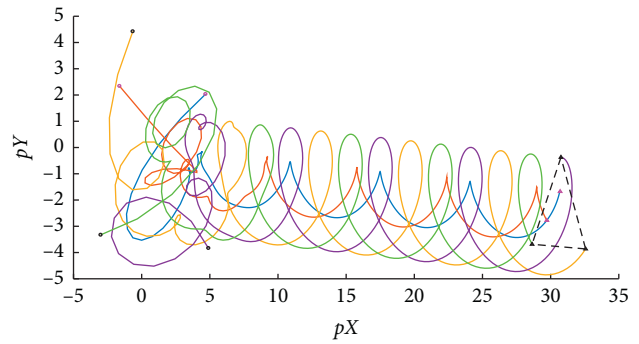


Figure 3: The trajectory of agents' positions over time with time delay.

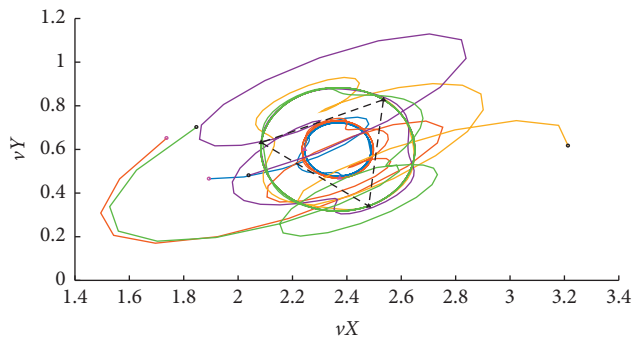


Figure 4: The trajectory of agents' velocities over time with time delay.

6. Conclusion

Time-varying formation-containment problems of the SOMAS were studied in this paper. Based on the fundamental theorems of graph theory and matrix theory, the formation-containment problem is transformed into the stability problem of the SOMAS. The PMIC protocols are designed, and the sufficient conditions for the time-varying formation-containment of the SOMAS are given by proving three theorems. The simulations show the correctness of the proposed theoretical results. It is expected that the future work will be carried out on the nonlinear MAS, and the conclusion of this paper will be applied to the actual multi-intelligent vehicle experimental platform.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant nos. 61972170 and 61973133).

References

[1] H. Su, X. Wang, and Z. Lin, "Flocking of multi-agents with a virtual leader," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 293–307, 2009.

[2] X. Dong, B. Yu, Z. Shi, and Y. Zhong, "Time-varying formation control for unmanned aerial vehicles: theories and applications," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 1, pp. 340–348, 2014.

[3] X. Hu, Z.-W. Liu, G. Wen, X. Yu, and C. Li, "Branch-wise parallel successive algorithm for online voltage regulation in distribution networks," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6678–6689, 2019.

[4] X. Hu, Z.-W. Liu, G. Wen, X. Yu, and C. Liu, "Voltage control for distribution networks via coordinated regulation of active and reactive power of dgs," *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4017–4031, 2020.

[5] X. Hu, H. Zhou, Z.-W. Liu, X. Yu, and C. Li, "Hierarchical distributed scheme for demand estimation and power reallocation in a future power grid," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 5, pp. 2279–2290, 2017.

[6] Z.-W. Liu, H. Hou, and Y.-W. Wang, "Formation-containment control of multiple underactuated surface vessels with sampling communication via hierarchical sliding mode approach," *ISA Transactions*, 2019.

[7] M. Chi, X.-L. Wang, D.-X. He, and Z.-W. Liu, "Multiconsensus of second-order multiagent networks via pulse-modulated intermittent control," *Complexity*, vol. 2020, Article ID 1059026, 2020.

[8] Z.-W. Liu, G. Wen, X. Yu, Z.-H. Guan, and T. Huang, "Delayed impulsive control for consensus of multiagent systems with switching communication graphs," *IEEE*

Transactions on Cybernetics, vol. 50, no. 7, pp. 3045–3055, 2020.

[9] H. Su, H. Wu, and J. Lam, "Positive edge-consensus for nodal networks via output feedback," *IEEE Transactions on Automatic Control*, vol. 64, no. 3, pp. 1244–1249, 2019.

[10] X. Dong, Y. Zhou, Z. Ren, and Y. Zhong, "Time-varying formation tracking for second-order multi-agent systems subjected to switching topologies with application to quadrotor formation flying," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 6, pp. 5014–5024, 2017.

[11] N. Huang, Z. Duan, and G. Chen, "Some necessary and sufficient conditions for consensus of second-order multi-agent systems with sampled position data," *Automatica*, vol. 63, pp. 148–155, 2016.

[12] Y. Wang, L. Cheng, Z. G. Hou, M. Tan, and M. Wang, "Containment control of multi-agent systems in a noisy communication environment," *Automatica*, vol. 50, no. 7, pp. 1922–1928, 2014.

[13] R. Rahimi, F. Abdollahi, and K. Naqshi, "Time-varying formation control of a collaborative heterogeneous multi agent system," *Robotics and Autonomous Systems*, vol. 62, no. 12, pp. 1799–1805, 2014.

[14] L. Han, X. Dong, Q. Li, and Z. Ren, "Formation-containment control for second-order multi-agent systems with time-varying delays," *Neurocomputing*, vol. 218, pp. 439–447, 2016.

[15] Z.-W. Liu, Z.-H. Guan, X. Shen, and G. Feng, "Consensus of multi-agent networks with aperiodic sampled communication via impulsive algorithms using position-only measurements," *IEEE Transactions on Automatic Control*, vol. 57, no. 10, pp. 2639–2643, 2012.

[16] H. Su, Y. Liu, and Z. Zeng, "Second-order consensus for multiagent systems via intermittent sampled position data control," *IEEE Transactions on Cybernetics*, vol. 50, no. 5, pp. 2063–2072, 2020.

[17] P. C. Parks and V. Hahn, *Stability Theory*, Prentice-Hall, Inc., Hoboken, NJ, USA, 1993.

[18] W. Ren and R. W Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Transactions on Automatic Control*, vol. 50, no. 5, pp. 655–661, 2005.

[19] Z. Meng, W. Ren, and Z. You, "Distributed finite-time attitude containment control for multiple rigid bodies," *Automatica*, vol. 46, no. 12, pp. 2092–2099, 2010.

[20] Z.-W. Liu, X. Yu, Z.-H. Guan, B. Hu, and C. Li, "Pulse-modulated intermittent control in consensus of multiagent systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 5, pp. 783–793, 2017.

Hindawi
Complexity
Volume 2021, Article ID 5525067, 11 pages
https://doi.org/10.1155/2021/5525067

WILEY | Hindawi

Research Article

Disturbance Observer-Based Robust Formation-Containment of Discrete-Time Multiagent Systems with Exogenous Disturbances

Chengjie Xu,[1,2] **Bofan Li**,[1] **and Yi Yuan**[3]

[1]*School of Science, Hunan University of Technology, Zhuzhou 412008, China*
[2]*The Key Laboratory for Electric Drive Control and Intelligent Equipment of Hunan Province, Hunan University of Technology, Zhuzhou 412008, China*
[3]*School of Computer Science, Hunan University of Technology, Zhuzhou 412008, China*

Correspondence should be addressed to Yi Yuan; 13637338678@sina.cn

Received 27 January 2021; Revised 30 March 2021; Accepted 3 April 2021; Published 6 May 2021

Academic Editor: Xiao Ling Wang

This paper investigates robust formation-containment control of discrete-time multiagent systems (MASs) with exogenous disturbances. Based on the discrete-time disturbance observer method, both state feedback and output feedback control protocols are proposed. Formation-containment conditions are obtained and convergency analysis is given according to Lyapunov stability theory. And, the corresponding control gains are obtained by solving some discrete-time algebraic Riccati equations. Numerical simulations are presented to illustrate the theoretical findings.

1. Introduction

Recently, the distributed cooperative control of the MASs has drawn much attention from various disciplines. In distributed cooperative control issues, consensus plays a significant role, which means that the final states of all agents can reach a common value. Due to its widespread applications, many works about consensus have been reported in the past decades [1–6]. Relative research studies refer to synchronization [7, 8], flocking [9], formation [10, 11], and so on.

In recent years, as an important extension of consensus, containment of MASs has been intensively studied for its wide applications in real world. Containment means that there exist multiple leaders in a network, and all followers can asymptotically enter into the convex hull spanned by the leaders. Li et al. [12, 13] addressed the distributed containment control of MASs with general linear dynamics. Ma and Miao [14] proposed the distributed dynamic output feedback control law by using relative output information of neighboring agents. Containment of second-order MASs was studied by using sampled-data position under continuous communication and intermittent communication topology, respectively [15, 16]. Containment was investigated for discrete-time linear MASs with input saturation and intermittent communication [17].

All the works mentioned above do not take the formation of the leaders into account. However, in some real networks, there may be information exchange between the leaders and the leaders will form a special formation, such as robot football. Therefore, formation-containment as a new research topic has attracted much attention, which combines the property of containment control and formation control. Han et al. [18] addressed formation-containment control of second-order dynamics MASs with time-varying. Formation-containment control protocol of high-order linear systems with time-delayed and time-invariant was analysed and designed [19]. In [20], formation-containment control of continuous-time nonlinear Euler–Lagrange MASs with input saturation was studied. Zuo et al. and Wang et al. [21–23] investigated the out formation-containment control, that is, the outputs of all followers converge to the convex hull spanned by the outputs of all leaders, and the outputs of all leaders can maintain a formation structure. Zuo et al. [21] proposed the distributed static and dynamic output feedback control protocols for homogeneous and

heterogeneous MASs with time-varying. Wang et al. [22, 23] used the intermittent control and impulsive control to achieve the output formation-containment control of heterogeneous MASs, respectively.

However, the above papers mainly investigated formation-containment control problem for continuous-time MASs without disturbances. Disturbances often exist and destroy the performance of the controlled systems. Therefore, it is very important and meaningful to research the coordination of MASs with exogenous disturbances. In [24], the finite-time leaderless consensus of double-integrator MASs under the fixed network topology with external bounded disturbances was investigated. The leader-following consensus of first-order and second-order nonlinear MASs with unknown bounded external disturbances were discussed [25]. In [26], identical and nonidentical external disturbances were investigated for the leader-following output consensus of discrete-time linear MASs with input saturation. Containment of continuous-time MASs with exogenous disturbances was investigated in virtue of disturbance observer (DO) technique [27]. Du and Li [28] employed the event-triggered control approach to deal with the robust stabilization problem of delayed systems with parameter uncertainties and exogenous disturbances.

Inspired by the above literatures, this paper concentrates on the formation-containment control of high-order discrete-time MASs with exogenous disturbances. The main contributions are as follows:

(i) Discrete-time MASs is discussed in this paper, which can be used to model more plants under computer control technology. And, the stability analysis is more challenging than the continuous-time MASs.

(ii) Formation-containment of MASs is investigated in this paper. The formation of the leader is considered, which has more applications than ordinary containment problems.

(iii) Exogenous disturbances are considered, and a discrete-time disturbance observer method is proposed for attenuating the disturbances.

The rest of the paper is organized as follows. Section 2 states the model considered in the paper and gives some basic definitions, lemmas, and assumptions. In Section 3, discrete-time DO-based state feedback formation-containment protocol is proposed. In Section 4, discrete-time DO-based output feedback formation-containment protocol is offered. Numerical examples are included to demonstrate the proposed protocols in Section 5. Finally, Section 6 gives a conclusion for this paper.

Notation: R denotes the set of real numbers. $R^{N\times M}$, I_N, and 1_N represent the set of $N \times M$ real matrices, $N \times N$ identity matrix, and the N-dimension column vector that all the elements are 1, respectively. A^T (or x^T) represents the transpose of the matrix A (or the vector x). $\|*\|$ represents the Euclid norm of $*$. $\otimes$ denotes the Kronecker product.

2. Preliminaries

A graph $\mathcal{G} = (V, E, A)$ represents a network topology, which includes a set of nodes $\mathcal{V} = \{1, 2, \ldots, N + M\}$, a set of edges $\mathcal{E} \subseteq V \times V$, and an adjacent matrix $\mathcal{A} = [a_{ij}]$. For a directed graph, $a_{ij} > 0 \Leftrightarrow (j, i) \in \mathcal{E}$, that is, j sends information to i. G is an undirected graph if $a_{ij} = a_{ji}$. $\mathcal{N}_i = \{j | (j, i) \in \mathcal{E}\}$ is a neighbor set of the ith node. $L = D - \mathcal{A} = [l_{ij}]$ is the Laplacian matrix, where $D = \mathrm{diag}(\sum_{j=1, j\neq i}^{N+M} a_{ij})$. Therefore, $l_{ii} = \sum_{j=1, j\neq i}^{N+M} a_{ij}$ and $l_{ij} = -a_{ij}, i \neq j$.

$\mathcal{F} = \{1, 2, \ldots, N\}$ and $\mathcal{L} = \{N + 1, N + 2, \ldots, N + M\}$ represent a set of followers and leaders, respectively. Suppose that the edges among followers and among leaders are undirected, and edges between leaders and followers are directed. Thus, L can be rewritten as

$$\begin{bmatrix} L_1 & L_2 \\ 0 & L_3 \end{bmatrix}. \tag{1}$$

The dynamics of the ith agent are described by

$$\begin{aligned} x_i(k+1) &= Ax_i(k) + B\big(u_i(k) + d_i(k)\big), \\ y_i(k) &= Cx_i(k), \quad i \in \mathcal{V}, \end{aligned} \tag{2}$$

where $x_i \in R^n$, $u_i \in R^m$, $d_i \in R^m$, and $y_i \in R^p$ denote the state, the control input, the exogenous disturbance, and the measurement output of the ith agent, respectively. $A \in R^{n\times n}$, $B \in R^{n\times m}$, and $C \in R^{p\times n}$ are the constant matrices.

It is assumed that the disturbance $\omega_i(k)$, $i = 1, \ldots, N + M$ are generated by the following exogenous system:

$$\begin{aligned} \omega_i(k+1) &= S\omega_i(k), \\ d_i(k) &= F\omega_i(k), \quad i \in \mathcal{V}, \end{aligned} \tag{3}$$

where $\omega_i \in R^l$ is the disturbance state of the ith agent of the exogenous system and $S \in R^{l\times l}$ and $F \in R^{m\times l}$ are the constant matrices of the exogenous system.

Suppose there is a virtual leader in the network, whose dynamics are described as

$$\begin{aligned} x_0(k+1) &= Ax_0(k), \\ y_0(k) &= Cx_0(k), \end{aligned} \tag{4}$$

where $x_0 \in R^n$ and $y_0 \in R^p$ are the state and the measured output of the virtual leader, respectively. Then, one has $h_i(k + 1) = Ah_i(k)$, $i \in \mathcal{L}$, where $h_i \in R^n$ is the desired relative position between the leader i and the virtual leader. Let $h_l(k) = (h_{N+1}^T(k), h_{N+2}^T(k), \ldots, h_{N+M}^T(k)^T)$; then, it follows $h_l(k + 1) = (I_M \otimes A)h_l(k)$.

The following definitions, assumptions, and lemmas are necessary for the main results of this paper.

Definition 1 (see [29]). Let $\mathcal{C}$ be a subset of R^n, the set $\mathcal{C}$ is said to be convex if any x and y in $\mathcal{C}$ and any $\alpha \in [0, 1]$, and the point $(1 - \alpha)x + \alpha y \in \mathcal{C}$. The convex hull of a set of points $X = \{x_1, x_2, \ldots, x_n\}$ is the minimal convex set containing all points in X. We denote the convex hull of X as $\mathrm{Co}(X)$.

Definition 2 (see [23]). The linear MASs (2) is said to achieve the formation-containment if, for any given initial states $x_i(0)$, $i \in \mathscr{V}$, there exist $\lim_{k\longrightarrow\infty}\|x_i(k) - h_i(k) - x_0(k)\| = 0$, $i \in \mathscr{L}$ and $\lim_{k\longrightarrow\infty} \inf_{x_r \in Co(X)} \|x_i(k) - x_r\| = 0$, $i \in \mathscr{F}$.

Assumption 1. Suppose that the edges among the followers are undirected, i.e., all the followers can access each other's information. Moreover, for each follower, there exists at least one leader that has a directed path to that follower.

Assumption 2. The matrix pair (A, B) is stabilizable.

Assumption 3. The matrix pair (A, C) is detectable.

Lemma 1 (see [30]). *Under Assumption 2, there exists a unique positive defined matrix P, satisfying the modified algebraic Riccati equation:*

$$A^T\mathrm{PA} - P - A^T\mathrm{PB}\left(B^T\mathrm{PB} + I\right)^{-1}B^T\mathrm{PA} + I = 0. \tag{5}$$

Lemma 2 (see [12]). *Under Assumption 1, L_1 is positive definite, each entry of $-L_1^{-1}L_2$ is nonnegative, and each row of $-L_1^{-1}L_2$ has a sum equal to one.*

3. DO-Based State Feedback of Formation-Containment

In this section, DO-based distributed state feedback protocol of the formation-containment is proposed and the formation-containment criteria are derived. The dynamics of discrete-time DO based on the state are proposed as follows:

$$\begin{aligned} v_i(k+1) &= (S + \mathrm{HBF})\left(v_i(k) - Hx_i(k)\right) + H\left(Ax_i(k) + Bu_i(k)\right),\\ \widehat{\omega}_i(k) &= v_i(k) - Hx_i(k),\\ \widehat{d}_i(k) &= F\widehat{\omega}_i(k), \quad i \in \mathscr{V}, \end{aligned} \tag{6}$$

where $v_i \in R^l$ is the internal state variable of the observer, $\widehat{d}_i$ and $\widehat{\omega}_i$ are the estimated value of d_i and ω_i, respectively, and $H \in R^{l\times n}$ is the gain matrix of the observer.

Remark 1. The disturbance is assumed to be generated by a exogenous system, which leads to that the agents in the network cannot get the information of the disturbances. For attenuating the disturbance, all agents have to estimate the value of the exogenous disturbances only using the relative state or relative output information.

According to (2) and (6), one has

$$\widehat{\omega}_i(k+1) = (S + \mathrm{HBF})\widehat{\omega}_i(k) - \mathrm{HBF}\omega_i(k). \tag{7}$$

Then, denote $\delta_i(k) = \omega_i(k) - \widehat{\omega}_i(k)$, and from (3) and (7), one has

$$\delta_i(k+1) = (S + \mathrm{HBF})\delta_i(k). \tag{8}$$

Let $\delta_f(k) = \delta_1^T(k), \delta_2^T(k), \ldots, \delta_N^T(k)^T$ and $\delta_l(k) = \delta_{N+1}^T(k), \delta_{N+2}^T(k), \ldots, \delta_{N+M}^T(k)^T$; then, (8) can be rewritten as the following form:

$$\begin{aligned} \delta_f(k+1) &= \left(I_N \otimes (S + \mathrm{HBF})\right)\delta_f(k),\\ \delta_l(k+1) &= \left(I_M \otimes (S + \mathrm{HBF})\right)\delta_l(k). \end{aligned} \tag{9}$$

The distributed DO-based state feedback formation-containment control protocol is proposed as

$$\begin{aligned} u_i(k) &= K_1\left(\sum_{j\in N_i} a_{ij}\left(x_i(k) - x_j(k)\right)\right) - \widehat{d}_i(k), \quad i \in \mathscr{F},\\ u_i(k) &= K_2\left\{\sum_{j\in N_i}\left(a_{ij}\left(x_i(k) - h_i(k) - \left(x_j(k) - h_j(k)\right)\right) + g_i\left(x_i(k) - h_i(k) - x_0(k)\right)\right) - \widehat{d}_i(k), \quad i \in \mathscr{L}, \end{aligned} \tag{10}$$

where K_1 and K_2 are the gain matrices to be designed and g_i represents the information interaction between the leader i and the virtual leader, where $g_i = 1$ if the virtual leader can send information to the leader i, otherwise $g_i = 0$.

Remark 2. The formation-containment protocol depends on the estimated value of d_i and the local relative information, which consists of two parts, the local relative information and the estimated value $\widehat{d}_i$.

Substituting control protocol (10) into system (2), one has

$$\begin{aligned} x_i(k+1) &= Ax_i + BK_1\left(\sum_{j\in N_i} a_{ij}\big(x_i(k)-x_j(k)\big)\right) + \mathrm{BF}\delta_i(k), \quad i\in\mathscr{F},\\ x_i(k+1) &= Ax_i(k) + BK_2\left\{\sum_{j\in N_i}\Big(a_{ij}\big(x_i(k)-h_i(k)-x_j(k)-h_j(k)\big)+g_i\big(x_i(k)-h_i(k)-x_0(k)\big)\Big)\right\} + \mathrm{BF}\delta_i(k), \quad i\in\mathscr{L}. \end{aligned} \tag{11}$$

Let $x_f(k) = x_1^T(k), x_2^T(k), \ldots, x_N^T(k)^T$ and $x_l(k) = x_{N+1}^T(k), x_{N+2}^T(k), \ldots, x_{N+M}^T(k)^T$; then, (11) can be rewritten as follows:

$$\begin{aligned} x_f(k+1) &= \big(I_N\otimes A + L_1\otimes \mathrm{BK}_1\big)x_f(k) + \big(L_2\otimes \mathrm{BK}_1\big)x_l(k) + \big(I_N\otimes \mathrm{BF}\big)\delta_f(k),\\ x_l(k+1) &= \big(I_M\otimes A\big)x_l(k) + \big(J\otimes \mathrm{BK}_2\big)\big(x_l(k)-h_l(k)-\overline{x}_0(k)\big) + \big(I_M\otimes \mathrm{BF}\big)\delta_l(k), \end{aligned} \tag{12}$$

where $J = L_3 + G$ is a positive definite matrix, $G = \mathrm{diag}\{g_{N+1}, g_{N+2}, \ldots, g_{N+M}\}$, and $\overline{x}_0(k) = 1_M\otimes x_0(k)$. On the basis of equation (4), one has $\overline{x}_0(k+1) = (I_M\otimes A)\overline{x}_0(k)$.

The following theorem gives a sufficient and necessary condition to achieve the formation-containment via state feedback protocol.

Theorem 1. *Suppose Assumptions 1 and 2 hold. Under DO-based state feedback protocol (10) with* $K_1 = -1/\lambda_1(B^TPB+I)^{-1}B^TPA$ *and* $K_2 = -1/\mu_1(B^TPB+I)^{-1}B^TPA$, *the formation-containment of system (2) can be realized if and only if* $S + HBF$ *is Schur stable, where* $P>0$ *is the unique solution of the algebraic Riccati equation (5) and* λ_1 *and* μ_1 *are the minimum eigenvalue of* L_1 *and* J, *respectively.*

Proof. Denote the error of the system as $e_i(k) = \sum_{j\in N_i} a_{ij}(x_i(k)-x_j(k))$, $i\in\mathscr{F}$, and $\psi_i(k) = \sum_{j\in N_i}(a_{ij}((x_i(k)-h_i(k)-x_j(k)-h_j(k))) + g_i(x_i(k)-h_i(k)-x_0(k)))$, $i\in\mathscr{L}$. Let $e(k) = e_1^T(k), e_2^T(k), \ldots, e_N^T(k)^T$ and $\psi(k) = \psi_{N+1}^T(k), \psi_{N+2}^T(k), \ldots, \psi_{N+M}^T(k)^T$; then, it follows

$$\begin{aligned} e(k) &= \big(L_1\otimes I_n\big)x_f(k) + \big(L_2\otimes I_n\big)x_l(k),\\ \psi(k) &= \big(J\otimes I_n\big)\big(x_l(k)-h_l(k)-\overline{x}_0(k)\big). \end{aligned} \tag{13}$$

Substituting (12) into $e(k+1)$ and $\psi(k+1)$ of (13), one derives

$$\begin{aligned} e(k+1) &= \big(I_N\otimes A + L_1\otimes \mathrm{BK}_1\big)e(k) + \big(L_2\otimes \mathrm{BK}_2\big)\psi(k) + \big(L_1\otimes \mathrm{BF}\big)\delta_f(k) + \big(L_2\otimes \mathrm{BF}\big)\delta_l(k),\\ \psi(k+1) &= \big(I_M\otimes A + J\otimes \mathrm{BK}_2\big)\psi(k) + (J\otimes \mathrm{BF})\delta_l(k). \end{aligned} \tag{14}$$

Then, the error system can be rewritten as the following form by (9) and (14):

$$\begin{pmatrix} e(k+1)\\ \psi(k+1)\\ \delta_f(k+1)\\ \delta_l(k+1)\end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14}\\ 0 & A_{22} & 0 & A_{24}\\ 0 & 0 & A_{33} & 0\\ 0 & 0 & 0 & A_{44}\end{pmatrix}\begin{pmatrix} e(k)\\ \psi(k)\\ \delta_f(k)\\ \delta_l(k)\end{pmatrix}, \tag{15}$$

where $A_{11} = I_N\otimes A + L_1\otimes \mathrm{BK}_1$, $A_{12} = L_2\otimes \mathrm{BK}_2$, $A_{13} = L_1\otimes \mathrm{BF}$, $A_{14} = L_2\otimes \mathrm{BF}$, $A_{22} = I_M\otimes A + J\otimes \mathrm{BK}_2$, $A_{24} = J\otimes \mathrm{BF}$, $A_{33} = I_N\otimes(S+\mathrm{HBF})$, and $A_{44} = I_M\otimes(S+\mathrm{HBF})$.

(15) is Schur stable if and only if A_{11}, A_{22}, A_{33}, and A_{44} are Schur stable. On the one side, $A_{11} = I_N\otimes A + L_1\otimes \mathrm{BK}$ and $A_{22} = I_M\otimes A + J\otimes \mathrm{BK}_2$ are Schur stable if and only if $A+\lambda_i\mathrm{BK}_1$ and $A+\mu_j\mathrm{BK}_2$ are Schur stable, where λ_i, $i=1,\ldots,N$, and μ_j, $j=1,\ldots,M$, are the eigenvalues of L_1 and J, respectively. According to Lemma 2, under Assumption 1, L_1 is positive definite and of course is nonsingular. Then, one has $\lambda_i>0$, $i=1,\ldots,N$. Consider the following discrete-time system as $x(k+1) = (A+\lambda_i\mathrm{BK}_1)x(k)$ with $K_1 = -1/\lambda_1(B^T\mathrm{PB}+I)^{-1}B^T\mathrm{PA}$, where λ_1 is the minimum eigenvalue of L_1. Choosing the discrete-time Lyapunov function as $V(k) = x^T(k)Px(k)$, where P is the unique solution of the modified discrete-time algebra Riccati equation (5), then one has

$$\begin{aligned} V(k+1)-V(k) &= x^T(k+1)Px(k+1)-x^T(k)Px(k)\\ &= x^T(k)\Big(\big(A+\lambda_i \mathrm{BK}_1\big)^T P\big(A+\lambda_i \mathrm{BK}_1\big)-P\Big)x(k)\\ &= x^T(k)\Big(A^T\mathrm{PA}+2\lambda_i A^T\mathrm{PBK}_1+\lambda_i^2 K_1^T B^T\mathrm{PBK}_1-P\Big)x(k)\\ &\le x^T(k)\Big(A^T\mathrm{PA}-P-2A^T\mathrm{PB}\big(B^T\mathrm{PB}+I\big)^{-1}B^T\mathrm{PA}\Big)\\ &\Big(+A^TPB\big(B^T\mathrm{PB}+I\big)^{-1}B^T\mathrm{PB}\big(B^T\mathrm{PB}+I\big)^{-1}B^T\mathrm{PA}\Big)x(k)\\ &\le x^T(k)\Big(A^TPA-P-A^T\mathrm{PB}\big(B^T\mathrm{PB}+I\big)^{-1}B^T\mathrm{PA}\Big)x(k)\\ &= -x^T(k)x(k).\\ &< 0. \end{aligned} \tag{16}$$

Hence, $V(k)\longrightarrow 0$ as $k\longrightarrow\infty$, and then, one can conclude that $A+\lambda_i \mathrm{BK}$ is Schur stable for $i=1,\ldots,N$. This leads to that $A_{11}=I_N\otimes A+L_1\otimes \mathrm{BK}$ is Schur stable. Similarly, one has $A_{22}=I_M\otimes A+J\otimes \mathrm{BK}_2$ is Schur stable with $K_2=-1/\mu_1(B^T\mathrm{PB}+I)^{-1}B^T\mathrm{PA}$, where μ_1 is the minimum eigenvalue of J.

On the other side, $S+\mathrm{HBF}$ is Schur stable if and only if $A_{33}=I_N\otimes(S+\mathrm{HBF})$ and $A_{44}=I_M\otimes(S+\mathrm{HBF})$ are Schur stable. Because A_{11}, A_{22}, A_{33}, and A_{44} are Schur stable, the error system (15) is Schur stable. Therefore, $\lim_{k\longrightarrow\infty}\|e(k)\|=0$, $\lim_{k\longrightarrow\infty}\|\psi(k)\|=0$, $\lim_{k\longrightarrow\infty}\|\delta_f(k)\|=0$, and $\lim_{k\longrightarrow\infty}\|\delta_l(k)\|=0$, that is, $\lim_{k\longrightarrow\infty}\|x_f(k)+(L_1^{-1}L_2\otimes I_n)x_l(k)\|=0$ and $\lim_{k\longrightarrow\infty}\|x_l(k)-h_l(k)-\overline{x}_0(k)\|=0$; thus, $\lim_{k\longrightarrow\infty}\inf_{x_r\in Co(X)}\|x_i(k)-x_r\|=0$, $i\in\mathscr{F}$, and $\lim_{k\longrightarrow\infty}\|x_i(k)-h_i(k)-x_0(k)\|=0$, $i\in\mathscr{L}$, on the basis of Definition 1, which means Theorem 1 holds. □

4. DO-Based Output Feedback of Formation-Containment

When the state of all agents cannot be obtained, the state observer can be used to estimate the state. Therefore, in this section, the discrete-time state observer and disturbance observer are given. Meanwhile, DO-based output feedback formation-containment protocol is proposed.

The state observer is designed as

$$\begin{aligned} \hat{x}_i(k+1) &= A\hat{x}_i(k)+B\big(u_i(k)+d_i(k)\big)+D\big(C\hat{x}_i(k)-y_i(k)\big),\\ \hat{y}_i(k) &= C\hat{x}_i(k),\quad i\in\mathscr{V}, \end{aligned} \tag{17}$$

where $\hat{x}_i\in R^n$ and $\hat{y}_i\in R^p$ are the estimated value of x_i and y_i, respectively, $D\in R^{n\times p}$ is the constant matrix, and the rest variable are same as the previous part.

The discrete-time disturbance observer based on output information is proposed as

$$\begin{aligned} v_i(k+1) &= (S+\mathrm{HBF})\big(v_i(k)-H\hat{x}_i(k)\big)+H\big(A\hat{x}_i(k)+Bu_i(k)+D\big(C\hat{x}_i(k)-y_i(k)\big)\big),\\ \hat{\omega}_i(k) &= v_i(k)-H\hat{x}_i(k),\\ \hat{d}_i(k) &= F\hat{\omega}_i(k), \end{aligned} \tag{18}$$

where $v_i\in R^l$ is the internal state variable of the observer, $\hat{\omega}_i$ and $\hat{d}_i$ are the estimated value of ω_i and d_i, respectively, and $H\in R^{l\times n}$ is the gain matrix of the observer.

Remark 3. For the case that the state of each agent cannot be obtained, the state observer can be used to estimate the state. Moreover, the disturbances exist in the subsystems. One has to design corresponding controller to attenuate the disturbances. Discrete-time output-based disturbance observer (18) is proposed.

Denoting $\delta_i(k)=\omega_i(k)-\hat{\omega}i(k)$, by (3) and (18), one has

$$\delta_i(k+1)=(S+\mathrm{HBF})\delta_i(k). \tag{19}$$

Let $\delta_f(k)=\delta_1^T(k),\delta_2^T(k),\ldots,\delta_N^T(k)^T$ and $\delta_l(k)=\delta_{N+1}^T(k),\delta_{N+2}^T(k),\ldots,\delta_{N+M}^T(k)^T$; then, $\delta_i(k+1)$ of (19) can be rewritten as follows:

$$\begin{aligned} \delta_f(k+1) &= \big(I_N\otimes(S+\mathrm{HBF})\big)\delta_f(k),\\ \delta_l(k+1) &= \big(I_M\otimes(S+\mathrm{HBF})\big)\delta_l(k). \end{aligned} \tag{20}$$

And, the formation-containment control protocol depends on output information and can be designed as

$$u_i(k) = K_3\left(\sum_{j\in N_i} a_{ij}\left(\widehat{x}_i(k) - \widehat{x}_j(k)\right)\right) - \widehat{d}_i(k), \quad i \in \mathscr{F},$$
$$u_i(k) = K_4\left\{\sum_{j\in N_i}\left(a_{ij}\left(\widehat{x}_i(k) - h_i(k) - \widehat{x}_j(k) - h_j(k)\right) + g_i\left(\widehat{x}_i(k) - h_i(k) - x_0(k)\right)\right)\right\} - \widehat{d}_i(k), \quad i \in \mathscr{L}, \tag{21}$$

where K_3 and K_4 are the gain matrices to be designed. g_i is mentioned in the previous section.

Substituting output feedback law (21) into the system of state observer (17) and denoting $\widetilde{x}_i(k) = x_i(k) - \widehat{x}_i(k)$, $i \in \mathscr{V}$, as the error of the real value and the estimated value of the state of the ith agent, it follows

$$\widehat{x}_i(k+1) = A\widehat{x}_i(k) + \mathrm{BK}_3\left(\sum_{j\in N_i} a_{ij}\left(\widehat{x}_i(k) - \widehat{x}_j(k)\right)\right) + \mathrm{BF}\delta_i(k) - \mathrm{DC}\widetilde{x}_i(k), \quad i \in \mathscr{F},$$
$$\widehat{x}_i(k+1) = A\widehat{x}_i(k) + \mathrm{BK}_4\left\{\sum_{j\in N_i}\left(a_{ij}\left(\widehat{x}_i(k) - h_i(k) - \widehat{x}_j(k) - h_j(k)\right) + g_i\left(\widehat{x}_i(k) - h_i(k) - x_0(k)\right)\right)\right\} + \mathrm{BF}\delta_i(k) - \mathrm{DC}\widetilde{x}_i(k), \quad i \in \mathscr{L},$$
$$\widetilde{x}_i(k+1) = (A + \mathrm{DC})\widetilde{x}_i(k), \quad i \in \mathscr{V}. \tag{22}$$

Let $\widehat{x}_f(k) = \widehat{x}_1^T(k), \widehat{x}_2^T(k), \ldots, \widehat{x}_N^T(k)^T$, $\widehat{x}_l(k) = \widehat{x}_{N+1}^T(k), \widehat{x}_{N+2}^T(k), \ldots, \widehat{x}_{N+M}^T(k)^T$, $\widetilde{x}_f(k) = \widetilde{x}_1^T(k), \widetilde{x}_2^T(k), \ldots, \widetilde{x}_N^T(k)^T$, and $\widetilde{x}_l(k) = \widetilde{x}_{N+1}^T(k), \widetilde{x}_{N+2}^T(k), \ldots, \widetilde{x}_{N+M}^T(k)^T$; then, (22) can be rewritten as follows:

$$\begin{aligned}
\widehat{x}_f(k+1) &= \left(I_N \otimes A + L_1 \otimes \mathrm{BK}_3\right)\widehat{x}_f(k) + \left(L_2 \otimes \mathrm{BK}_3\right)\widehat{x}_l(k) + \left(I_N \otimes \mathrm{BF}\right)\delta_f(k) - \left(I_N \otimes \mathrm{DC}\right)\widetilde{x}_f(k),\\
\widehat{x}_l(k+1) &= \left(I_M \otimes A\right)\widehat{x}_l(k) + \left(J \otimes \mathrm{BK}_4\right)\left(\widehat{x}_l(k) - \widehat{h}_l(k) - \overline{x}_0(k)\right) + \left(I_M \otimes \mathrm{BF}\right)\delta_l(k) - \left(I_M \otimes \mathrm{DC}\right)\widetilde{x}_l(k),\\
\widetilde{x}_f(k+1) &= \left(I_N \otimes (A + \mathrm{DC})\right)\widetilde{x}_f(k),\\
\widetilde{x}_l(k+1) &= \left(I_M \otimes (A + \mathrm{DC})\right)\widetilde{x}_l(k),
\end{aligned} \tag{23}$$

where related concept of J and $\overline{x}_0(k)$ are the same as Section 3.

A sufficient and necessary condition to achieve the formation-containment via output feedback law is given by the following theorem.

Theorem 2. *Suppose Assumptions 1–3 hold. Under DO-based output feedback protocol (21) with* $K_3 = -1/\lambda_1 (B^T PB + I)^{-1} B^T PA$ *and* $K_4 = -1/\mu_1 (B^T PB + I)^{-1} B^T PA$, *the formation-containment of system (2) can be realized if and only if both* $S + HBF$ *and* $A + DC$ *are Schur stable, where* $P > 0$ *is the unique solution of the algebraic Riccati equation (5) and* λ_1 *and* μ_1 *are the minimum eigenvalue of* L_1 *and* J, *respectively.*

Proof. Consider the error of the state observer system as $\widehat{e}_i(k) = \sum_{j\in N_i} a_{ij}(\widehat{x}_i(k) - \widehat{x}_j(k))$, $i \in \mathscr{F}$, and $\widehat{\psi}_i(k) = \sum_{j\in N_i}(a_{ij}(\widehat{x}_i(k) - h_i(k) - \widehat{x}_j(k) - h_j(k)) + g_i(\widehat{x}_i(k) - h_i(k) - x_0(k)))$, $i \in \mathscr{L}$. Let $\widehat{e}(k) = \widehat{e}_1^T(k), \widehat{e}_2^T(k), \ldots, \widehat{e}_N^T(k)^T$ and $\widehat{\psi}(k) = \widehat{\psi}_{N+1}^T(k), \widehat{\psi}_{N+2}^T(k), \ldots, \widehat{\psi}_{N+M}^T(k)^T$; then, one has

$$\begin{aligned}
\widehat{e}(k) &= \left(L_1 \otimes I_n\right)\widehat{x}_f(k) + \left(L_2 \otimes I_n\right)\widehat{x}_l(k),\\
\widehat{\psi}(k) &= \left(J \otimes I_n\right)\left(x_l(k) - h_l(k) - \overline{x}_0(k)\right).
\end{aligned} \tag{24}$$

Substituting (23) into (24), then $\widehat{e}(k+1)$ and $\widehat{\psi}(k+1)$ have the following form:

$$\begin{aligned}\widehat{e}(k+1) &= \left(I_N \otimes A + L_1 \otimes \mathrm{BK}_3\right)\widehat{e}(k) + \left(L_2 \otimes \mathrm{BK}_4\right)\widehat{\psi}(k) + \left(L_1 \otimes \mathrm{BF}\right)\delta_f(k) + \left(L_2 \otimes \mathrm{BF}\right)\delta_l(k) - \left(L_1 \otimes \mathrm{DC}\right)\widetilde{x}_f(k) - \left(L_2 \otimes \mathrm{DC}\right)\widetilde{x}_l(k),\\ \widehat{\psi}(k+1) &= \left(I_M \otimes A + J \otimes \mathrm{BK}_4\right)\psi(k) + \left(J \otimes \mathrm{BF}\right)\delta_l(k) - \left(J \otimes \mathrm{DC}\right)\widetilde{x}_l(k).\end{aligned} \tag{25}$$

According to (20), (23), and (25), the error system can be rewritten as the form of matrix multiplication:

$$\begin{pmatrix} \widehat{e}(k+1) \\ \widehat{\psi}(k+1) \\ \delta_f(k+1) \\ \delta_l(k+1) \\ \widetilde{x}_f(k+1) \\ \widetilde{x}_l(k+1) \end{pmatrix} = \begin{pmatrix} B_{11} & B_{12} & B_{13} & B_{14} & B_{15} & B_{16} \\ 0 & B_{22} & 0 & B_{24} & 0 & B_{26} \\ 0 & 0 & B_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & B_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & B_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & B_{66} \end{pmatrix} \begin{pmatrix} \widehat{e}(k) \\ \widehat{\psi}(k) \\ \delta_f(k) \\ \delta_l(k) \\ \widetilde{x}_f(k) \\ \widetilde{x}_l(k) \end{pmatrix}, \tag{26}$$

where $B_{11} = I_N \otimes A + L_1 \otimes \mathrm{BK}_3$, $B_{12} = L_2 \otimes \mathrm{BK}_4$, $B_{13} = L_1 \otimes \mathrm{BF}$, $B_{14} = L_2 \otimes \mathrm{BF}$, $B_{15} = -L_1 \otimes \mathrm{DC}$, $B_{16} = -L_2 \otimes \mathrm{DC}$, $B_{22} = I_M \otimes A + J \otimes \mathrm{BK}_4$, $B_{24} = J \otimes \mathrm{BF}$, $B_{26} = -J \otimes \mathrm{DC}$, $B_{33} = I_N \otimes (S + \mathrm{HBF})$, $B_{44} = I_M \otimes (S + \mathrm{HBF})$, $B_{55} = I_N \otimes (A + \mathrm{DC})$, and $B_{66} = I_M \otimes (A + \mathrm{DC})$.

(26) is Schur stable if and only if B_{11}, B_{22}, B_{33}, B_{44}, B_{55}, and B_{66} are Schur stable. According to (16) of Theorem 1, we similarly have $B_{11} = I_N \otimes A + L_1 \otimes \mathrm{BK}_3$ with $K_3 = -1/\lambda_1 (B^T \mathrm{PB} + I)^{-1} B^T \mathrm{PA}$ and $B_{22} = I_M \otimes A + J \otimes \mathrm{BK}_4$ with $K_4 = -1/\mu_1 (B^T \mathrm{PB} + I)^{-1} B^T \mathrm{PA}$ are Schur stable. And, $S + \mathrm{HBF}$ and $A + \mathrm{DC}$ are Schur stable if and only if $B_{33} = I_N \otimes (S + \mathrm{HBF})$, $B_{44} = I_M \otimes (S + \mathrm{HBF})$, $B_{55} = I_N \otimes (A + \mathrm{DC})$, and $B_{66} = I_M \otimes (A + \mathrm{DC})$ are Schur stable.

By the above analysis, it is easy to say that (26) is Schur stable, which means that $\lim_{k\longrightarrow\infty}\|\widehat{e}(k)\| = 0$, $\lim_{k\longrightarrow\infty}\|\widehat{\psi}(k)\| = 0$, $lim_{k\longrightarrow\infty}\|\delta_f(k)\| = 0$, $\lim_{k\longrightarrow\infty}\|\delta_l(k)\| = 0$, $\lim_{k\longrightarrow\infty}\|\widetilde{x}_f(k)\| = 0$, and $\lim_{k\longrightarrow\infty}\|\widetilde{x}_l(k)\| = 0$; by the definitions of $\widehat{e}(k)$, $\widehat{\psi}(k)$, $\widetilde{x}_f(k)$, and $\widetilde{x}_l(k)$, one has $\lim_{k\longrightarrow\infty}\|\widetilde{x}_f(k) + (L_1^{-1}L_2 \otimes I_n)\widetilde{x}_l(k)\| = 0$, $\lim_{k\longrightarrow\infty}\|\widetilde{x}_l(k) - h_l(k) - \overline{x}_0(k)\| = 0$, $\lim_{k\longrightarrow\infty}\|x_i(k) - \widehat{x}_i(k)\| = 0$, $i \in \mathscr{V}$, that is, $\lim_{k\longrightarrow\infty}\inf_{x_r \in Co(X)}\|x_i(k) - x_r\| = 0$, $i \in \mathscr{F}$, and $\lim_{k\longrightarrow\infty}\|x_i(k) - h_i(k) - x_0(k)\| = 0$, $i \in \mathscr{L}$. Therefore, the formation-containment via output feedback is achieved, which means Theorem 2 holds. □

5. Simulations

In this section, numerical simulations are given to illustrate the theoretical results. The topology of the network is shown as the graph in Figure 1, which consists of six followers labeled as $\mathscr{F} = \{1, 2, \ldots, 6\}$, three leaders labeled as $\mathscr{L} = \{7, 8, 9\}$, and one virtual leader labeled as 0.

Choose the system matrices A, B, C, D, F, and S as follows:

$$\begin{aligned} A &= \begin{bmatrix} 1.0 & -2.0 \\ 0 & -1.0 \end{bmatrix}, B = \begin{bmatrix} 0.03 \\ -0.05 \end{bmatrix}, \\ C &= \begin{bmatrix} 0 & -1.0 \\ 1.0 & 2.0 \end{bmatrix}, D = \begin{bmatrix} 0.3 & 0.2 \\ 0.2 & 0.3 \end{bmatrix}, \\ F &= \begin{bmatrix} -0.02 & 1.8 \end{bmatrix}, S = \begin{bmatrix} -1.0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned} \tag{27}$$

Then, one has the following dynamics: $i \in \mathscr{V} = \{1, 2, \ldots, 8, 9\}$, $x_{i1}(k+1) = x_{i1}(k) - 2x_{i2}(k) + 0.03(u_i(k) + d_i(k))$ and $x_{i2}(k+1) = -x_{i2}(k) - 0.05(u_i(k) + d_i(k))$. For virtual formation structure, $x_{01}(k+1) = x_{01}(k) - 2x_{02}(k)$, $x_{02}(k+1) = -x_{02}(k)$, and $i \in \mathscr{L}$, $h_{i1}(k+1) = h_{i1}(k) - 2h_{i2}(k)$ and $h_{i2}(k+1) = -h_{i2}(k)$. It is easy to verify that (A, B) is stabilizable, and (A, C) is detectable. The eigenvalues of A are 1.0 and −1.0, respectively. The eigenvalues of S are −1.0 and 0, respectively, which means that A and S are not Schur stable. Choose the gain matrix as

$$H = \begin{bmatrix} -0.2 & 0 \\ 0 & -0.5 \end{bmatrix}. \tag{28}$$

And, the eigenvalues of $A + \mathrm{DC}$ are −0.7899 and 0.1899, respectively, and the eigenvalues of $S + \mathrm{HBF}$ are −0.9999 and 0.045, respectively, which means that both $A + \mathrm{DC}$ and $S + \mathrm{HBF}$ are Schur stable. Hence, the conditions of Theorem 1 and Theorem 2 are satisfied. According to algebraic Ricatti equation (5), one has a positive definite matrix as follows:

$$P = \begin{bmatrix} 13.3479 & -13.5550 \\ -13.5550 & 43.9266 \end{bmatrix}. \tag{29}$$

Then, one can obtain the corresponding controller gain matrices as $K_1 = [-0.8931, -0.3699]$, $K_2 = [-1.9822, -0.8210]$, $K_3 = [-0.8931, -0.3699]$, and $K_4 = [-1.9822, -0.8210]$. By state feedback, one can have Figures 2–4, and by output feedback, one has Figures 5–7. According to Figures 2 and 5, one can know all followers converge to the convex hull spanned by the leaders, which means $\lim_{k\longrightarrow\infty}\inf_{x_r \in Co(X)}\|x_i(k) - x_r\| = 0$, $i \in \mathscr{F}$. In the figures, red lines denote the leaders' trajectories. Figures 3 and 6 denote $\lim_{k\longrightarrow\infty}\|x_i(k) - h_i(k) - x_0(k)\| = 0$, $i \in \mathscr{L}$. By Figures 4 and 7, we can obtain three leaders which can maintain a triangular structure. Form which, the formation-containment control via state feedback and output feedback control protocols are achieved.

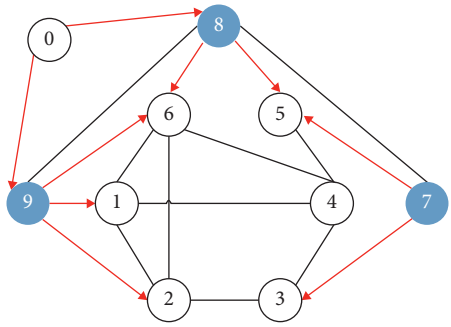


Figure 1: Topology graph.

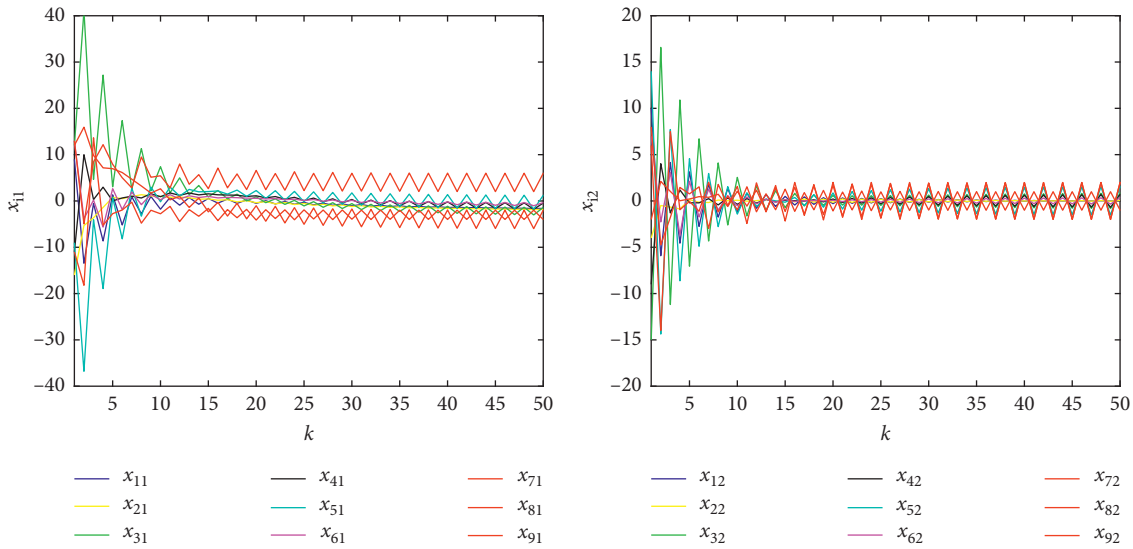


Figure 2: The trajectories of the state x_i, $i = \{1, 2, \ldots, 9\}$, via state feedback.

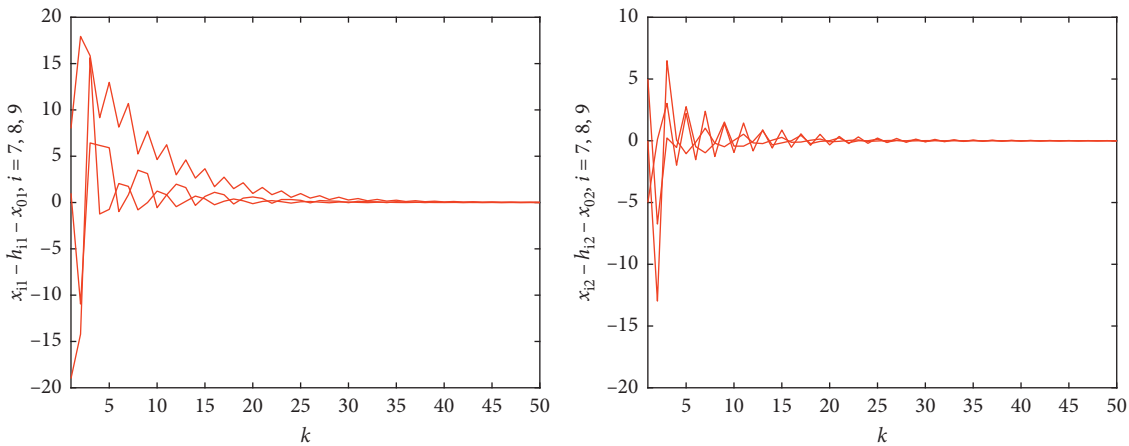


Figure 3: The trajectories of $x_i - h_i - x_0$, $i = \{7, 8, 9\}$, via state feedback.

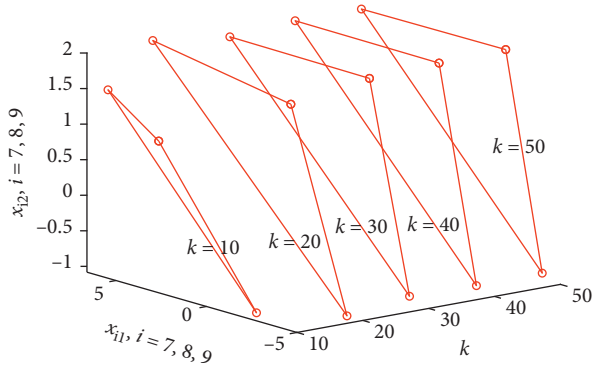


Figure 4: The formation structure of three leaders via state feedback.

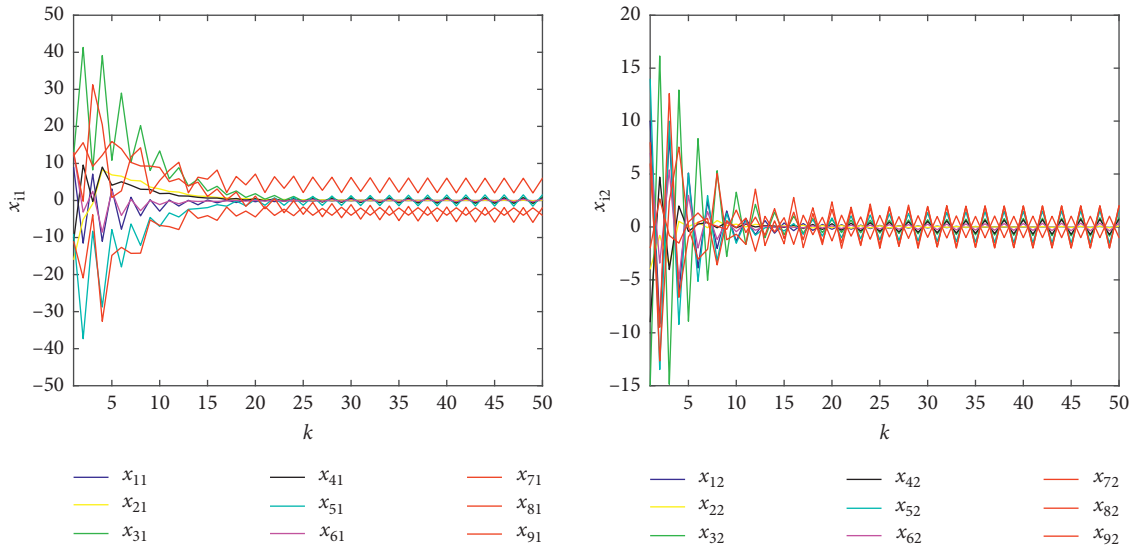


Figure 5: The trajectories of the state x_i, $i = \{1, 2, \ldots, 9\}$. via output feedback.

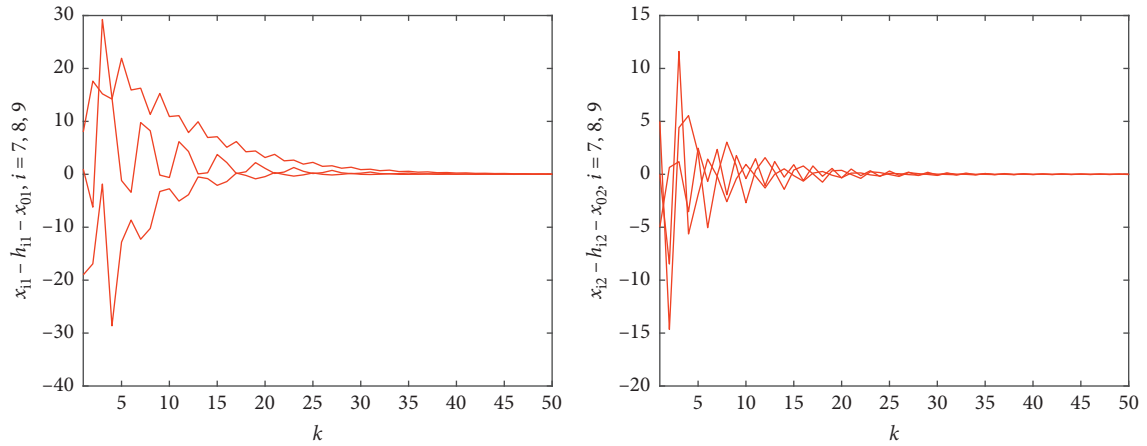


Figure 6: The trajectories of the $x_i - h_i - x_0$, $i = \{7, 8, 9\}$, via output feedback.

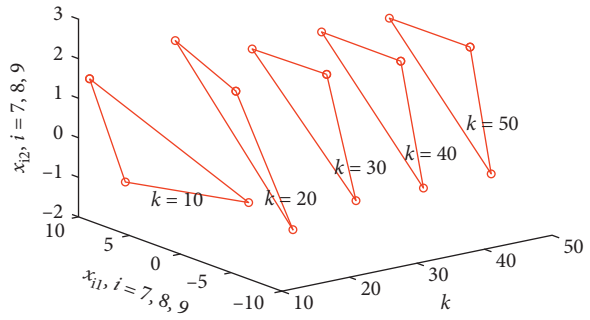


Figure 7: The formation structure of three leaders via output feedback.

6. Conclusions

In this paper, formation-containment control is investigated for high-order discrete-time MASs with exogenous disturbances. Two protocols are proposed via DO-based state feedback and output feedback, respectively. Formation-containment conditions are available, formation-containment analysis is given, and the gain matrices are obtained by solving the discrete-time algebraic Riccai equations. Future works will focus on the output formation-containment of heterogeneous MASs.

Data Availability

No data were used to support the findings of the study.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of the study.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China under Grant no. 61703154, Natural Science Foundation of Hunan Province (2019JJ60061), Scientific Research Foundation of Hunan Provincial Education Department (19C0601), and China Postdoctoral Science Foundation (2016M602298).

References

[1] X. Wang, X. Wang, H. Su, and L. James, "Coordination control for uncertain networked systems using interval observers," *Institute of Electrical and Electronics Engineers Transactions on Cybernetics*, vol. 50, no. 9, pp. 4008–4019, 2020.

[2] C. Xu, H. Xu, H. Su, and C. Liu, "Disturbance-observer based consensus of linear multi-agent systems with exogenous disturbance under intermittent communication," *Neurocomputing*, vol. 404, pp. 26–33, 2020.

[3] J. Qu, Z. Ji, C. Lin, and H. Yu, "Fast consensus seeking on networks with antagonistic interactions," *Complexity*, vol. 201815 pages, 2018.

[4] X. Wang and H. Su, "Completely model-free RL-based consensus of continuous-time multi-agent systems," *Applied Mathematics and Computation*, vol. 382, Article ID 125312, 2020.

[5] H. Su, Y. Liu, and Z. Zeng, "Second-order consensus for multiagent systems via intermittent sampled position data control," *Institute of Electrical and Electronics Engineers Transactions on Cybernetics*, vol. 50, no. 5, pp. 2063–2072, 2020.

[6] X. Wang, G.-P. Jiang, H. Su, and Z. Zeng, "Consensus-based distributed reduced-order observer design for LTI systems," *Institute of Electrical and Electronics Engineers Transactions on Cybernetics*, vol. 202011 pages, 2020.

[7] C. Xu, Y. Zhao, B. Qin, and H. Zhang, "Adaptive synchronization of coupled harmonic oscillators under switching topology," *Journal of the Franklin Institute*, vol. 356, no. 2, pp. 1067–1087, 2019.

[8] C. Xu, H. Su, C. Liu, and G. Zhang, "Robust adaptive synchronization of complex network with bounded disturbances," *Advances in Difference Equations*, vol. 2019, p. 483, 2019.

[9] S. Chen, H. Pei, Q. Lai, and H. Yan, "Multitarget tracking control for coupled heterogeneous inertial agents systems based on flocking behavior," *Institute of Electrical and Electronics Engineers Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 12, pp. 2605–2611, 2019.

[10] Z. Lin, L. Wang, Z. Han, and M. Fu, "Distributed formation control of multi-agent systems using complex Laplacian," *Institute of Electrical and Electronics Engineers Transactions on Automatic Control*, vol. 59, no. 7, pp. 1765–1777, 2014.

[11] X. Ge and Q.-L. Han, "Distributed formation control of networked multi-agent systems using a dynamic event-triggered communication mechanism," *Institute of Electrical and Electronics Engineers Transactions on Industrial Electronics*, vol. 64, no. 10, pp. 8118–8127, 2017.

[12] Z. Li, W. Ren, X. Liu, and M. Fu, "Distributed containment control of multi-agent systems with general linear dynamics in the presence of multiple leaders," *International Journal of Robust and Nonlinear Control*, vol. 23, no. 5, pp. 534–547, 2013.

[13] B. Li, H.-Y. Yang, Z.-Q. Chen, and Z.-X. Liu, "Distributed containment control of multi-agent systems with general linear dynamics and time-delays," *International Journal of Control, Automation and Systems*, vol. 16, no. 6, pp. 2718–2726, 2018.

[14] Q. Ma and G. Miao, "Distributed containment control of linear multi-agent systems," *Neurocomputing*, vol. 133, pp. 399–403, 2014.

[15] Y. Liu and H. Su, "Some necessary and sufficient conditions for containment of second-order multi-agent systems with sampled position data," *Neurocomputing*, vol. 378, pp. 228–237, 2020.

[16] Y. Liu and H. Su, "Containment control of second-order multi-agent systems via intermittent sampled position data communication," *Applied Mathematics and Computation*, vol. 362, Article ID 124522, 2019.

[17] C. Xu, B. Li, and L. Yang, "Semi-global containment of discrete-time high-order multi-agent systems with input saturation via intermittent control," *IET Control Theory & Applications*, vol. 14, no. 16, pp. 2303–2309, 2020.

[18] L. Han, X. Dong, Q. Li, and Z. Ren, "Formation-containment control for second-order multi-agent systems with time-varying delays," *Neurocomputing*, vol. 218, pp. 439–447, 2016.

[19] X. Dong, Z. Shi, G. Lu, and Y. Zhong, "Formation-containment analysis and design for high-order linear time-invariant

swarm systems," *International Journal of Robust and Nonlinear Control*, vol. 25, no. 17, pp. 3439–3456, 2015.

[20] C. Li, L. Chen, Y. Guo, and G. Ma, "Formation-containment control for networked Euler-Lagrange systems with input saturation," *Nonlinear Dynamics*, vol. 91, no. 2, pp. 1307–1320, 2018.

[21] S. Zuo, Y. Song, F. L. Lewis, and A. Davoudi, "Time-varying output formation containment of general linear homogeneous and heterogeneous multiagent systems," *Institute of Electrical and Electronics Engineers Transactions on Control of Network Systems*, vol. 6, no. 2, pp. 537–548, 2018.

[22] Y.-W. Wang, X.-K. Liu, J.-W. Xiao, and X. Lin, "Output formation-containment of coupled heterogeneous linear systems under intermittent communication," *Journal of the Franklin Institute*, vol. 354, no. 1, pp. 392–414, 2017.

[23] Y.-W. Wang, X.-K. Liu, J.-W. Xiao, and Y. Shen, "Output formation-containment of interacted heterogeneous linear systems by distributed hybrid active control," *Automatica*, vol. 93, pp. 26–32, 2018.

[24] X. He and Q. Wang, "Distributed finite-time leaderless consensus control for double-integrator multi-agent systems with external disturbances," *Applied Mathematics and Computation*, vol. 295, pp. 65–76, 2017.

[25] M.-C. Fan and Y. Wu, "Global leader-following consensus of nonlinear multi-agent systems with unknown control directions and unknown external disturbances," *Applied Mathematics and Computation*, vol. 331, pp. 274–286, 2018.

[26] H. Su, Y. Qiu, and L. Wang, "Semi-global output consensus of discrete-time multi-agent systems with input saturation and external disturbances," *ISA Transactions*, vol. 67, pp. 131–139, 2017.

[27] C. Xu, Y. Zheng, H. Su, and H.-B. Zeng, "Containment for linear multi-agent systems with exogenous disturbances," *Neurocomputing*, vol. 160, pp. 206–212, 2015.

[28] Y. Du and B. Li, "Event-based robust stabilization for delayed systems with parameter uncertainties and exogenous disturbances," *Advances in Difference Equations*, vol. 2020, no. 1, 16 pages, 2020.

[29] Z. Zhao, W. Yang, and H. Shi, "Semi-global containment control of discrete-time linear systems with actuator position and rate saturation," *Neurocomputing*, vol. 349, no. 15, pp. 173–182, 2019.

[30] Z. Lin, *Low Gain Feedback, Lecture Notes in Control and Information Sciences*, Vol. 240, Springer, London, UK, 1998.

Hindawi
Complexity
Volume 2021, Article ID 5517062, 15 pages
https://doi.org/10.1155/2021/5517062

Research Article

SPA-Based Modified Local Reachability Density Ratio wSVDD for Nonlinear Multimode Process Monitoring

Zhaojing Wang, Weidong Yang, Hong Zhang, and Ying Zheng

The Key Laboratory of Image Information Processing and Intelligent Control, School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430074, China

Correspondence should be addressed to Weidong Yang; yangwd@mail.hust.edu.cn

Received 7 February 2021; Revised 15 March 2021; Accepted 2 April 2021; Published 21 April 2021

Academic Editor: Xiao Ling Wang

Many industrial processes are operated in multiple modes due to different manufacturing strategies. Multimodality of process data is often accompanied with nonlinear and non-Gaussian characteristics, which makes data-driven monitoring more complicated. In this paper, statistics pattern analysis (SPA) is introduced to extract low- and high-order statistics from raw process data. Support vector data description (SVDD), which can deal with nonlinear and non-Gaussian problems, is applied to monitor multimode process in this paper. To improve detection performance of SVDD for training multimode data with outliers, modified local reachability density ratio (mLRDR) is proposed as a weight factor to be embedded in the weighted-SVDD (wSVDD) model, in which the local neighbors in terms of both space and time are considered. Finally, the effectiveness and superiority of our proposed method are demonstrated by the Tennessee-Eastman (TE) process and wastewater treatment process (WWTP).

1. Introduction

The operation conditions of industrial processes will inevitably change with diverse customer requirements, set-points variation, and different intrinsic features, which leads to multiple modes [1–3]. Because of the complexity of multimodal process, it is difficult to obtain satisfactory monitoring results. In the past decades, effective monitoring of multimode processes has attracted a lot of attention.

The data-driven monitoring methods of multimode process were recently reviewed extensively by Quinones-Grueiro et al. [4]. They indicated that multimodel scheme is one of the widely applied monitoring methods, in which the identification of different modes should be performed in advance. For instance, Du et al. [5] applied K-means to address the multimode clustering problem. Similarly, Zhang and Zhao [6] clustered multimode data into different clusters through the kernel fuzzy c-means (KFCM) algorithm. In these methods, the cluster number should be confirmed in advance. As an alternative, the moving window-based methods, for example, recursive local outlier factor (LOF) [7], measure the similarity between windows with the spatial and temporal information of the features to identify the clusters. After mode identification, each mode is monitored by such traditional methods as principal component analysis/independent component analysis (PCA/ICA) and their extensions and so forth. [8–10]. Recently, several efforts have been made on applying the multiple models with Bayesian fusion to multimode monitoring. Specifically, the Gaussian Mixture Model (GMM) was deployed to accomplish fault diagnosis by considering multiple models simultaneously [11, 12].

Compared with the multimodel scheme, single-model-based method simplifies the modeling and monitoring procedure. Hidden Markov model- (HMM-) based methods were proposed in [13, 14], whereas their parameter setting is still a complex task. Zhu et al. [15] proposed a recursive mixture factor analyzer for multimode time-variant process modeling and monitoring. Considering that there are interconnections among different modes, a strategy was developed to grasp common characteristics in them to construct a monitoring model. Hwang and Han [16]

proposed the super principal component analysis (PCA) method, which can be considered as the first attempt based on this idea. Then, Zhang et al. [17] built the common and individual monitoring model by PCA-based and partial least squares- (PLS-) based scheme. Besides, Zhang et al. [17] utilized the common basis vectors to extract the common features among multimode data. For nearest-neighbor approaches, different modes should be standardized [18]. For example, Ma et al. [19] presented a neighborhood-based global coordination (NBGC) method by aligning the local models into a global one for multimode process monitoring.

Among the single-model-based methods, SVDD model and its extensions are widely applied. Developed from support vector machine [20], SVDD is proposed based on the idea of one-class classification. SVDD projects the data into high-dimensional feature space without the requirement of Gaussian distribution [21]. Therefore, the SVDD method can deal with both nonlinear and non-Gaussian data [22]. However, SVDD is insensible to a very small fraction of outlier, and the trained hypersphere will sometimes tend to encircle outliers that are far away from the normal dataset. Besides, the normal samples with different density have no different effects on SVDD modeling. In the multimode training dataset, the sampling number and data density vary dramatically with different modes. Thus, SVDD is not suitable to be directly applied to multimode monitoring. The wSVDD method was developed to solve this problem, in which the weight factor is usually determined by nearest-neighbor information. To reduce the impact of outliers on modeling in single-mode process, Chen et al. [23] proposed robust-SVDD by introducing cutoff-distance-based local density of each data sample and the ε-insensitive loss function with negative samples. Wang and Lan [24] used the SD outlyingness to assign lower weight values to outliers. For multimode process, Zhao et al. utilized the weighted mean and standard deviation of each sample's neighbors to standardize the dataset and applied weighted local standardization (WLS) strategy to wSVDD [25]. Li et al. [26] proposed a local density ratio weighted support vector data description (LDR-wSVDD) to maintain the monitoring efficiency of a single hypersphere model. Nevertheless, the spatial information is still not fully mined, and the temporal information is not employed in the construction of the weight for wSVDD, which may lead to unsatisfactory performance of the existing methods. Consequently, wSVDD still needs to be further studied.

Recently, Breunig et al. [27] defined local reachability density (LRD) and indicated that the LRD of a sample is related to the local information of not only itself but also its nearest neighbors. By containing more comprehensive density information, LRD has the potential to be transformed into a new weight factor for wSVDD. However, LRD varies drastically with different modes; thus, it does not yet reflect their local information equally. In addition, conventional LRD only considers spatial information and does not take temporal information into account.

Moreover, the data in different modes may overlap each other in actual industrial process, so that the original data sometimes fails to embody the unique characteristics of certain modes. Statistics pattern analysis (SPA) was firstly introduced by Wang and He [28], which is able to mine underlying statistic features of data and has been applied in process monitoring [29]. Zhu and Gu integrated SPA into local kernel principal component analysis (LKPCA) to enhance fault detection performance [30].

In this paper, SPA is applied to obtain statistical characteristics for fully mining the underlying information of multimode dataset firstly. Then, wSVDD is selected as a monitoring model, and modified local reachability density ratio (mLRDR) is proposed as a new weight for wSVDD to remove the multimodality of process and reduce the impact of outliers. The local information in terms of both space and time is considered in mLRDR calculation to further enhance monitoring performance. Ultimately, a novel scheme called SPA-based modified local reachability density ratio weighted support vector data description (SmLRDR-wSVDD) for nonlinear multimode monitoring is established.

The remainder of the paper is organized as follows: in Section 2, the proposed nonlinear multimode process monitoring scheme is elaborated in detail; in Section 3, the results and discussion of TE process and WWTP are presented; finally, conclusions are given in Section 4.

2. Methodology

In this section, SmLRDR-wSVDD method is proposed to monitor nonlinear multimode process, in which statistics pattern dataset is formed by SPA, and a new weight factor named mLRDR is proposed based on LRD for wSVDD monitoring method.

2.1. Construction of Statistics Pattern. Although specific characteristics of different modes may not be embodied from the original data, they can be expressed in terms of the statistics. In the SPA framework [28], different statistics patterns for observed variables are selected to capture the dominant process characteristics such as dynamics and nonlinearity and are used as the modeling and monitoring object in the proposed method.

Suppose that $\mathbf{X} \in \mathbb{R}^{n\times m}$ is the original measured dataset with n samples and m variables, and a window of samples is denoted as

$$\begin{aligned}\mathbf{X}_t^h &= \left[\mathbf{x}_{t,1}^h, \mathbf{x}_{t,2}^h, \ldots, \mathbf{x}_{t,m}^h\right] \\ &= \begin{bmatrix} x_1(t-h+1) & x_2(t-h+1) & \ldots & x_m(t-h+1) \\ x_1(t-h+2) & x_2(t-h+2) & \ldots & x_m(t-h+2) \\ \vdots & \vdots & \vdots & \vdots \\ x_1(t) & x_2(t) & \ldots & x_m(t) \end{bmatrix},\end{aligned} \tag{1}$$

where h is the window width and t is the time index. Generally, statistics pattern (SP) consists of three groups of statistics: first-order, second-order, and high-order statistics, which can be expressed as

$$\mathbf{S} \equiv [\mu|\mathbf{V}|\Xi], \tag{2}$$

where μ denotes the first-order statistics, that is, the mean, and its elements are calculated from the data in a window:

$$\mu_i(t) = \frac{1}{h}\sum_{q=0}^{h-1} x_i(t-q), \tag{3}$$

where $\mathbf{V}$ denotes the second-order statistics, which includes variance (v_i), correlation ($r_{i,j}$), autocorrelation (r_i), and cross-correlation ($r_{i,j}$) of different variables in the window as

$$v_i(t) = \frac{1}{h}\sum_{q=0}^{h-1}\left[x_i(t-q) - \mu_i(t)\right]^2, \tag{4}$$

$$r_{i,j}(t) = \frac{1}{h}\frac{\sum_{q=0}^{h-1}\left[x_i(t-q) - \mu_i(t)\right]\left[x_j(t-q) - \mu_i(t)\right]}{\sqrt{v_i(t)v_j(t)}}, \tag{5}$$

$$r_i^d(t) = \frac{1}{h-d}\frac{\sum_{q=d}^{h-1}\left[x_i(t-q) - \mu_i(t)\right]\left[x_i(t+d-q) - \mu_i(t)\right]}{v_i(t)}, \tag{6}$$

$$r_{i,j}^d(t) = \frac{1}{h-d}\frac{\sum_{q=d}^{h-1}\left[x_i(t-q) - \mu_i(t)\right]\left[x_j(t+d-q) - \mu_i(t)\right]}{\sqrt{v_i(t)v_j(t)}}, \tag{7}$$

with d denoting the time lag between the variables; Ξ denotes the high-order statistics including skewness (γ_i) and kurtosis (κ_i) to measure the degree of nonlinearity and quantify the non-Gaussianity of the process variables [31] as

$$\gamma_i(t) = \frac{(1/h)\sum_{q=0}^{h-1}\left[x_i(t-q) - \mu_i(t)\right]^3}{\left(v_i(t)\right)^{(3/2)}}, \tag{8}$$

$$\kappa_i(t) = \frac{(1/h)\sum_{q=0}^{h-1}\left[x_i(t-q) - \mu_i(t)\right]^4}{\left(v_i(t)\right)^2} - 3. \tag{9}$$

The SP is constructed by stacking λ selected statistics by equations (1)–(9) into a row vector. In the paper, three kinds of statistics $\mu_i(t)$, $v_i(t)$, and $\kappa_i(t)$ are selected from one-order, two-order, and high-order statistics to construct statistics matrix $\mathbf{S} \in \mathbb{R}^{N\times M}$ with $N = n - h + 1$, $M = \lambda \times m$, and $\lambda = 3$.

2.2. mLRDR Weighted SVDD. The main idea of SVDD is to project data into high-dimensional space for constructing a minimum hypersphere. The training dataset $\mathbf{S} = [\mathbf{s}_1^T, \mathbf{s}_2^T, \ldots, \mathbf{s}_N^T]^T \in \mathbb{R}^{N\times M}$ is obtained according to equation (2), and it is normalized as $\overline{\mathbf{S}}$. In the procedure of SVDD modeling, $\overline{\mathbf{S}}$ is firstly mapped from the original space to a higher feature space by a nonlinear transformation function $\Phi(\overline{\mathbf{s}}_i)$. For conventional SVDD, all the training data has the same impact on the model construction, which makes it insensible to outliers and data density. Therefore, wSVDD is adopted by introducing a weight for each training data, and the corresponding hypersphere can be calculated by solving the following primal optimization problem:

$$\begin{aligned} &\min_{R,a,\xi} \mathbf{R}^2 + C\sum_{i=1}^{N} w_i\xi_i \\ &\text{s.t.}\left\|\mathbf{\Phi}(\overline{\mathbf{s}}_i) - \mathbf{a}\right\| \le \mathbf{R}^2 + \xi_i, \quad \xi_i \ge 0, \end{aligned} \tag{10}$$

where $\mathbf{a}$ and $\mathbf{R}$ are the center and the radius of hypersphere, respectively; trade-off parameter $C = (1/(p \times N))$ is introduced to make the hypersphere as small as possible while preventing the misclassification of the samples, in which p denotes the percentage of outliers permitted in the training set [26]; ξ_i is slack variable that allows outliers of the hypersphere; and w_i is weight of the i^{th} data, where smaller w_i indicates more possibility of an outlier.

For multimode processes, the weight factor of wSVDD should fully reflect the local distribution characteristics of each piece of training data. Density is an indicator reflecting the data distribution, which is widely used in the design of weight factor. Local reachability density (LRD) [27] of a data point depends on its distance to not only the point's neighbors but also its neighbors' neighbors.

Besides the space density, the current data is also related to the previous and the next points in time sequence, and monitoring performance will be improved by taking both the temporal information and spatial information into account [32]. Consequently, the modified LRD is proposed in this paper by considering the local neighbors in both time and space to comprehensively extract the local features of the data.

Firstly, all K-nearest neighbors of $\overline{s}_i\,(i = 1, 2, \ldots, N)$ are selected to construct a set $\text{KNN}(\overline{s}_i)$, which satisfies $K = K_1 + K_2$, where K_1 and K_2 are the numbers of neighbors in space and time, respectively.

Suppose that $K - \text{distance}(\overline{s}_i)$ is the Euclidean distance from $\overline{s}$ to its K^{th}-nearest neighbor. Then the distance $d(\overline{s}_i, \overline{s}_i^k)$ between $\overline{s}_i$ and $\overline{s}_i^k$ ($k = 1, 2, \ldots, \text{K}$) is computed, where $\overline{s}_i^k$ is the k^{th}-nearest neighbor of $\overline{s}_i$. The reachability distance of $\overline{s}_i$ with respect to $\overline{s}_i^k$ is defined as

$$\text{reach} - d\left(\overline{s}_i, \overline{s}_i^k\right) = \max\left\{K - \text{distance}\left(\overline{s}_i^k\right), d\left(\overline{s}_i, \overline{s}_i^k\right)\right\}. \tag{11}$$

Thus, mLRD is constructed as

$$\text{mLRD}\left(\overline{s}_i\right) = \frac{K}{\sum_{k=1}^{K}\text{reach} - d\left(\overline{s}_i, \overline{s}_i^k\right)}. \tag{12}$$

The main property of weight factor for multimode process should equalize different mode and reduce the influence of outlier in modeling. According to equation (12), if distribution characteristics vary greatly with different modes, mLRD in different modes will change a lot; and some outliers will be hardly separated from normal data; that is, the local density of normal data with lower global density may be similar to that of the outliers around the data with

higher global density. Accordingly, to weaken the multimodality feature and widen the gap between outlier and the normal data, a modified local reachability density ratio (mLRDR) is developed to be the weight factor of wSVDD as

$$w(\bar{s}_i) = \mathrm{mLRDR}(\bar{s}_i) = \frac{\mathrm{mLRD}(\bar{s}_i)}{\max_{\bar{s}_j \in \mathrm{KNN}(\bar{s}_i)}\{\mathrm{mLRD}(\bar{s}_i), \mathrm{mLRD}(\bar{s}_j)\}}. \tag{13}$$

It is obvious that $w(\bar{s}_i)$ is kept within $[0, 1]$.

According to equation (13), if $\bar{s}_i$ is a normal point, its mLRD is close to its neighbors and the corresponding mLRDR approaches 1. If $\bar{s}_i$ is an outlier, its mLRD is significantly smaller than that of its normal neighbors and mLRDR is close to 0, which indicates that outliers will have trivial influence on the determination of the hypersphere.

Compared with the existing weight factors such as LDR [26], mLRDR contains more comprehensive spatial and additional temporal information, and it can distinguish normal data with lower density and the outliers around the data points with higher density data; thus it will improve the monitoring performance of wSVDD.

For wSVDD modeling, equation (10) can be transformed into a dual-optimization problem by introducing Lagrange multipliers $\alpha = [\alpha_1, \alpha_1, \ldots, \alpha_N]^T$ as follows:

$$\begin{aligned} &\min_{\alpha_i} \sum_{i=1}^{N} \alpha_i \mathrm{Ker}(\bar{s}_i, \bar{s}_i) - \sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i \alpha_j \mathrm{Ker}(\bar{s}_i, \bar{s}_j) \\ &\text{s.t.} \sum_{i=1}^{N} \alpha_i = 1, \quad 0 \le \alpha_i \le w_i C, \end{aligned} \tag{14}$$

where $\mathrm{Ker}(\bar{s}_i, \bar{s}_j) = \langle \Phi(\bar{s}_i) \cdot \Phi(\bar{s}_j) \rangle = e^{[-(\bar{s}_i - \bar{s}_j)^T (\bar{s}_i - \bar{s}_j)/\sigma^2]}$, is kernel function and σ can be determined via detecting the "tightness" of the decision boundaries [33]. After solving equation (14), only objects $\bar{s}_i$ with $\alpha_i > 0$ are called the support vectors (SVs), and their subscripts set is $\mathrm{SV} = \{i | \alpha_i > 0,\ i = 1, 2, \ldots, N\}$. Hence, the center and radius of the hypersphere are

$$\mathbf{a} = \sum_{i \in SV} \alpha_i \Phi(\bar{s}_i), \tag{15}$$

$$\mathbf{R} = \sqrt{\mathrm{Ker}(\bar{s}_{sv}, \bar{s}_{sv}) - 2\sum_{i=1}^{N} \alpha_i \mathrm{Ker}(\bar{s}_{sv}, \bar{s}_i) + \sum_{i=1}^{N}\sum_{j=1}^{N} \mathrm{Ker}(\bar{s}_i, \bar{s}_j)}, \tag{16}$$

respectively.

When h new samples are obtained for a window, SPs are calculated by equation (2) and normalized as $\bar{s}_{\mathrm{new}}$, and the distance between $\bar{s}_{\mathrm{new}}$ and the center $\mathbf{R}$ can be calculated by

$$\mathrm{Dist}(\bar{s}_{\mathrm{new}}) = \sqrt{\mathrm{Ker}(\bar{s}_{\mathrm{new}}, \bar{s}_{\mathrm{new}}) - 2\sum_{i=1}^{N} \alpha_i \mathrm{Ker}(\bar{s}_{\mathrm{new}}, \bar{s}_i) + \sum_{i=1}^{N}\sum_{j=1}^{N} \mathrm{Ker}(\bar{s}_i, \bar{s}_j)}. \tag{17}$$

If $\mathrm{Dist}(\bar{s}_{\mathrm{new}}) \le \mathbf{R}$, the system is considered in normal state. Otherwise, a fault occurs.

2.3. The Procedure of SmLRDR-wSVDD Scheme. In this subsection, the proposed SmLRDR-wSVDD scheme is illustrated in Figure 1. The specific offline modeling and online monitoring steps are listed as follows:

Offline modeling procedure:

(i) Step 1: adopt SPA on a multimode dataset $\mathbf{X} \in \mathbb{R}^{N \times M}$ to obtain statistics matrix $\mathbf{S} \equiv [\mu | \mathbf{V} | \Xi] \in \mathbb{R}^{N \times M}$ by equations (2)–(9).
(ii) Step 2: normalize statistics matrix $\mathbf{S}$ and obtain $\overline{\mathbf{S}}$.
(iii) Step 3: compute the mLRDR as a weight factor $\mathbf{W}(\bar{s}_i)$ by equations (11)–(13).
(iv) Step 4: set up weighted-SVDD model with the weight $\mathbf{W}$, solve equation (14), and obtain the center and the radius of the hypersphere by equations (15) and (16), respectively.

Online monitoring procedure:

(i) Step 1: for a current window dataset $\mathbf{x}_{\mathrm{new}}$, statistics vector $\mathbf{s}_{\mathrm{new}} \in \mathbb{R}^M$ is calculated by equations (2)–(9).
(ii) Step 2: normalize statistics vector $\mathbf{s}_{\mathrm{new}}$ to obtain $\bar{s}_{\mathrm{new}}$.
(iii) Step 3: calculate the distance Dist between $\bar{\mathbf{s}}_{\mathrm{new}}$ and the center of the hypersphere by (17fd17).
(iv) Step 4: if $\mathrm{Dist}(\bar{\mathbf{s}}_{\mathrm{new}}) \le \mathbf{R}$, $\mathbf{x}_{\mathrm{new}}$ is considered to be normal; otherwise, it is in a faulty state.

In this paper, if Dist**s** of 3successive windows are over the control limit **R**, the system is considered in faulty condition, and the last sample in the first faulty window is determined as the beginning of the fault.

It must be noted that the identification of new normal modes is not considered in the proposed method. Although a new normal mode will be detected as a fault by SmLRDR-wSVDD, it may be redefined by analyzing the stationarity of sampling data after enough data are collected.

In summary, the proposed method adopts SPA to mine more features from different-order statistics of the original data compared with other existing SVDD-based methods. Besides, a new weight factor is proposed for wSVDD modeling, in which the additional neighbors' local spatial information and temporal information are considered, which makes the proposed method more sensitive to outliers and density compared with other methods.

3. Case Study

In this section, the proposed SmLRDR-wSVDD method is applied on Tennessee-Eastman (TE) benchmark process and wastewater treatment process (WWTP). In both cases, the proposed method is compared with the existing methods including KPCA [34], conventional SVDD, and LDR-wSVDD [26].

3.1. Tennessee-Eastman Benchmark Process. Tennessee-Eastman process (TE process) was put forward in 1993 by Downs and Vogel [35], which has been widely adopted for scientific research [36]. There are five units,

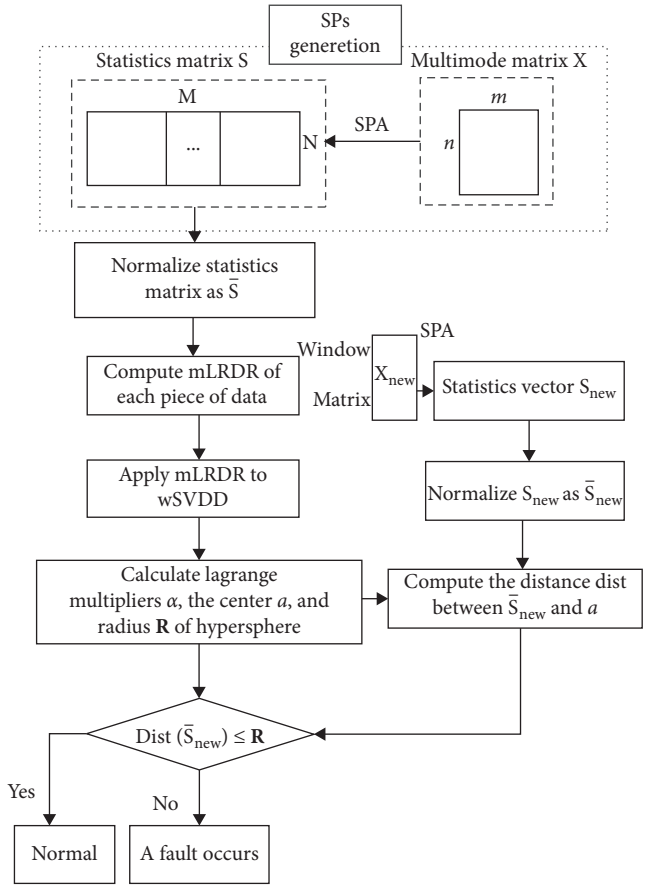


Figure 1: Flowchart of SmLRDR-wSVDD scheme for multimode monitoring.

reactor, condenser, compressor, stripper, and separator, in this nonlinear process. According to different (G/H) mass ratios (G and H are two products), 6 steady modes can be obtained as listed in Table 1.

The control scheme and normal operations were provided by Ricker [37] and the corresponding simulation platform can be downloaded from http://depts.washington.edu/control/LARRY/TE/download.html. Mode 1 and Mode 3 are simulated to generate modeling and testing datasets. There are 53 variables: 12 manipulated variables, 22 continuous process variables, and 19 composition measurements. Since the steady-state values of the recycle value and steam value in Mode 1 and the agitator rates in both modes do not change, they are not included in the monitored variables, and the remaining 9 manipulated variables and 22 easily measured continuous variables are chosen in this case. In addition, TE process simulation platform contains 20 faults. Excluding faults 3, 9, and 15, which are slight disturbances, and unknown faults 16–20, 12 faults given in Table 2 are employed; and the sampling interval is set as 3 minutes.

For model training, 800 normal samples from Mode 1 and 200 normal samples from Mode 3 are collected, and 20 outliers are mixed with them. For online monitoring, each test dataset contains 1000 samples, in which the first 200 samples are normal, and the fault occurs at the 201^{th} sample. The confidence level of KPCA is chosen as 95%. For the other three algorithms, the Gaussian kernel width parameter σ is determined by the method in [33], and $p = 5\%$ is selected for the penalty parameter C. In the proposed method, the window width h is assigned as 10, and the numbers of neighbors in space and time K_1 and K_2 are chosen as 8 and 2, respectively. The above-mentioned parameters are the same in the two cases.

The four methods, that is, KPCA, SVDD, LDR-wSVDD, and the proposed SmLRDR-wSVDD, are applied and compared on the TE process. The missed alarm rates (MARs) [38] for each fault type of Mode 1 and Mode 3 are shown in Table 3, in which the best results are marked in bold, and the false alarm rates (FAR) are shown in Table 4. Specifically, MARs of KPCA on faults 4, 5, 6, 7, 10, 11, 12, and 14 of Mode 1 and faults 4, 5, 6, 7, 10, 11, and 14 of Mode 3 are all more than 0.5, indicating that a single KPCA model is not suitable to monitor multimode process. Among the four methods, the proposed SmLRDR-wSVDD method achieves the lowest MARs for almost all the faults except for fault 5 of Mode 1 and fault 2 of Mode 3. For example, the proposed method greatly decreases the MARs of fault 10 for both Mode 1 and Mode 3 and fault 12 for Mode 1 by over 9% compared to LDR-wSVDD. It must be noted that all methods have high MAR for fault 5 of Mode 1, since it is actually a slight disturbance.

Further, the detailed monitoring results of four methods for different faults of two modes are shown in Figures 2–5. For fault 11 of Mode 1, the SPE of KPCA and Dist of SVDD fluctuate violently and some of them are under the limits shown in Figures 2(a) and 2(b); as illustrated in Figure 2(c), Dist fluctuates slightly, and some fall below the limit by LDR-wSVDD; meanwhile, with the proposed method shown in Figure 2(d), all the statistics are far above the control limit. Similarly, the proposed SmLRDR-wSVDD can separate normal condition and fault more obviously than the other 3 methods as shown in Figures 3 and 4. The monitoring result of fault 13 for Mode 3 is shown in Figure 5. For this fault, there are 48, 47, 27, and 9 samples of detection delay by KPCA, SVDD, LDR-wSVDD, and SmLRDR-wSVDD. Thus, one can infer that SmLRDR-wSVDD is more sensitive to the four faults than the other methods.

FARs of all methods are summarized in Table 4. For all faults of Mode 1, FARs of KPCA and LDR-wSVDD are within $[0.01, 0.03]$ and $[0.01, 0.02]$, respectively, while those of conventional SVDD and SmLRDR-wSVDD are within $[0, 0.01]$. In Mode 3, FARs of SPE in KPCA are the largest for all faults, sometimes even larger than 0.05, and those of LDR-wSVDD for all faults range from 0.015 to 0.035. Note that all FARs of the proposed SmLRDR-wSVDD method are less than 0.02, though most of them are higher than those of conventional SVDD within 0.01.

Figure 6 shows LRD and the weight of each training sample based on LDR and the proposed mLRDR, respectively, in which blue, green, and red dots are from Mode 1, Mode 3, and outliers, respectively. In Figure 6(a), it is easy to find that the normal modes vary greatly in terms of LRD, and there is no distinctive gap between Mode 3 and outliers. As shown in Figures 6(b) and 6(c), LDRs and mLRDRs of samples in two normal modes distribute around 0.6 and 0.9 evenly. Compared with LDRs, mLRDRs of outliers are quite

Table 1: The parameters of 6 modes.

Mode	G (%)	H (%)	Product yield
1	50	50	G,H: 7038 kgh^{-1}
2	10	90	G:140 $8kgh^{-1}$, H:1266 $9kgh^{-1}$
3	90	10	G:1000 $0kgh^{-1}$, H:111 $1kgh^{-1}$
4	50	50	Maximum productivity
5	10	90	Maximum productivity
6	90	10	Maximum productivity

Table 2: The parameters of 6 modes.

No.	Description	Type
1	A/C feed ratio, B composition constant (stream 4)	Step
2	Composition, A/C ratio constant (stream 4)	Step
4	Reactor cooling water inlet temperature	Step
5	Condenser cooling water inlet temperature	Step
6	A feed loss (stream 1)	Step
7	C header pressure loss-reduced availability (stream 4)	Step
8	A, B, C feed composition (stream 4)	Step
10	C feed temperature (stream 4)	Random variation
11	Reactor cooling water inlet temperature	Random variation
12	Condenser cooling water inlet temperature	Random variation
13	Reaction kinetics	Slow drift
14	Reactor cooling water valve	Sticking

Table 3: MAR for all test conditions on TE process.

Fault	Mode 1					Mode 3				
	KPCA		SVDD	LDR-SVDD	SmLRDR-SVDD	KPCA		SVDD	LDR-SVDD	SmLRDR-SVDD
	T^2	SPE				T^2	SPE			
1	0.9513	0.005	0.0025	0.0025	**0**	0.9925	0.0138	0.0088	0.0075	**0.005**
2	0.8725	0.0075	0.005	0.0025	**0.0013**	0.9083	0.27	0.0388	**0.0075**	0.0363
4	0.9863	0.9713	0	0	**0**	0.9838	0.875	0	0	**0**
5	0.9863	0.9775	0.965	**0.9063**	0.9313	0.9775	0.9975	0.0838	0.0325	**0.0025**
6	0.9854	0.5773	0	0	**0**	0.9877	0.5802	0	0	**0**
7	0.9838	0.9338	0	0	**0**	0.9763	0.8825	0	0	**0**
8	0.63	0.0188	0.0175	0.0175	**0.0112**	0.565	0.0163	0.0163	0.015	**0.01**
10	0.9663	0.96	0.8338	0.5333	**0.4413**	0.9588	0.7863	0.565	0.2875	**0.105**
11	0.94	0.5372	0.1625	0.015	**0**	0.9475	0.7213	0.235	0.015	**0**
12	0.9588	0.8113	0.6363	0.4888	**0.2688**	0.6075	0.0213	0.0175	0.0075	**0.0025**
13	0.4663	0.0238	0.0225	0.0225	**0.0213**	0.4675	0.07	0.0587	0.0262	**0.0112**
14	0.9875	0.9538	0.7075	0.1062	**0**	0.985	0.8725	0.6038	0.0075	**0**

Table 4: FAR for all test conditions on TE process.

Fault	Mode 1					Mode 3				
	KPCA		SVDD	LDR-SVDD	SmLRDR-SVDD	KPCA		SVDD	LDR-SVDD	SmLRDR-SVDD
	T^2	SPE				T^2	SPE			
1	0.03	0.02	0.01	0.02	0.01	0.01	0.025	0.01	0.025	0.01
2	0.03	0.01	0.005	0.015	0.005	0.015	0.025	0.005	0.02	0.015
4	0.02	0.015	0.005	0.015	0.01	0.01	0.025	0.005	0.02	0.01
5	0.02	0.015	0.01	0.01	0.01	0.02	0.05	0.05	0.015	0.015
6	0.01	0.015	0.005	0.01	0.01	0.015	0.05	0.005	0.02	0.02
7	0.02	0.02	0.005	0.01	0.01	0.02	0.055	0.005	0.02	0.02
8	0.025	0.015	0.01	0.02	0.005	0.015	0.05	0.01	0.025	0.015
10	0.025	0.02	0.005	0.02	0.005	0.015	0.04	0.005	0.025	0.01
11	0.02	0.01	0.005	0.015	0.005	0.01	0.04	0.005	0.03	0.015
12	0.02	0.015	0.005	0.015	0.005	0.02	0.04	0.01	0.025	0.01
13	0.02	0.015	0.005	0.02	0.01	0.01	0.065	0.005	0.025	0.01
14	0.02	0.01	0.01	0.02	0.005	0.01	0.07	0.01	0.03	0.01

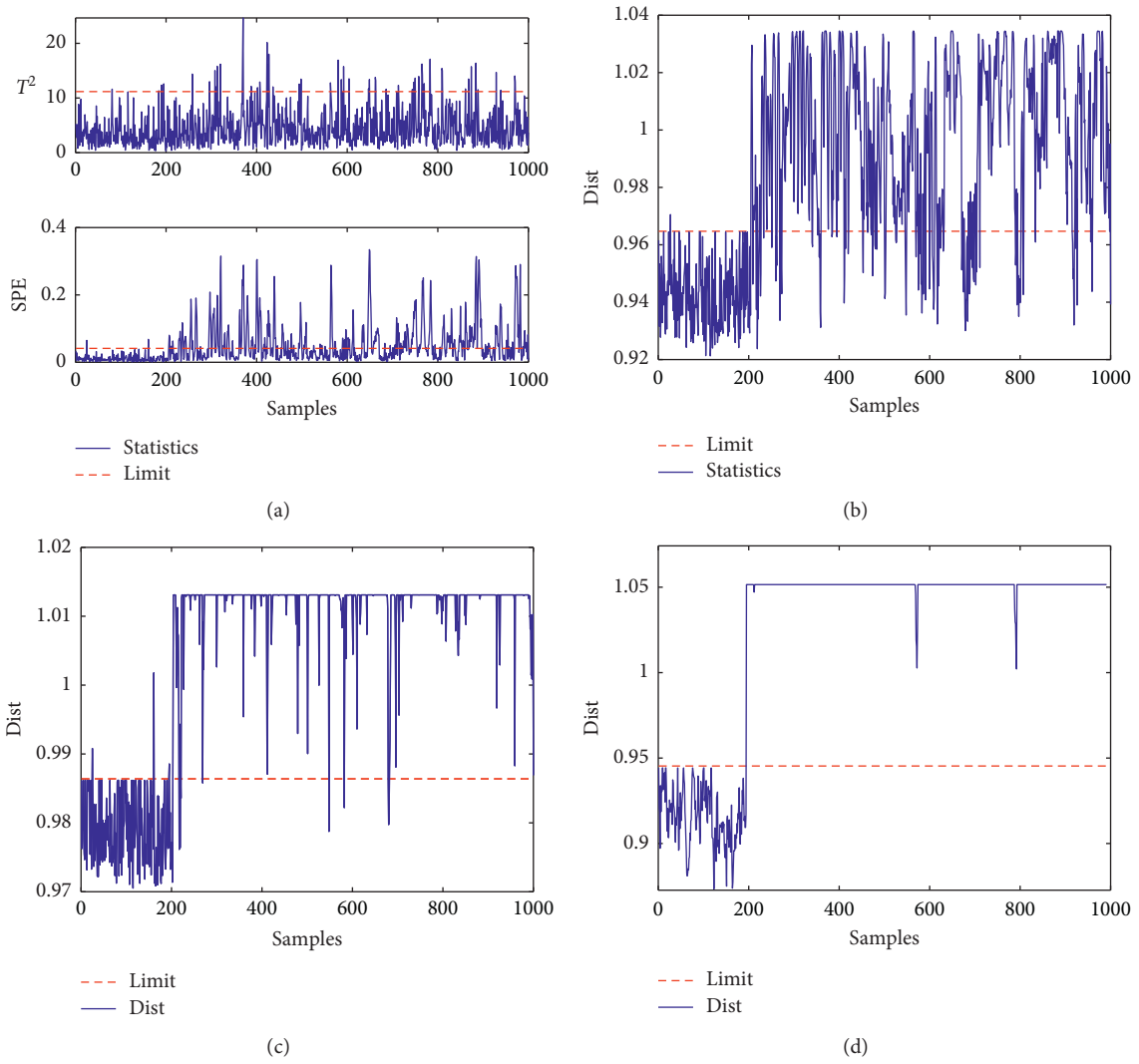


Figure 2: Monitoring results for fault 11 of Mode 1 by (a) KPCA, (b) conventional SVDD, (c) LDR-wSVDD, and (d) SmLRDR-wSVDD.

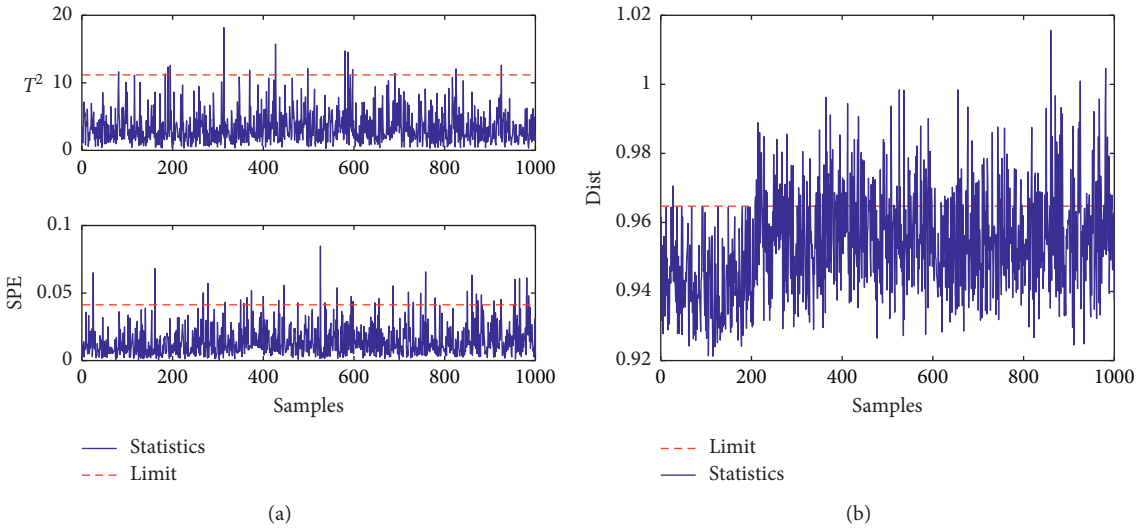


Figure 3: Continued.

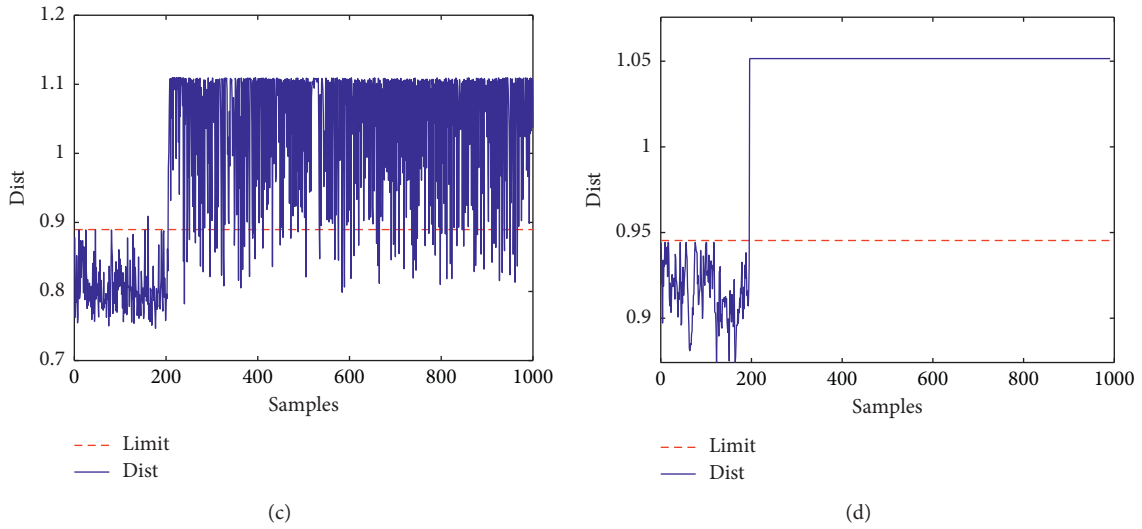


Figure 3: Monitoring results for fault 14 of Mode 1 by (a) KPCA, (b) conventional SVDD, (c) LDR-wSVDD, and (d) SmLRDR-wSVDD.

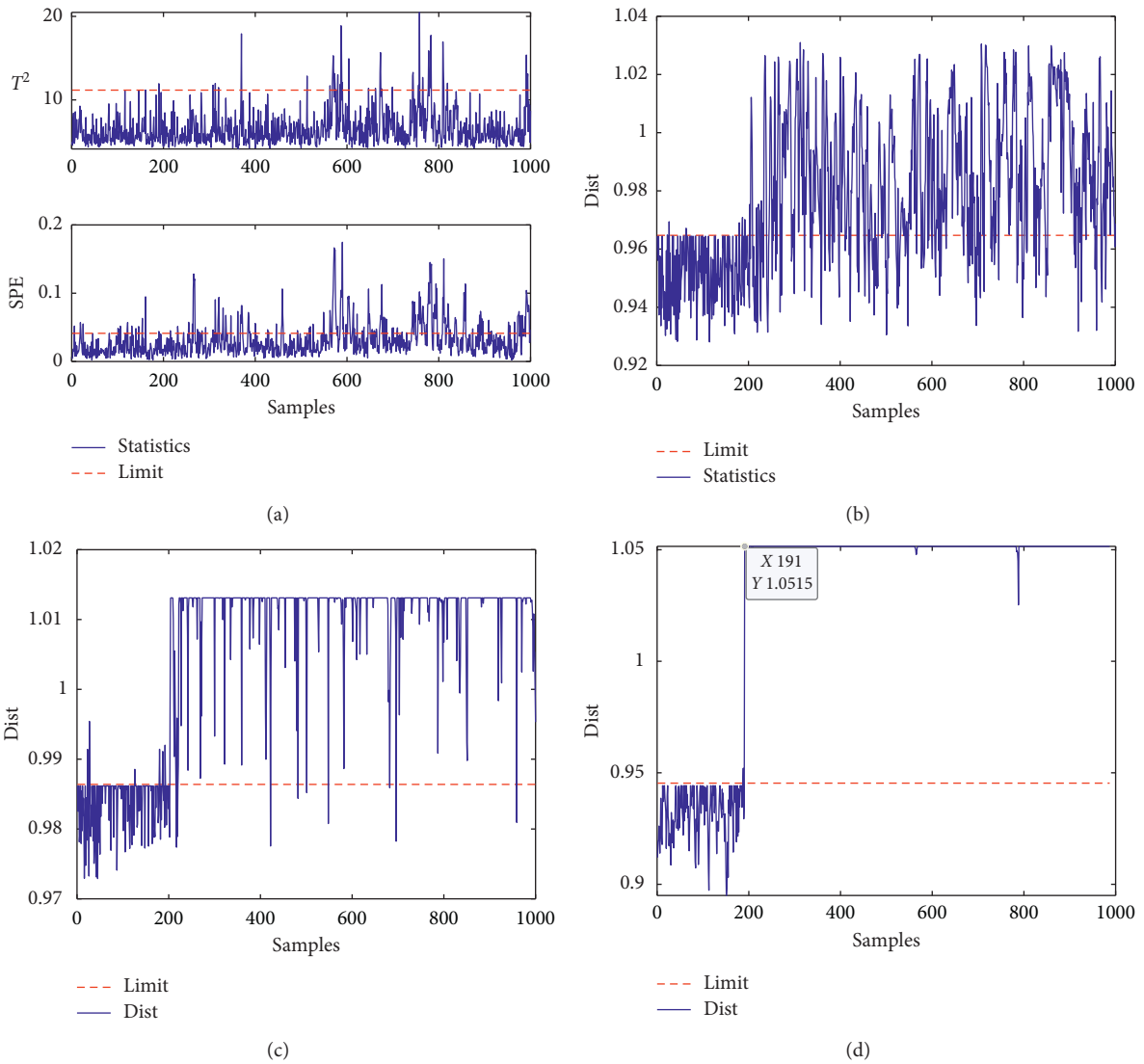


Figure 4: Monitoring results for fault 11 of Mode 3 by (a) KPCA, (b) conventional SVDD, (c) LDR-wSVDD, and (d) SmLRDR-wSVDD.

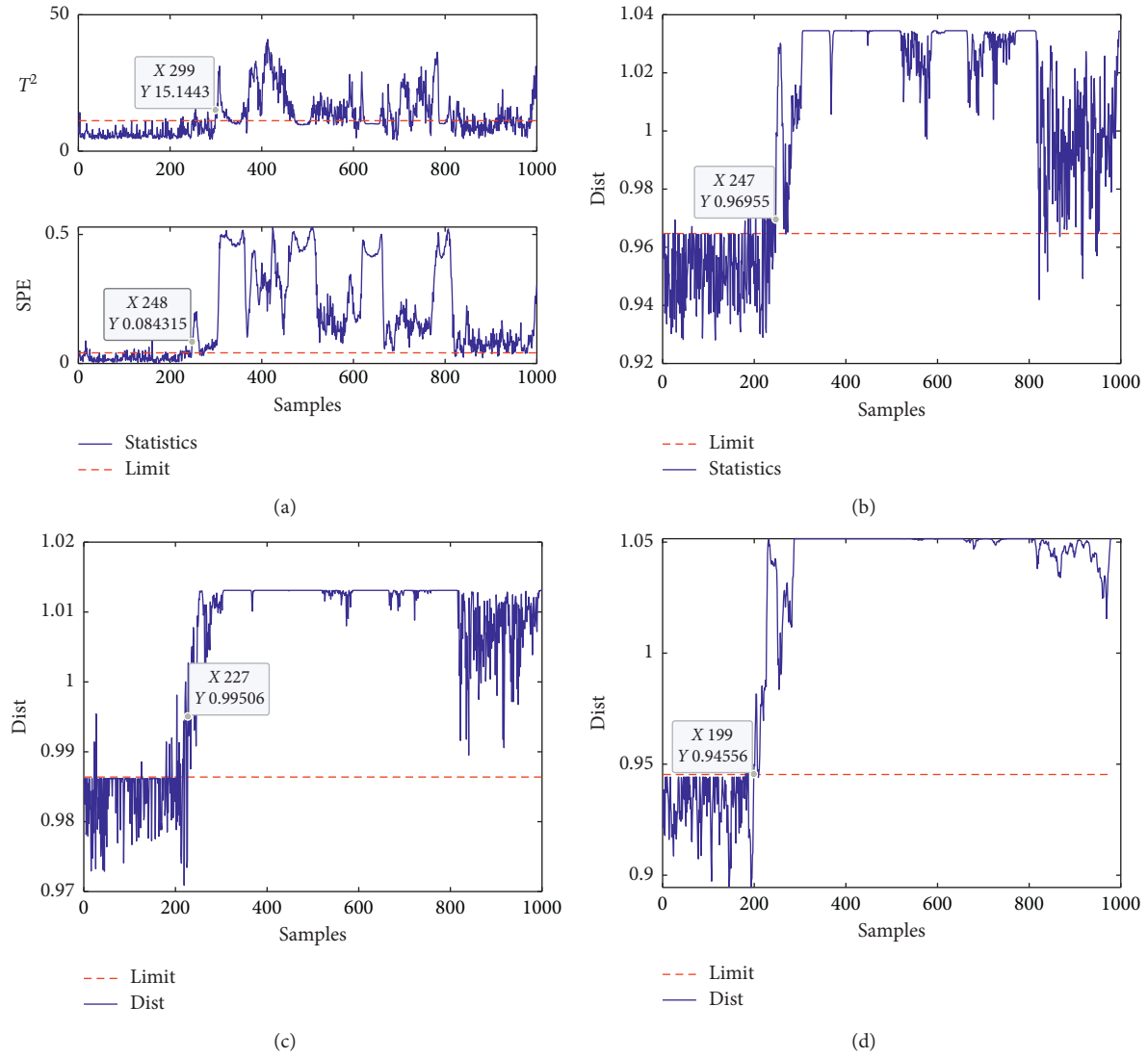


Figure 5: Monitoring results for fault 13 of Mode 3 by (a) KPCA, (b) conventional SVDD, (c) LDR-wSVDD, and (d) SmLRDR-wSVDD.

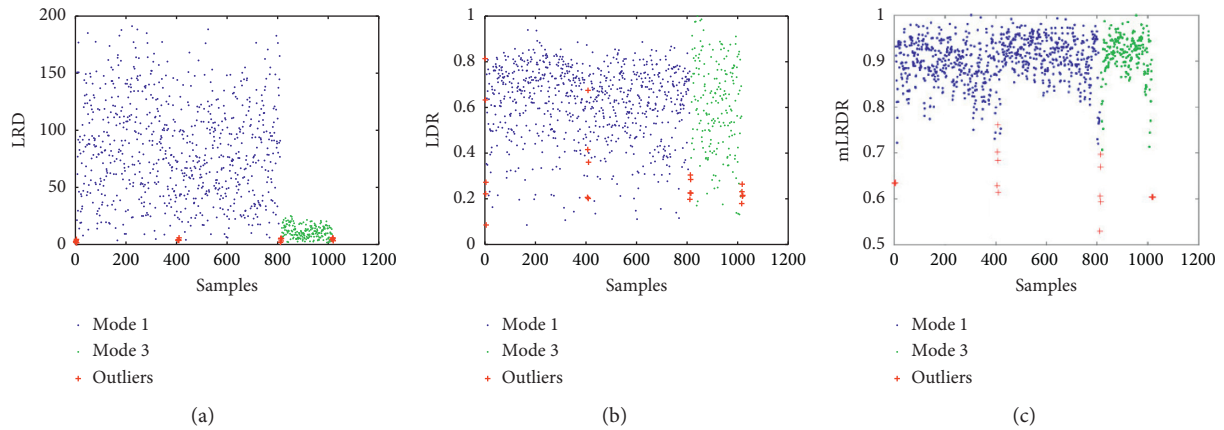


Figure 6: (a) LRD; the weight of each training data based on (b) LDR and (c) mLRDR.

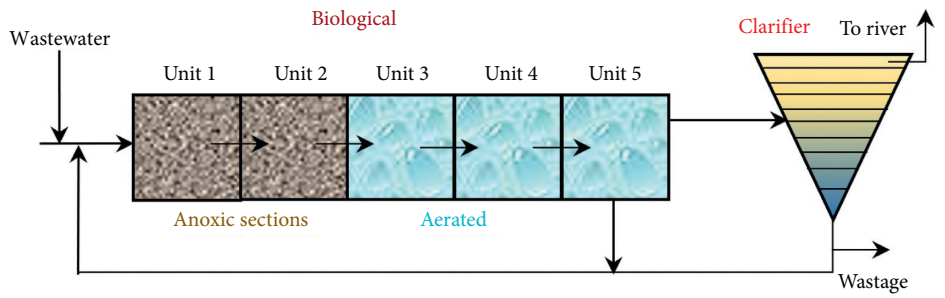


Figure 7: Process layout of the WWTP.

Table 5: FAR for all test conditions on TE process.

Parameter	Mode 1	Mode 2
Influent flow rate (m^3/day)	18446	9223
Recycle flow rate (m^3/day)	18446	9223
Internal recycle flow rate (m^3/day)	55338	27669
KLa-Tank 3 (h^{-1})	10	7.5
KLa-Tank 4 (h^{-1})	10	7.5
KLa-Tank 5 (h^{-1})	3.5	3.5

Table 6: Description and notation of the monitoring variables.

Description	Notation		
Readily biodegradable substrate	$S_{S,3}$ (Tank 3)	$S_{S,4}$ (Tank 4)	$S_{S,5}$ (Tank 5)
Oxygen	$S_{O,3}$ (Tank 3)	$S_{O,4}$ (Tank 4)	$S_{O,5}$ (Tank 5)
Nitrate and nitrite nitrogen	$S_{NO,3}$ (Tank 3)	$S_{NO,4}$ (Tank 4)	$S_{NO,5}$ (Tank 5)
$NH_4^+ + NH_3$ nitrogen	$S_{NH,3}$ (Tank 3)	$S_{NH,4}$ (Tank 4)	$S_{NH,5}$ (Tank 5)
Alkalinity	$S_{ALK,3}$ (Tank 3)	$S_{ALK,4}$ (Tank 4)	$S_{ALK,5}$ (Tank 5)
Total suspended solid	TSS_4 (Tank 4)		
Oxygen transfer coefficient	KLa_5 (Tank 4)		
Nitrate and nitrite nitrogen	Settler underflow $S_{NO,u}$		Plant exit $S_{NO,e}$
The advanced biological nitrogen removal process	Influent ammonium concentration $S_{NH,in}$		

Table 7: Six simulation cases in the BSM1 benchmark.

Disturbance type	Samples		Simulation conditions
None	Dry weather	$1^{st} - 672^{nd}$	Normal Mode 1
		$673^{rd} - 1344^{th}$	Normal Mode 2
External	Case 1	$1^{st} - 800^{th}$, $1001^{st} - 1344^{th}$	Normal Mode 1
		$801^{st} - 1000^{th}$	Rain
Internal	Case 2	$1^{st} - 672^{nd}$	Normal Mode 1
		$673^{rd} - 1344^{th}$	Specific growth rate for autotrophs: μ_A from 0.5 to 0.3 each day in a linear fashion
	Case 3	$1^{st} - 672^{nd}$	Normal Mode 1
		$673^{rd} - 1344^{th}$	Maximum heterotrophic growth rate μ_H has a step change decreasing from 4.0 d^{-1} to 2.0 d^{-1}
	Case 4	$1^{st} - 672^{nd}$	Normal Mode 1
		$673^{rd} - 1344^{th}$	DO controller setpoint: changed from 2 to 1 mg/L (operation)
	Case 5	$1^{st} - 672^{nd}$	Normal Mode 2
		$673^{rd} - 1344^{th}$	μ_H has a step change decreasing from 4.0 d^{-1} to 3.0 d^{-1}, and maximum heterotrophic decay rate (b_H) has a step change increasing from 0.3 d^{-1} to 0.1 d^{-1}

Table 8: MAR under five cases on WWTP.

Cases no.	KPCA		SVDD	LDR-wSVDD	SmLRDR-wSVDD
	T^2	SPE			
Case 1	1	0.03	0.04	0.035	0
Case 2	1	0.2113	0.1741	0.1265	0.0164
Case 3	0.9732	0.4137	0.0729	0.003	0
Case 4	0.7589	0	0	0	0
Case 5	0.9643	0.2411	0.1518	0.0313	0

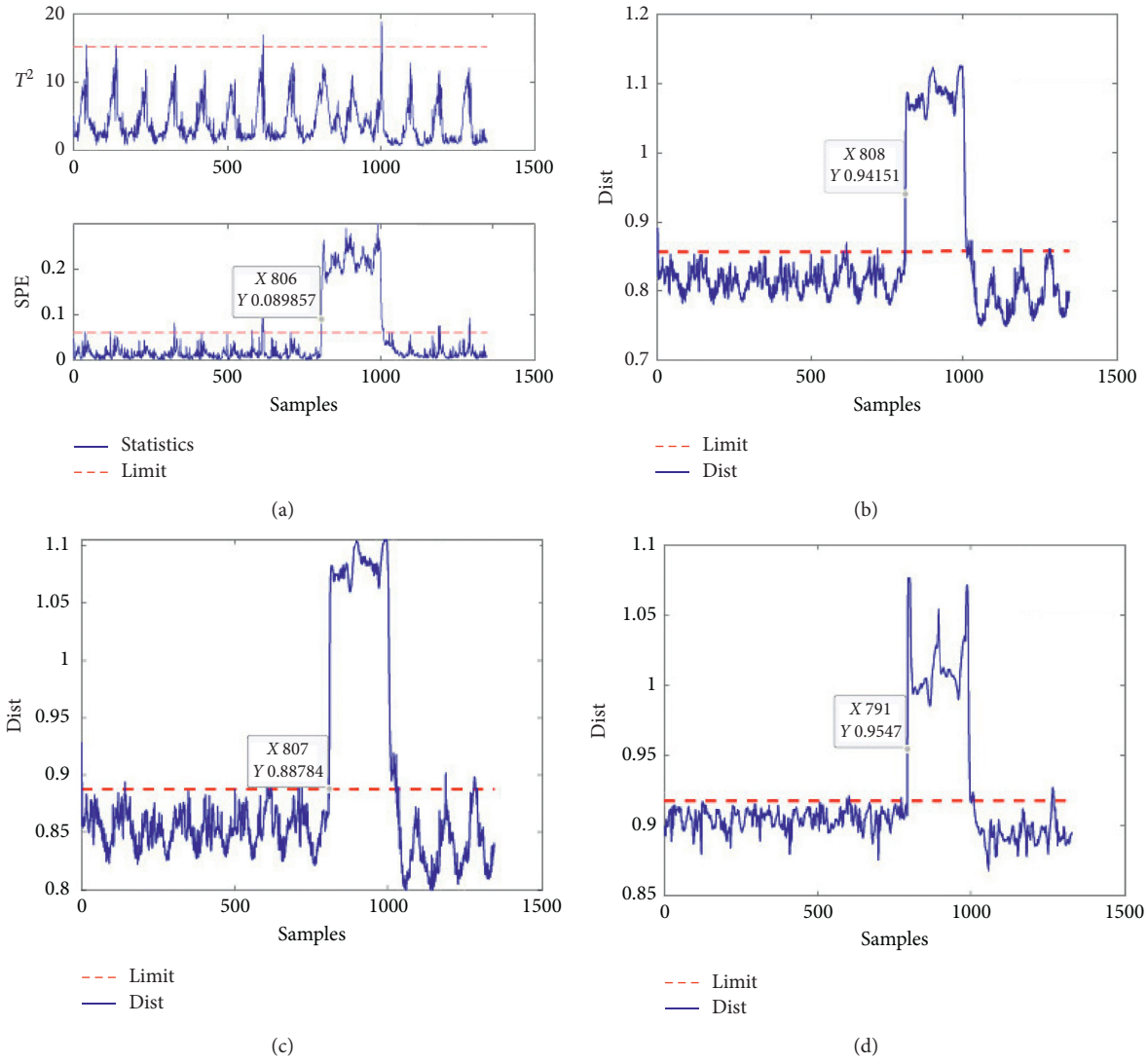


Figure 8: Monitoring results for Case 1 by (a) KPCA, (b) conventional SVDD, (c) LDR-wSVDD, and (d) SmLRDR-wSVDD.

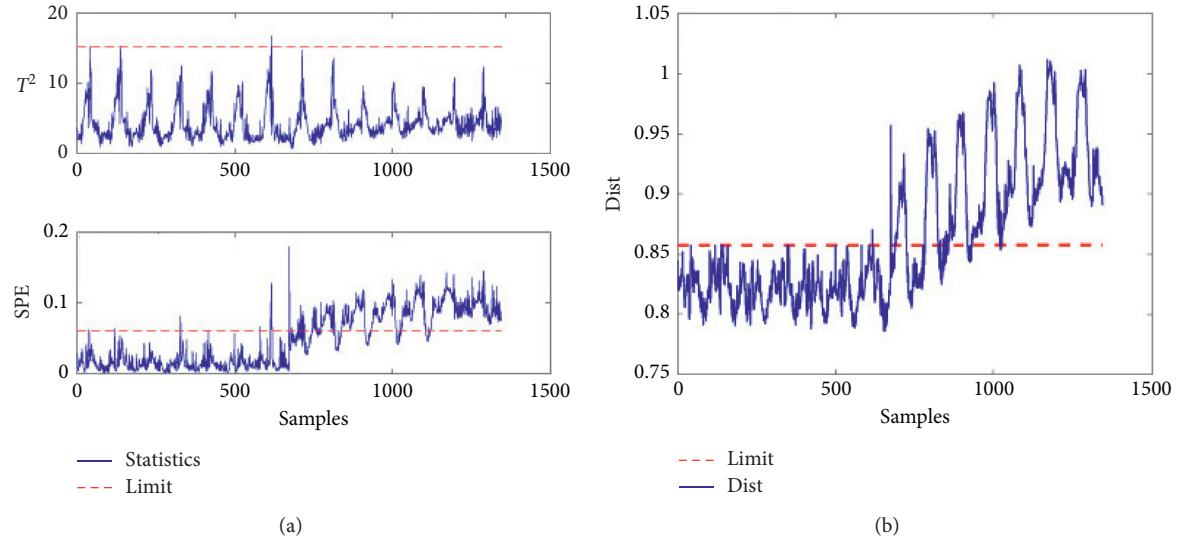


Figure 9: Continued.

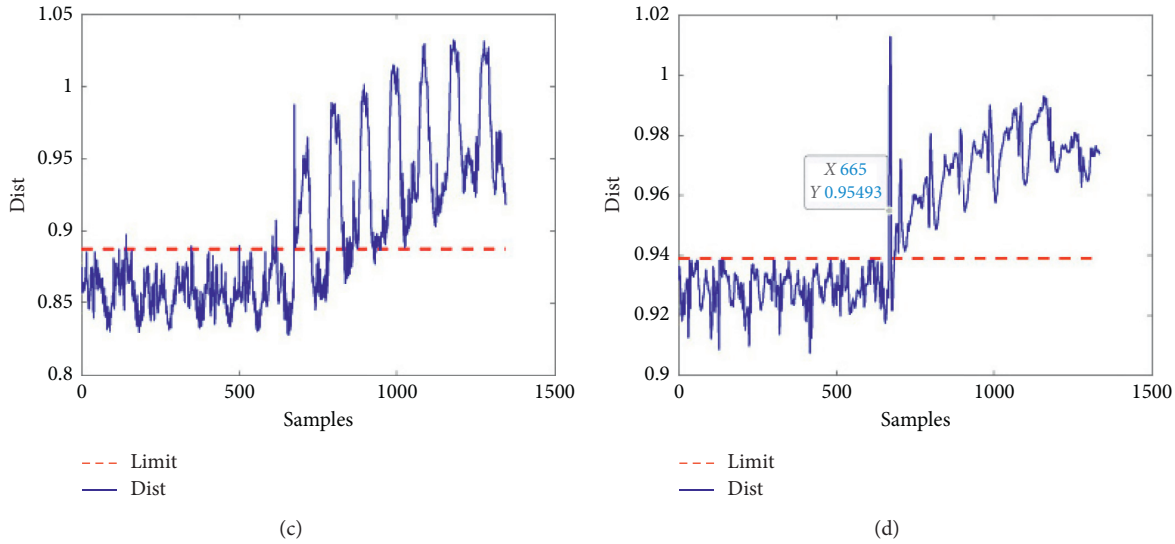


Figure 9: Monitoring results for Case 2 by (a) KPCA, (b) conventional SVDD, (c) LDR-wSVDD, and (d) SmLRDR-wSVDD.

lower than those of normal data. Therefore, the decision boundary is not susceptible to outliers by the proposed method, which will improve its monitoring performance compared with other methods.

3.2. Wastewater Treatment Process (WWTP). In this subsection, benchmark simulation model no. 1 (BSM1) of WWTP [39] is considered. This benchmark is a complicated nonlinear system, which consists of five compartments (two anoxic and three aerobic) in the biological reactor and a ten-layer clarifier. The layout of the plant is shown in Figure 7, and more detailed information about BSM1 can be found at http://www.benchmarkwwtp.org/.

By changing the flow rates and oxygen transfer coefficients in biological reactor, three different modes can be obtained [40]. Without loss of generality, Mode 1 and Mode 2 are chosen in this case, and the corresponding parameters are shown in Table 5. The process contains 223 variables, 20 of them is related to the biological phenomena, and they are selected in this example as shown in Table 6.

There are two disturbance types in the benchmark simulation [41]: external disturbances and internal disturbances. Specifically, dry weather, storm events, and prolonged rain are three weather situations, among which the first one is regarded as normal and the latter two are external disturbances, which can be detected by the influent characteristics. As for the internal disturbances, decreasing nitrification, decreasing settling velocity, nitrate sensor failure, and setpoint change of DO controller were mentioned in [41].

In this example, the samples are collected every 15 minutes. 1344 samples of dry weather of Mode 1 and Mode 2 shown in Table 7 are mixed with 20 additional outliers as training dataset. 5 simulation cases (Case 1–Case 5) are designed for the purpose of testing [40, 41], which are also illustrated in Table 7.

The four above-mentioned algorithms are performed for the 5 cases, and the parameters of them are the same as those in the TE process study. The MARs for all faults are shown in Table 8. For each case, the best result is marked in bold. All the four methods have low MARs in Cases 1 and 4. Meanwhile the MARs of KPCA are beyond 0.2 in Cases 2, 3, and 5, and those of conventional SVDD are more than 0.1 in Cases 2 and 5. Most MARs of LDR-wSVDD are below 0.05, except 0.1265 in Case 2. It must be noted that the MARs of the proposed method in Cases 1, 3, 4, and 5 are 0 and only 0.0164 in Case 2, all of which are the lowest.

Specifically, the monitoring results for Case 1 are shown in Figure 8. As shown in Figures 8(a)–8(c), the fault is detected with 6, 8, and 7 points' delay by SPE of KPCA, the conventional SVDD, and LDR-wSVDD, respectively; meanwhile the proposed method can detect the fault immediately after its occurrence, as shown in Figure 8(d). For Case 2 shown in Figure 9, the fault occurs after the 672nd sample, and it changes slowly away from normal. From Figures 9(a) and 9(b), there are a lot of statistics falling below the control limits after the 672nd sample under KPCA and the conventional SVDD. As shown in Figure 9(c), though the weight factor, that is, LDR, is applied in wSVDD, many statistics are still under the limit at the beginning of the fault. As shown in Figure 9(d), the fault is detected quickly by the proposed method, and the following statistics are far above the control limit. Compared with other methods, SmLRDR-wSVDD is more effective for the faults with slow drift in the budding stage.

As for FARs shown in Table 9, the proposed SmLRDR-wSVDD method has the lowest FARs in Cases 2, 3, and 4. Although its FARs are slightly higher than those of conventional SVDD for Cases 1 and 5, both of them are fewer than 0.03, which is too trivial to be considered.

Similarly, LRD, LDR, and mLRDR of each training sample are visualized in Figure 10 to explain the superiority of our proposed method. Blue, green, and red dots are from Mode 1, Mode 2, and outliers, respectively. Obviously, LRD changes with different modes, and that of outliers is close to normal in Figure 10(a). Although LDRs can remove multimodality as shown in Figure 10(b), the LDRs of outliers

Table 9: FAR under five cases on WWTP.

Cases no.	KPCA		SVDD	LDR-wSVDD	SmLRDR-wSVDD
	T^2	SPE			
Case 1	0.0061	0.0227	0.0208	0.0411	0.0236
Case 2	0.006	0.0149	0.0015	0.0104	0.006
Case 3	0.0045	0.0119	0.0045	0.0134	0.003
Case 4	0.006	0.0149	0.003	0.0149	0.0015
Case 5	0.0045	0.0104	0.0015	0.0015	0.0119

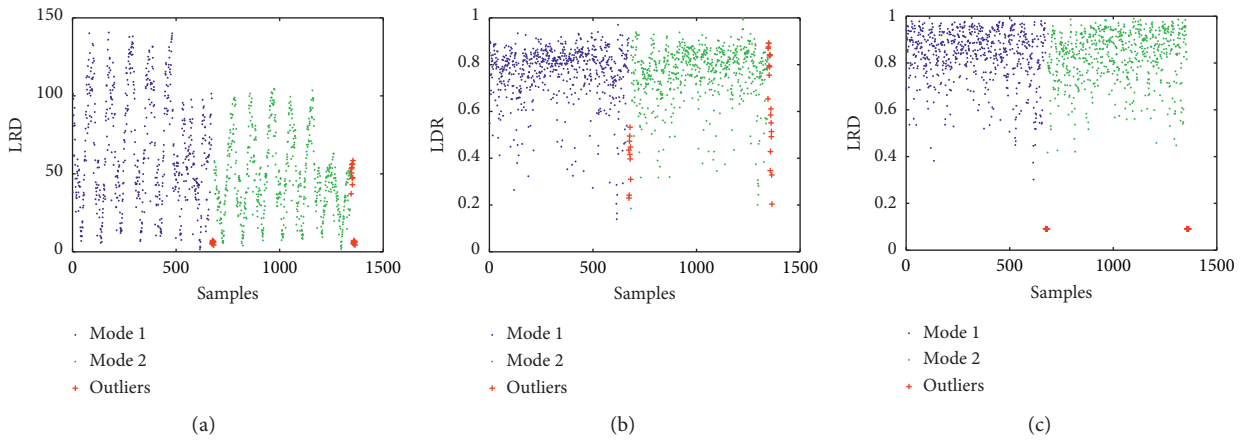


Figure 10: (a) LRD; the weight of each piece of training data based on (b) LDR and (c) mLRDR.

and normal samples are not distinguished. As a comparison, mLRDRs of outliers are far away from those of normal as shown in Figure 10(c). It is demonstrated once again that the proposed method can deal with multiple modes evenly and reduce the influence of outliers on modeling.

4. Conclusion

In this paper, a nonlinear multimode process monitoring method named SmLRDR-wSVDD is proposed to further improve the monitoring performance. Specifically, the SPA is used first to extract low- and high-order statistics; then, weight-SVDD is adopted to build a monitoring model, in which a new weight factor mLRDR is proposed based on LRD. Compared with the traditional methods, more features that are difficult to display are mined from raw data by the proposed method, and more comprehensive local space and time information is integrated into the weight factor, which can remove the multimodality uniformly and greatly weaken the influence of outliers on the boundary of the hypersphere. The applications in the TE process and WWTP have demonstrated that the proposed method is effective in nonlinear multimode process monitoring with lower MAR/FAR and higher fault sensitivity compared with the other three existing methods.

Data Availability

The data in the manuscript are generated according to http://depts.washington.edu/control/LARRY/TE/download.html and http://www.benchmarkwwtp.org/.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant 61873102.

References

[1] J. Zhu, Y. Yao, D. Li, and F. Gao, "Monitoring big process data of industrial plants with multiple operating modes based on Hadoop," *Journal of the Taiwan Institute of Chemical Engineers*, vol. 91, pp. 10–21, 2018.

[2] Y. Zheng, Y. Wang, H. Yan, Y. Wang, W. Yang, and B. Tao, "Density peaks clustering-based steady/transition mode identification and monitoring of multimode processes," *The Canadian Journal of Chemical Engineering*, vol. 98, no. 10, pp. 2137–2149, 2020.

[3] Z. Wang, Y. Zheng, and D. S.-H. Wong, "Trajectory-based operation monitoring of transition procedure in multimode process," *Journal of Process Control*, vol. 96, pp. 67–81, 2020.

[4] M. Quinones-Grueiro, A. Prieto-Moreno, C. Verde, and O. Lanes-Santiago, "Data-driven monitoring of multimode continuous processes: a review," *Chemometrics and Intelligent Laboratory Systems*, vol. 18, pp. 56–71, 2019.

[5] W. Du, Y. Fan, and Y. Zhang, "Multimode process monitoring based on data-driven method," *Journal of The Franklin Institute*, vol. 354, no. 6, pp. 2613–2627, 2017.

[6] S. Zhang and C. Zhao, "Stationarity test and Bayesian monitoring strategy for fault detection in nonlinear

multimode processes," *Chemometrics and Intelligent Laboratory Systems*, vol. 168, pp. 45–61, 2017.

[7] B. Song, S. Tan, and H. Shi, "Key principal components with recursive local outlier factor for multimode chemical process monitoring," *Journal of Process Control*, vol. 47, pp. 136–149, 2016.

[8] D. Ha, U. Ahmed, H. Pyun, C.-J. Lee, K. H. Baek, and C. Han, "Multi-mode operation of principal component analysis with k-nearest neighbor algorithm to monitor compressors for liquefied natural gas mixed refrigerant processes," *Computers & Chemical Engineering*, vol. 106, pp. 96–105, 2017.

[9] Y. Xu and X. Deng, "Fault detection of multimode non-Gaussian dynamic process using dynamic Bayesian independent component analysis," *Neurocomputing*, vol. 200, pp. 70–79, 2016.

[10] X. Peng, Y. Tang, W. Du, and F. Qian, "Multimode process monitoring and fault detection: a sparse modeling and dictionary learning method," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 6, pp. 4866–4875, 2017.

[11] Q. Jiang, B. Huang, and X. Yan, "GMM and optimal principal components-based Bayesian method for multimode fault diagnosis," *Computers & Chemical Engineering*, vol. 84, pp. 338–349, 2016.

[12] L. Ma, J. Dong, and K. Peng, "Root cause diagnosis of quality-related faults in industrial multimode processes using robust Gaussian mixture model and transfer entropy," *Neurocomputing*, vol. 285, pp. 60–73, 2018.

[13] L. Wang, C. Yang, and Y. Sun, "Multimode process monitoring approach based on moving window hidden markov model," *Industrial & Engineering Chemistry Research*, vol. 57, no. 1, pp. 292–301, 2017.

[14] M. Fang, H. Kodamana, B. Huang, and N. Sammaknejad, "A novel approach to process operating mode diagnosis using conditional random fields in the presence of missing data," *Computers & Chemical Engineering*, vol. 111, pp. 149–163, 2018.

[15] J. Zhu, Z. Ge, and Z. Song, "Recursive mixture factor analyzer for monitoring multimode time-variant industrial processes," *Industrial & Engineering Chemistry Research*, vol. 55, no. 16, pp. 4549–4561, 2016.

[16] D.-H. Hwang and C. Han, "Real-time monitoring for a process with multiple operating modes," *Control Engineering Practice*, vol. 7, no. 7, pp. 891–902, 1999.

[17] K. Zhang, K. Peng, and J. Dong, "A common and individual feature extraction-based multimode process monitoring method with application to the finishing mill process," *IEEE Transactions On Industrial Informatics*, vol. 14, no. 11, pp. 4841–4850, 2018.

[18] B. Song, S. Tan, H. Shi, and B. Zhao, "Fault detection and diagnosis via standardized k nearest neighbor for multimode process," *Journal of the Taiwan Institute of Chemical Engineers*, vol. 106, pp. 1–8, 2020.

[19] Y. Ma, B. Song, H. Shi, and Y. Yang, "Neighborhood based global coordination for multimode process monitoring," *Chemometrics and Intelligent Laboratory Systems*, vol. 139, pp. 84–96, 2014.

[20] Q. He, Q. Zhang, H. Wang, and C. Zhang, "Local similarity-based fuzzy multiple kernel one-class support vector machine," *Complexity*, vol. 2020, Article ID 8853277, 12 pages, 2020.

[21] D. M. J. Tax and R. P. W. Duin, "Support vector domain description," *Pattern Recognition Letters*, vol. 20, no. 11–13, pp. 1191–1199, 1999.

[22] Y. Zhang and X. Li, "Two-step support vector data description for dynamic, non-linear, and non-Gaussian processes monitoring," *The Canadian Journal of Chemical Engineering*, vol. 98, no. 10, pp. 2109–2124, 2020.

[23] G. Chen, X. Zhang, Z. J. Wang, and F. Li, "Robust support vector data description for outlier detection with noise or uncertain data," *Knowledge-Based Systems*, vol. 90, pp. 129–137, 2015.

[24] K. Z. Wang and H. Lan, "Robust support vector data description for novelty detection with contaminated data," *Engineering Applications of Artificial Intelligence*, vol. 91, pp. 1–10, 2020.

[25] F.-z. Zhao, B. Song, and H.-b. Shi, "Multi-mode process monitoring based on a novel weighted local standardization strategy and support vector data description," *Journal of Central South University*, vol. 23, no. 11, pp. 2896–2905, 2016.

[26] H. Li, H. Wang, and W. Fan, "Multimode process fault detection based on local density ratio-weighted support vector data description," *Industrial & Engineering Chemistry Research*, vol. 56, no. 9, pp. 2475–2491, 2017.

[27] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "Lof: Identifying density-based local outliersLof," *ACM SIGMOD Record*, vol. 29, no. 2, pp. 93–104, 2000.

[28] J. Wang and Q. P. He, "Multivariate statistical process monitoring based on statistics pattern analysis," *Industrial & Engineering Chemistry Research*, vol. 49, no. 17, pp. 7858–7869, 2010.

[29] Q. P. He, J. Wang, and D. Shah, "Feature space monitoring for smart manufacturing via statistics pattern analysis," *Computers & Chemical Engineering*, vol. 126, pp. 321–331, 2019.

[30] B. Zhou and X. Gu, "Multi-block statistics local kernel principal component analysis algorithm and its application in nonlinear process fault detection," *Neurocomputing*, vol. 376, pp. 222–231, 2020.

[31] X. Deng and X. Tian, "Nonlinear process fault pattern recognition using statistics kernel PCA similarity factor," *Neurocomputing*, vol. 121, pp. 298–308, 2013.

[32] B. Song and H. Shi, "Temporal-spatial global locality projections for multimode process monitoring," *IEEE Access*, vol. 6, pp. 9740–9749, 2018.

[33] Y. Xiao, H. Wang, L. Zhang, and W. Xu, "Two methods of selecting Gaussian kernel parameters for one-class SVM and their application to fault detection," *Knowledge-Based Systems*, vol. 59, pp. 75–84, 2014.

[34] F. Zhao, I. Rekik, S.-W. Lee, J. Liu, J. Zhang, and D. Shen, "Two-phase incremental kernel PCA for learning massive or online datasets," *Complexity*, vol. 2019, Article ID 5937274, 17 pages, 2019.

[35] J. J. Downs and E. F. Vogel, "A plant-wide industrial process control problem," *Computers & Chemical Engineering*, vol. 17, no. 3, pp. 245–255, 1993.

[36] Y. Wang, Y. Yao, Y. Zheng, and D. S. H. Wong, "Multi-objective monitoring of closed-loop controlled systems using adaptive Lasso," *Journal of the Taiwan Institute of Chemical Engineers*, vol. 56, pp. 84–95, 2015.

[37] N. L. Ricker, "Optimal steady-state operation of the Tennessee Eastman challenge process," *Computers & Chemical Engineering*, vol. 19, no. 9, pp. 949–959, 1995.

[38] J. Xu, J. Wang, I. Izadi, and T. Chen, "Performance assessment and design for univariate alarm systems based on FAR, MAR, and AAD," *IEEE Transactions on Automation Science and Engineering*, vol. 9, pp. 296–307, 2012.

[39] J. Alex, L. Benedetti, J. Copp, et al., Benchmark Simulation Model no.1 (BSM1), pp. 1–62, 2008.

[40] B. Wang, Z. Li, Z. Dai, N. Lawrence, and X. Yan, "Data-driven mode identification and unsupervised fault detection for nonlinear multimode processes," *IEEE Transactions On Industrial Informatics*, vol. 16, no. 6, pp. 3651–3661, 2020.

[41] C. K. Yoo, S. W. Choi, and I.-B. Lee, "Dynamic monitoring method for multiscale fault detection and diagnosis in MSPC," *Industrial & Engineering Chemistry Research*, vol. 41, no. 17, pp. 4303–4317, 2002.

Hindawi
Complexity
Volume 2021, Article ID 5539047, 11 pages
https://doi.org/10.1155/2021/5539047

Research Article

An Empirical Study on the Agglomeration Characteristics of China's Construction Industry Based on Spatial Autocorrelation and Spatiotemporal Transition

Likun Zhao, Junsen Tian, Yanqi Liu, and Rui Liu

Department of Civil Engineering, North China University of Technology, Beijing 100144, China

Correspondence should be addressed to Junsen Tian; 1412026763@qq.com

Received 28 January 2021; Revised 29 March 2021; Accepted 2 April 2021; Published 13 April 2021

Academic Editor: Hou-Sheng Su

The spatiotemporal agglomeration of industries is the most prominent geographical feature of economic activities. Based on the analysis of the spatiotemporal distribution of China's construction industry agglomeration, this paper analyzes the characteristics and evolution trend of the spatiotemporal agglomeration of construction industry in 31 provinces and cities of China from 2010 to 2019 by using Moran's index and the spatiotemporal transition measurement model. The findings are as follows: (1) China's construction industry has experienced two stages in terms of time: steady rise and turbulent rise. Spatially, China's construction industry, as a whole, the space takes the shape of one horizontal and two vertical, similar to the letter "H" being crossed. And the difference of "East-West" two ends of the industrial agglomeration level is obvious. (2) The Yangtze River Delta Urban Agglomerations (Shanghai, Jiangsu, and Zhejiang), the Pearl River Delta Urban Agglomerations (Guangdong), Beijing-Tianjin-Hebei Urban Agglomerations, and the western region (Xinjiang and Tibet) have significant local features. The four major types of China's construction industry cluster, which are H-H, H-L, L-H, and L-L, are formed. (3) The time-space transition of China's construction industry is dominated by the "stable transition" mode. The transition inertia is significant. The regional development has strong path dependence and spatial locking characteristics.

1. Introduction

Industrial geographic agglomeration is one of the significant characteristics of industrial spatial distribution. According to statistics, in France, the United Kingdom, and the United States, 75–95% of industrial production activities have been centralized, and the proportion of decentralized production activities is less than 15% (Maurel, 1999; Wen, 2004) [1, 2]. However, once industrial agglomeration is formed, it will not be fixed forever. When the industry gathers to a certain extent, the crowding cost will be generated, thus forming the centrifugal force to promote the industry outward transfer.

At present, the research on the temporal and spatial evolution of industrial agglomeration and the measurement of agglomeration level mainly focuses on the overall level of manufacturing industry, while the research on the evolution of space-time pattern of industrial agglomeration and the measurement of agglomeration level is less.

As for the research on the spatiotemporal evolution of industrial agglomeration, Krugman (1991) and Fujita et al. (1999) established FKV model to explain the spatial agglomeration phenomenon of industrial activities in the 1990s [3, 4]. Dirk (2006) and Ikeda et al. (2016) extended the clustering model to the geographical dimension and depicted the scale of spatial clustering area through endogenous changes in location [5, 6]. In addition, Hoffmann et al. (2017), Zhao et al. (2019), and Hu et al. (2019) analyzed the spatial clustering phenomenon in the service industry and other fields by using clustering index and other methods [7–10]. However, most statistics used to detect industrial spatial agglomeration cannot distinguish high and low spatial clustering. In order to fill this gap, Zhang and Lin (2010) introduced Moran's index to simulate the evolution

characteristics of real spatial agglomeration pattern [11], and Kang (2010) determined the spatiotemporal clustering threshold of agglomeration process based on Zhang [12]. The scholars comprehensively analyzed the agglomeration pattern of agriculture (Li, 2020; Zhao et al., 2019), manufacturing (He and Xie, 2006; Yi and Zhang, 2020), and service (Zhang et al., 2013; Li et al., 2019) by using relevant spatial statistical analysis methods [13–20]. Among them, some scholars also discussed the spatial clustering characteristics of the architectural industry (Tong et al., 2012; Sun et al., 2015; and Zhao et al., 2021) [21–23].

Measure study on industrial agglomeration level and single industrial concentration measurement method has a lot of literature research; there are mainly location entropy (Xie et al., 2015) [24], industrial geographical concentration (Chris and Declan, 2013) [25], Herfindahl–Hirschman index (Bruckmann, 1971) [26], locational Gini-coefficient (Dagum, 1986) [27], E-G index (Ellison and Glaeser, 1999) [28], multivariable intelligent analysis (Su et al., 2019; Su et al., 2020) [29, 30], K function (Ripley, 1977) [31], D-O index (Yuan et al., 2014) [32], Moran's I (Moran, 1950) [33], and Standard Deviational Ellipse (Wong, 1999) [34]. The research on common agglomeration measurement method is relatively late. Duranton and Overman (2008), Marcon and Puech (2009), and Lang et al. (2015), respectively, proposed the corresponding common agglomeration measure methods based on K function and M function (Overman, 2005) [35–38]. On the foundation of the study of Howard's coclustering index (Howard et al., 2016) [39], Kopczewska (2016) proposed a distance-based coclustering index [40]. Liu et al. used the above method to measure the spatial agglomeration level of manufacturing industry and other industries (Zheng, 2012; Qiu and Fang, 2013; Guo and Zhang, 2017; and Liu, 2019) [41–44].

It is found by the above literatures that the researches on theoretical innovation in foreign countries and China are relatively late. In the aspect of industrial agglomeration level and characteristics, the systematic results have not been formed. The research field mainly focuses on manufacturing, service industry, and agriculture, while the construction industry is involved less and mainly stays at the level of industrial development status. Few scholars have studied the spatial evolution characteristics of the building industry agglomeration, lacking analysis on spatial correlation and heterogeneity of the building industry.

Based on the building output value of panel data from China's 31 provinces, autonomous regions, and municipalities directly under the Central Government and Xinjiang Production and Construction Corps (excluding Hong Kong, Macao, and Taiwan) from 2010 to 2019, this paper uses spatial autocorrelation analysis to analyze the spatial agglomeration pattern of China's construction industry. Based on the LISA spatiotemporal transition analysis framework proposed by Rey et al., with the help of the comprehensive advantages of geography, this paper discusses the dynamic evolution of the local construction industry's spatial development. This study's conclusions have thoroughly investigated the temporal and spatial evolution characteristics of the agglomeration of the construction industry in China. The results of the study not only provide a data basis for the evaluation of the agglomeration development of the construction industry in China but also help to verify the implementation effects of the construction industry agglomeration policies and measures and determine the focus of future construction industry development regions and breakthrough points that enrich the theory of agglomeration of the construction industry. When discussing its regional differences and spatial structure, the time attribute and space attribute are combined, which more comprehensively reveals the dynamic differentiation law of the development of the construction industry and provides theoretical analysis of spatial economics and new economic geography to develop the construction industry perspective.

2. Modeling and Data Sources

2.1. Modeling. The empirical study on the spatiotemporal characteristics of China's construction industry agglomeration is divided into two parts: (1) The spatial and temporal distribution of construction industry is analyzed by natural breakpoint method. (2) Based on the spatiotemporal transition measure model and the global Moran's index and the local Moran's index in ESDA model, the spatial agglomeration characteristics of construction industry are summarized.

2.1.1. Spatial and Temporal Distribution Econometric Model of ArcGIS. The natural segment point method (Jenks, 1967) is a method of classifying data sets [45]. ArcGIS software was used to calculate the proportion of the output value of construction industry in 31 regions of China in the national output value of construction industry. The classification was carried out according to the natural segment point method, and the distribution of geographical space was analyzed. According to the natural segment point method, this article divides the indicators into five levels. The proportions of the representatives from the first to the fifth level are arranged from low to high as follows:

Grade 1: 0.001–0.009
Grade 2: 0.009–0.024
Grade 3: 0.024–0.048
Grade 4: 0.048–0.082
Grade 5: 0.082–0.133

2.1.2. Spatial Autocorrelation Analysis Model of Spatial Agglomeration Characteristics. Moran's index is an important index to measure spatial correlation which includes the global Moran's index and the Anselin local Moran's index. The range of Moran's index is [−1, 1]. Moran's index >0 means positive spatial correlation. The larger the value is, the more obvious the spatial correlation is. Moran's index <0 means negative spatial correlation. The smaller the value is, the greater the spatial difference is Moran index = 0; the spatial distribution is random.

In this paper, Moran's index is introduced to analyze the spatial correlation and dependence between construction industry clusters in China's neighboring regions:

(1) *The Global Moran's Index.* The global Moran's index (Anselin, 1995) describes the spatial characteristics of the construction industry development in the whole region [46]. It judges whether there is spatial agglomeration in the development of the construction industry. This is shown in the following formula:

$$I = \frac{\left(n\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\left(x_i - \overline{x}\right)\left(x_j - \overline{x}\right)\right)}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)}, \quad (1)$$

where Moran's index is the global spatial autocorrelation index of construction industry development. n is the 31 regions of China. X_i is the regional construction output value. X_i is the average value of construction output value. X_i is the spatial weight, which represents the potential interaction between regional construction industry development under the given significant level. Due to the special geographical location of Hainan, it is assumed that Hainan Province is adjacent to Guangdong Province and Guangxi Province.

(2) *The Anselin Local Moran's Index.* Compared with the global Moran's index, the local Moran's index (Anselin, 2010) can reflect the spatial dependence of an area and its surrounding areas [47]. The local Moran's index is used to analyze whether there is spatial heterogeneity in the development of construction industry. This is shown in the following formula:

$$I = \left[\frac{\left(x_i - \overline{x}\right)}{S}\right]\left[\sum_{j=1}^{n} w_{ij}\left(x_j - \overline{x}\right)\right], \quad (2)$$

where n, X_i, $\overline{x}$, and w_{ij} have the same meaning as in formula (1). The larger the absolute value of I, the higher the degree of spatial relevance of the subregions.

If I is positive, it means that local areas are low-low or high-high similarity value gathering areas; If I is negative, it means that local areas are low-high or high-low similarity value gathering areas.

The Moran's scatter plot is usually used to represent the local Moran's index. The first quadrant is H-H type; that is to say, the adjacent area of high concentration of construction industry is also high. The second quadrant is L-H type; that is to say, the adjacent areas of low industrial agglomeration area have higher agglomeration degree. The third quadrant is L-L type; that is to say, the agglomeration degree of neighboring areas of low industrial agglomeration areas is also low. The fourth quadrant is H-L type; that is to say, the agglomeration degree of adjacent areas of high industrial agglomeration areas is low. H-H type and L-L type indicate that the spatial correlation between regions is positive. L-H type and H-L type indicate that the spatial correlation between regions is negative.

2.1.3. Spatiotemporal Transition Measurement Model of Spatial Agglomeration Characteristics. Basing on the original research method, Rey (2010) embeds the distance, direction, concentration, and other attributes of the geographic coordinates of the research unit in Moran's I scatter chart into the traditional Markov chain in a specific time interval, proposing a local Markov transfer and spatiotemporal transitions, used to describe the spatiotemporal evolution of Moran's I scatter diagram, and divide the spatiotemporal transitions into 4 basic types: self-transitions (I), neighborhood transitions (II), and same-direction transitions according to the transition state between each region itself and its neighbors (III), and stable transition (IV) [48], as shown in Table 1.

Based on the division of temporal and spatial transition types in the development of the construction industry, the quantile regression model can be deeply nested with it. On the one hand, the quantile regression model of construction industry development can overcome the influence of strong distribution assumptions and outliers. On the other hand, the quantile regression results can not only examine the driving effect of the construction industry agglomeration on industrial development in stages, but also take into account the dynamic characteristics of the driving effect.

Assuming that the probability distribution of industrial agglomeration M *is* $F(M) = Prob(M \le m)$, the $\boldsymbol{\tau}$ quantile of M is defined as satisfying $q(\tau) = inf\{m|F\}(m) \ge \tau$, $0 < \tau < 1$, then the τ quantile $q(\tau)$ of $F(m)$ is based on the objective function that minimizes the absolute value of the weighted error. This is shown in the following formula:

$$q(\tau) = \arg\min \zeta \left\{\tau \int_{c>\zeta}^{1} |y - \zeta| dF(c) + (1-\tau)\int_{c<\zeta}^{2} |y - \zeta\ dF(c)|\right\}. \quad (3)$$

Specifically, according to the quantile division of the construction industry agglomeration pattern, this paper sets the high-level response and the low-level response in turn. Based on the different influence of the types of temporal and spatial transitions and driving factors in the development of the construction industry, the transition mechanism of high-low-level driving and high-low-level restriction can be constructed. The mechanism framework of temporal and spatial transition and nested analysis of construction industry development is shown in Figure 1.

2.2. Data Sources. The data in this paper come from two sources: (1) The spatial administrative boundary information comes from the National Basic Geographic Information Center. The raster data is extracted based on the research area vector boundary. (2) Social and economic data were obtained from the National Bureau of Statistics. Some missing data were supplemented by regional yearbook data.

Table 1: Basic classification types of spatiotemporal transition of construction industry development.

Type	Evolutionary manners	Type	Evolutionary manners
I: self-transition	$HH_t \longrightarrow LH_{t+1}$ $HL_t \longrightarrow LL_{t+1}$ $LH_t \longrightarrow HH_{t+1}$ $LL_t \longrightarrow HL_{t+1}$	III: same direction transition	$HH_t \longrightarrow LL_{t+1}$ $HL_t \longrightarrow LH_{t+1}$ $LL_t \longrightarrow HH_{t+1}$ $LH_t \longrightarrow HL_{t+1}$
II: neighborhood transition	$HH_t \longrightarrow HL_{t+1}$ $HL_t \longrightarrow HH_{t+1}$ $LH_t \longrightarrow LL_{t+1}$ $LL_t \longrightarrow LH_{t+1}$	IV: stable transition	$HH_t \longrightarrow HH_{t+1}$ $HL_t \longrightarrow HL_{t+1}$ $LH_t \longrightarrow LH_{t+1}$ $LL_t \longrightarrow LL_{t+1}$

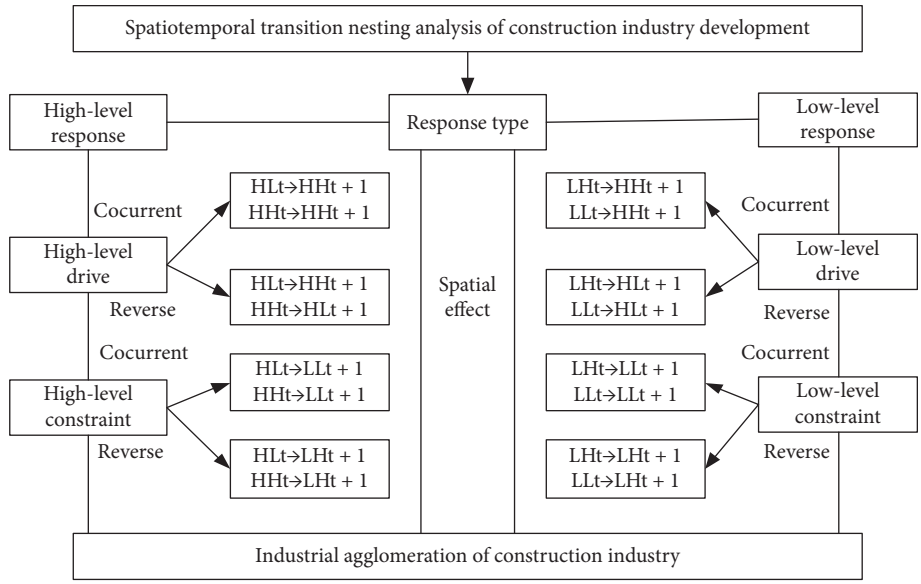


Figure 1: The mechanism of spatial and temporal transition and nested analysis of construction industry development.

This paper selects the relevant data of 31 provinces, autonomous regions, municipalities directly under the Central Government and Xinjiang Production and Construction Corps (excluding Hong Kong, Macao, and Taiwan, hereinafter referred to as the 31 regions) in China from 2010 to 2019 as the research scope.

3. The Present Situation of the Space-Time Distribution of China's Construction Industry Agglomeration

The development of the construction industry is influenced by geographical environment, economy, and other social factors. Based on the official data of The National Bureau of Statistics from 2010 to 2019, ArcGIS software was used to calculate the proportion of the output value of the construction industry in 31 regions of China in the national output value of construction industry, as shown in Figure 2.

Through the analysis of Figure 2, it is found the following:

On the spatial dimension, China's construction industry's overall spatial and temporal distribution presents a "丰" shape with one horizontal and two belts, and the "East-West" difference is obvious. Among them, "one horizontal line" refers to the northern horizontal line with Xinjiang, Gansu, Ningxia, Inner Mongolia, and Liaoning as the main areas. The "two belts" refer to the eastern coastal belt dominated by Shandong, Jiangsu, Shanghai, Zhejiang, Fujian, and Guangdong and the central axis belt dominated by Shaanxi, Hubei, and Chongqing. With the passage of time, the central region gradually showed "bulge" phenomenon.

The main reasons for the above phenomena are as follows:

(1) In the early stage, under the influence of eastern China's pioneering experiments and western development, the development of the construction industry in east of China, which was in the "gap," was relatively backward

(2) With the rise of central China and the further implementation of the western development strategy, China's construction industry began to shift to the west and central regions, and the concentration level of the construction industry in the central region was significantly improved, and the overall pattern showed an obvious "V" shape from decline to rise

In terms of time, the development trend of China's construction industry can be divided into two stages:

Phase I (2010–2015): during the period of steady rise, the proportion of construction output value increased steadily; Phase II (2016–2019): during the period of turbulence, the construction output ratio amplitude gradually increases, and the "instability" feature becomes increasingly prominent. In 2017, it shows a "cliff-like" decline.

The main reasons for the above phenomena are as follows:

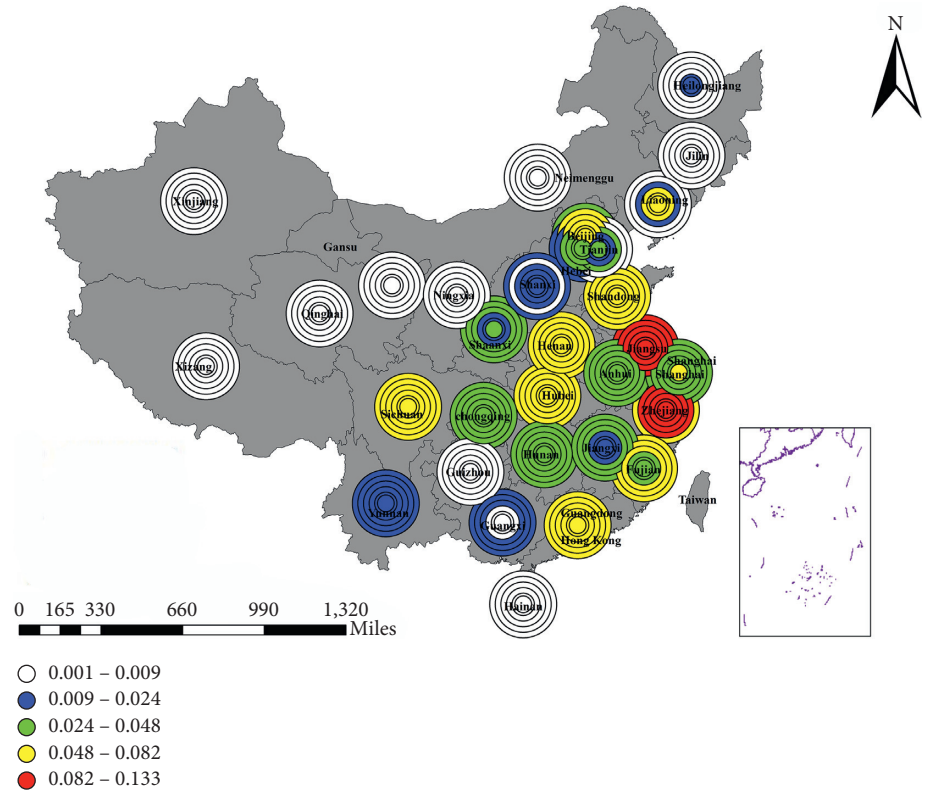


Figure 2: Time and space distribution of China's construction industry. *Note.* From the inside to the outside, each circle represents the period from 2010 to 2019, respectively. The proportion of construction output value decreases successively according to red, yellow, green, blue, and white.

(1) From 2010 to 2015, the eastern region took the lead in attracting the construction industry from Hong Kong, Macao, and developed countries by taking advantage of its unique geographical advantages and the opening-up policy. The construction industry developed rapidly and its agglomeration level reached the highest level in China. In recent years, with the gradual saturation of construction capital, the construction industry's development presents a downward trend.

(2) Given the most significant impact of construction dust on the regional environment, some regions should eliminate enterprises with large pollution emission in building materials and other industries. Affected by the economic downturn and the slowdown of real estate investment, the construction industry's development is slow.

4. Analysis of the Agglomeration Characteristics of China's Construction Industry

4.1. Analysis of the Spatial Correlation Characteristics of China's Construction Industry Agglomeration

4.1.1. Empirical Analysis of Global Spatial Correlation. Based on the statistical data of construction industry development in 31 regions of China from 2010 to 2019, the overall Moran's index is introduced to analyze the overall differentiation characteristics of China's construction industry spatial agglomeration. The measurement results are shown in Table 2.

According to the analysis in Table 2 about China's construction industry in the past 10 years,

(1) the global Moran's index is between 0 and 0.5, and the overall trend is increasing from small to large. It shows that the positive spatial autocorrelation characteristics of China's construction industry development area are apparent, the spatial correlation is gradually growing, and the regional connection is becoming closer;

(2) regional development as a whole presents the characteristics of high-high and low-low spatial agglomeration. It shows that the regional construction industry's development level is closely related to the development of the surrounding areas. On the one hand, it may be restricted by the development of the surrounding areas, and on the other hand, it will stimulate the development of the surrounding areas;

(3) the global Moran's index differs slightly at different times, further verifying that the development of the construction industry has good stability.

4.1.2. Empirical Analysis of Local Spatial Correlation. The global Moran's index mainly measures the overall agglomeration correlation degree of the development level of

Table 2: Global autocorrelation Moran's index of changes in China's construction industry.

Years	Moran's index	E (I)	P	Years	Moran's index	E (I)	P
2010	0.2714	−0.03	0.019	2015	0.4687	−0.03	0.001
2011	0.3494	−0.03	0.018	2016	0.2637	−0.03	0.020
2012	0.3874	−0.03	0.009	2017	0.2588	−0.03	0.020
2013	0.2199	−0.03	0.026	2018	0.3294	−0.03	0.013
2014	0.4458	−0.03	0.008	2019	0.3422	−0.03	0.024

China's construction industry, but it cannot reveal the agglomeration correlation degree of local areas in adjacent spaces. To further analyze the spatial agglomeration and differentiation characteristics of China's construction industry. This paper uses the local Moran's index to describe the spatial heterogeneity of the development of the construction industry in each study area and its adjacent areas. Generate Moran's scatter plot and LISA cluster plot to judge the spatial correlation degree of the construction industry's development level between regions. The measurement result is shown in Figure 3.

Through the analysis of Figure 3, the following is found:

Moran's index scatter plot shows obvious "high-low" and "low-high" agglomeration characteristics gradually weakening; "high-high" and "low-low" agglomeration characteristics are gradually strengthening. The spatial heterogeneity is weak, and it has a significant "Matthew effect." It shows that the agglomeration level of the regional construction industry is restricted and driven by the development of the surrounding area's construction industry and is highly consistent with the spatial agglomeration level of the surrounding area.

The LISA cluster map of China's construction industry is mainly divided into high-high agglomeration (H-H), low-high agglomeration (L-H), low-low agglomeration (L-L), and high-low agglomeration (H-L), as shown in Table 3.

(1) In terms of local space, Shandong, Jiangsu, Zhejiang, Shanghai, Xinjiang, Tibet, Beijing, Tianjin, Hebei, and Guangdong have significant statistical significance, indicating that the development of the construction industry in these regions has a guiding effect on the development of China's construction industry.

(2) H-H agglomerations are mainly distributed in the Yangtze River Delta Urban Agglomeration. These areas together form a "high-value club" for the development of the construction industry. L-L clusters are mainly distributed in northwest China, and the area is gradually shrinking. Hebei, Shanxi, and other areas in L-H have relatively high development in the construction industry, but they are easily affected by the negative effects of surrounding areas and weaken. Guangdong and other regions in the H-L region have a favorable location advantage of being "proliferated," and it is necessary to strengthen the connection with the surrounding provinces with low development of construction industry.

(3) Some regions fall into the second and fourth quadrants, indicating spatial differences in the development of China's construction industry. From the change of Moran's scatter chart and LISA cluster chart of the construction industry from 2010 to 2019, it can be seen that the development level index value of the construction industry falling in the first and third quadrants has increased, mainly in the eastern coastal cities. It shows that the eastern coastal areas of China rely on their unique location, economy, and transportation advantages; through the polarization effect, the building resources are continuously promoted to gather in the direction of "periphery-center"; as a result, the level of agglomeration of the construction industry continues to increase, the spatial connections between provinces and cities are getting closer, and the spatial heterogeneity continues to shrink.

4.2. China Building Industry Cluster Transition Characteristics of Time and Space Analysis. The researches on the spatial distribution characteristics of construction industry development are discussed in this paper; the spatial distribution is variable. Therefore, this paper adopts the time-space transition measure method for in-depth analysis and selects the stability index of spatial agglomeration of the construction industry, namely, spatial cohesion, to measure the transition state of the regional construction industry. Its value is equal to the proportion of a certain transition type in the study period's total number of transitions. According to the basic types of temporal and spatial transitions, the construction industry's local Moran's index transition probability matrix can be obtained, as shown in Table 4.

Through the analysis of Table 4, the following is found:

(1) The spatial structure of construction industry in 31 regions of China is stable as a whole, showing an obvious inert transition. The transition probability of type I, type II, and type III transition is only 0.18, while the space transition probability of type IV is 0.82, and the corresponding spatial cohesion is 82%, which indicates that the regional development of China's construction industry has strong path dependence and spatial locking.

(2) "Self-transition" and "neighborhood transition" are the two time-space transition modes that appear least frequently in China's construction industry's regional development. It shows that the development of the regional construction industry will have mutual influence on the surrounding areas. The construction industry's development factors will have a spillover effect and demonstration effect in the neighboring regions.

(3) The cotransition of both regions and their neighborhoods is the least common, which further verifies the atypical characteristics of the development of regional construction industry.

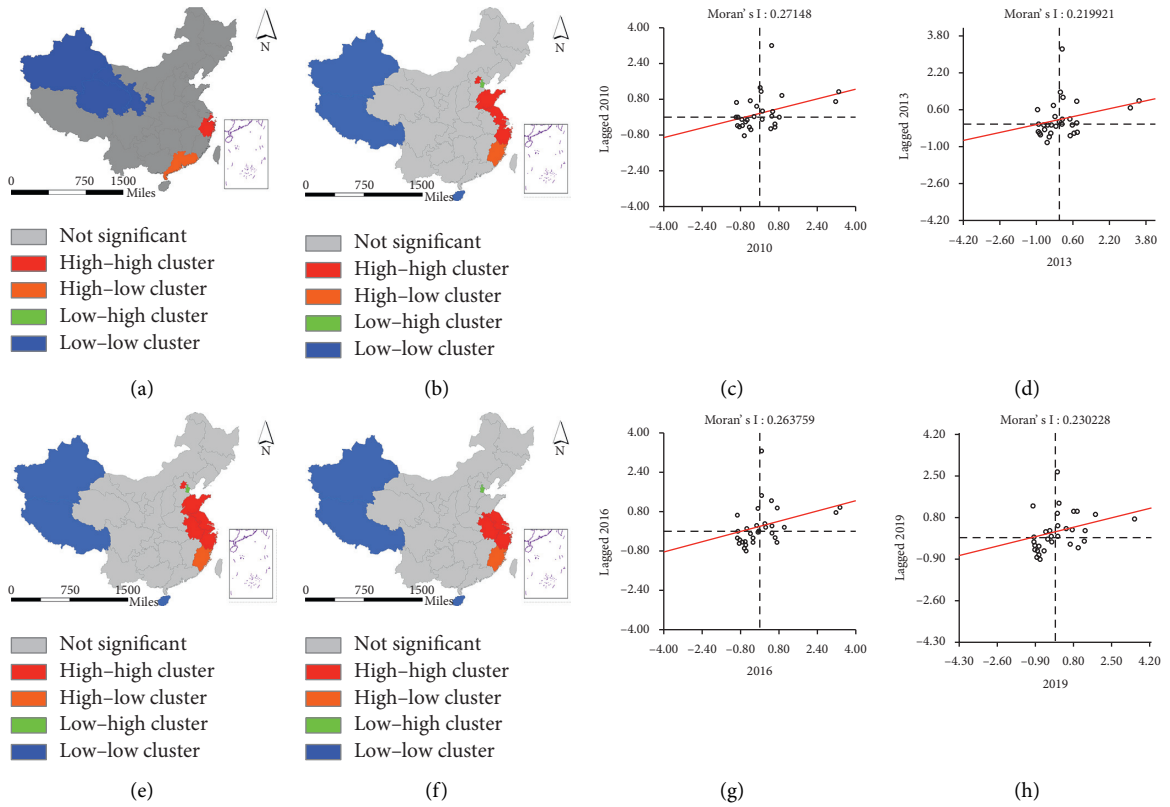


Figure 3: Spatial clustering characteristics of China's construction industry.

Table 3: Moran's index scattered distribution map province distribution.

Time	H-H	L-L	L-H	H-L	Cross quadrant
2010	Beijing, Tianjin, Shanghai, Zhejiang, Fujian, Jiangsu	Jilin, Heilongjiang, Inner Mongolia, Jiangxi, Hubei, Sichuan, Guizhou, Yunnan, Tibet, Xinjiang, Ningxia, Gansu, Qinghai, Shaanxi, Henan	Hubei, Shanxi, Guangxi	Anhui, Hunan, Guangdong, Chongqing, Liaoning	Hainan, Shandong
2013	−Fujian, +Shandong	+Anhui, +Hunan	+Tianjin	−Anhui, −Hunan	+Fujian, −Shandong
2016	+Shandong, −Tianjin	+Hainan, −Hubei, −Shaanxi		+Fujian, +Hubei, +Shaanxi	−Fujian, −Hainan
2019		+Liaoning, −Jiangxi		−Liaoning, +Jiangxi	

Table 4: Construction industry's local Moran's index transition probability matrix.

Type	H-H_{t+1}	H-L_{t+1}	L-L_{t+1}	L-H_{t+1}
H-H_t	IV (0.28)	II (0.011)	I (0.016)	III (0.005)
H-L_t	II (0.011)	IV (0.164)	III (0.003)	I (0.03)
L-L_t	I (0.008)	III (0.003)	IV (0.052)	II (0.016)
L-H_t	III (0.008)	I (0.03)	II (0.016)	IV (0.321)

Based on the nested analysis framework of spatiotemporal transition and spatial quantile regression of the construction industry, the driving paths of the construction industry in different regions are different, and the regions are further divided into four response modes, namely, low-level restriction, low-level drive, high-level restriction, and high-level drive. The empirical results are shown in Table 5.

Through the analysis of Table 4, the following is found: The development type of China's construction industry is mainly driven by high level and restricted by low level. The high-level types are mainly concentrated in the eastern developed areas, and the low-level types are mainly in the western region.

(i) Low-level constraints: except for Heilongjiang, other provinces are hindered by their own development and the construction industry in neighboring regions, and the development situation is not optimistic.

(ii) Low-level driving type: except for Sichuan, all other areas are driven in the same direction. At the initial stage, they are at a low level. However, both it and the neighboring areas are in good development

Table 5: China's construction industry development spatiotemporal response mode.

Pattern	Response patterns	The evolution way	Provinces
Self-transition	Low-level drive	$LL_t \longrightarrow HL_{t+1}$	Jiangxi, Shanxi
	High-level restriction	$HL_t \longrightarrow LL_{t+1}$ $HH_t \longrightarrow LH_{t+1}$	Hainan, Anhui, Hunan, Hubei
Neighborhood transition	High-level restriction	$HL_t \longrightarrow HH_{t+1}$	Shandong
	Low-level restriction	$LL_t \longrightarrow LL_{t+1}$	Ningxia, Henan, Qinghai, Gansu, Xinjiang, Tibet, Heilongjiang, Yunnan, Jilin, Inner Mongolia, Guizhou, Guangxi, Sichuan, Shanxi
Stable transition	Low-level drive	$LH_t \longrightarrow LH_{t+1}$	Hebei
	High-level restriction	$HL_t \longrightarrow HL_{t+1}$	Tianjin, Liaoning
	High-level drive	$HH_t \longrightarrow HH_{t+1}$	Beijing, Jiangsu, Zhejiang, Shanghai, Guangdong, Chongqing, Fujian

condition, and they are jointly transformed to a high-level state.

(iii) High-level constraint: Tianjin, Liaoning, and Henan all changed from high level to low level, but their adjacent units were not affected and remained at high level, which played a driving role in the development of low-level areas.

(iv) High-level driving type: except for Chongqing, other areas have a good development trend in the future, forming high-level agglomeration areas.

5. Conclusion and Discussion

5.1. Conclusion. Based on the relevant data of the total building output value of 31 regions in China, this study uses spatial correlation and temporal and spatial transition measurement models to study the spatial pattern of industrial agglomeration in China, and the following research conclusions are obtained:

(1) The agglomeration of China's construction industry presents strong spatial differences. Taking Heihe-Tengchong Line as the boundary, the distribution difference between southeast and northwest is obvious. The southeast always occupies a significant advantage.

(2) China's construction industry has significant positive spatial autocorrelation. Among them, H-H and L-L are the most important spatial clustering types. They show strong qualitative characteristics of spatial lock in spatial geographical distribution.

(3) The high-concentration areas of China's construction industry are mainly distributed in the Yangtze River Delta Urban Agglomeration and Pearl River Delta Urban Agglomeration, such as Guangdong, Zhejiang, and Jiangsu. The low-level agglomeration areas mainly include the Lanzhou-Xining Urban Agglomeration and the Ningxia Urban Agglomeration along the Yellow River. There are great regional differences in the development of the internal construction industry in western China, Sichuan, Shaanxi, and other regions that performed well. The construction industry in Tibet, Ningxia, and other regions has low-level agglomeration density and relatively backward development.

(4) The agglomeration of China's construction industry presents strong spatiotemporal path-dependent characteristics. In terms of time dimension, the construction industry in adjacent areas has significant positive correlation. In spatial dimension, the spillover effect of spatial agglomeration of construction industry is uncertain. It shows that China's construction industry can influence the industrial development of surrounding areas through demonstration. It also shows that China's construction industry is driven or restricted by the industrial development level of surrounding areas, forming the linkage effect between regions.

(5) The time-space transition of China's construction industry is dominated by "stable transition" mode. And it is supplemented by "self-transition" and "neighborhood transition" mode. The transition inertia is significant. The regional development has strong path dependence and spatial locking characteristics. It leads to the coexistence of "agglomeration" and "differentiation" in the space of the construction industry. The eastern region as a whole shows a high level of driving. The western region shows a stable low-level restriction state.

5.2. Discussion. Regarding the problem of industrial distribution and agglomeration, domestic and foreign researches are basically concentrated in general industries such as manufacturing and tertiary industries. Due to the uniqueness and uniqueness of the output of the construction industry, many scholars believe that the industrial agglomeration phenomenon of the general industry and the tertiary industry is unlikely to occur in the construction industry (Zhen et al., 2010) [49]. However, the economic interaction in regions, coupled with the factors such as industrial coordination and regional policies, has led to a gradually increasing in the regional development level of the construction industry (Liu et al., 2015) [50]. With the

upgrading and transformation of construction industrialization, as an important link in the process of industrialization, the phenomenon of industrial agglomeration has emerged. Xiao (2013), Zhang (2010), Dai (2010), Hong (2020), etc. used a variety of methods, respectively, to measure the development of the construction industry in different regions; it proved that the construction industry has a certain degree of agglomeration in space [51–54]. This provides theoretical and practical support for the research, but there are some shortcomings. There are few research results on the characteristics of the construction industry agglomeration. Most of the researches on the time evolution characteristics of the construction industry agglomeration lack spatial perspective, causing the research results to be more abstract.

This research mainly uses the latest construction industry-related data to describe the complete process of the spatial evolution of the Chinese construction industry in the past 10 years and summarizes the significant temporal and spatial differentiation of the spatial evolution of the Chinese construction industry in different regions. Through the analysis of the spatiotemporal pattern of construction industry agglomeration, it plays an important role in predicting the regional development trend, recognizing and grasping the regional development stage, adjusting the regional spatial development strategy, and improving the allocation of industrial resources in time, also in favor of promoting the urban industrial transformation and its spatial restructuring. When discussing its regional differences and spatial structure, it breaks the long-standing limitation of agglomeration characteristics with time as the axis. Based on the perspective of spatial and temporal dimensions, it provides theoretical analysis of spatial economics and new economic geography for the development of the construction industry, thus making up for the academic deficiencies.

Therefore, the empirical research conclusions of this paper have important guiding significance for promoting the improvement of the agglomeration level of China's construction industry, optimizing the path selection for the transition of the construction industry to a state of high agglomeration, and the layout of industrial development decisions.

First of all, the analysis of the space-time distribution of China's construction industry agglomeration will play a good role in guiding the macro layout of China's construction industry. This study makes a comprehensive comparison of the development of the construction industry in various regions of China and finds that the agglomeration of the construction industry in China presents strong spatial differences. Based on this difference, the regional development strategy of the construction industry should build a growth pole of the construction industry in the Yangtze River Delta Urban Agglomeration and Pearl River Delta Urban Agglomeration, with Jiangsu, Guangdong, and Zhejiang as the main regions. These areas are multicore central urban agglomeration with a higher degree of openness and a wider radiation range. Through the radiation linkage of multiple cores, they can promote the rapid development of the entire urban agglomeration and even across regions. With the implementation of the strategy of western development, the construction industry in western China has seen significant substantial growth, but there is still a gap between western China and eastern China. Therefore, the allocation of construction industry resources should be optimized and differentiated policies should be adopted to improve the unbalanced development of the construction industry. Sichuan in the western region should make use of its own advantages to become a pole of growth in the western region. By promoting the growth of other provinces in western China through the diffusion of resources and technologies, the construction industry can be transformed to a state of high concentration.

Secondly, through the analysis of the overall and local characteristics of the construction industry, it has guiding significance for the scientific formulation of the development policy of the construction industry. The research found that the spatial correlation between the construction industry in Chinese provinces was obvious. The development of the construction industry has formed a pattern of positive correlation between high output and high spatial correlation in the eastern region and low output and low spatial correlation in the central and western regions. The analysis results show that the neighbor effect is significant. Therefore, the development planning of the construction industry in each region should pay attention to the spatial characteristics of cross regions. From a broader point of view to study the coordinated development of regional architecture, actively seek cross-regional cooperation and mutual assistance to form a win-win situation. First of all, strengthening regional cooperation in the construction industry, especially in the developed areas to implement a development policy conducive to "rich neighbors," can bring win-win results. Secondly, low-level and low-growth areas appear in the form of agglomeration in space. Selecting agglomeration areas as the primary development object and adopting regional development policies that promote the common growth of agglomeration areas can produce better results. Thirdly, the spatial correlation of regional construction growth is affected by geographical factors such as relative location and distance. This indicates that sufficient attention should be paid to geographical factors (spatial correlation and heterogeneity) in the formulation of corresponding macroeconomic policies for construction. Only in this way can reasonable policies be formulated.

In sum, this research supports the exploration of China's construction industry agglomeration theory. Research perspectives and research methods are applicable to exploring the development of construction industry agglomeration in any country according to the development of national conditions to determine the specific research content.

A large part of the empirical research in this study uses spatial econometric analysis methods. This article selects the most commonly used distance spatial weight matrix in spatial measurement. This matrix only considers the distance relation between provinces and is not necessarily the most appropriate matrix form to measure the spatial arrangement consistency of the construction industry. Therefore, the future research can further explore the

construction and selection of spatial matrix and choose the most suitable analysis method for the actual situation of this study.

Data Availability

The total construction output value data of 31 regions in China can be obtained for free on the official website of the National Bureau of Statistics of China (https://data.stats.gov.cn/

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Beijing Municipal Natural Science Foundation (no. 9202006), the Excellent Talent Project of North China University of Technology in 2019 (no. 216051360020XN225/004), and the National Key Research and Development Program of China (no. 2016YFC0701808-02). The authors acknowledge and are grateful for the funding.

References

[1] F. Maurel and B. Sédillot, "A measure of the geographic concentration in French manufacturing industries," *Regional Science and Urban Economics*, vol. 29, no. 5, pp. 575–604, 1999.

[2] M. Wen, "Relocation and agglomeration of Chinese industry," *Journal of Development Economics*, vol. 73, no. 1, pp. 329–347, 2004.

[3] P. Krugman, "Increasing returns and economic geography," *Journal of Political Economy*, vol. 99, no. 3, pp. 483–499, 1991.

[4] M. Fujita, P.-R. Krugman, and A.-J. Venables, *The Spatial Economy: Cities, Regions and International Trade*, The MIT Press, Cambridge, MA, USA, 1999.

[5] S. Dirk, "Erratum: where do cities form? A geographical agglomeration model for Europe," *Journal of Regional Science*, vol. 46, no. 3, pp. 602-603, 2006.

[6] K. Ikeda, K. Murota, and Y. Takayama, "Stable economic agglomeration patterns in two dimensions: beyond the scope of central place theory," *Journal of Regional Science*, vol. 57, no. 1, pp. 132–172, 2017.

[7] G. H. Searle, "Changes in producer services location, Sydney: globalisation, technology and labour," *Asia Pacific Viewpoint*, vol. 39, no. 2, pp. 237–255, 1998.

[8] J. Hoffmann, S. Hirsch, and J. Simons, "Identification of spatial agglomerations in the German food processing industry," *Papers in Regional Science*, vol. 96, no. 1, pp. 139–162, 2017.

[9] X. Zhao, M. Zhang, Y. Li, X. Huang, and L. Zhang, "Urban residential land expansion and agglomeration in China: a spatial analysis approach," *Environment Development and Sustainability*, vol. 22, no. 1, pp. 5317–5335, 2019.

[10] C.-H. Hu, W.-D. Liu, Y.-Q. Jia, and Y.-Y. Jin, "Characterization of territorial spatial agglomeration based on POI data: a case study of Ningbo City, China," *Sustainability*, vol. 11, no. 18, pp. 5083–5097, 2019.

[11] T.-L. Zhang and L. Ge, "A supplemental indicator of high-value or low-value spatial clustering," *Geographical Analysis*, vol. 38, no. 2, pp. 209–225, 2010.

[12] H. Kang, "Detecting agglomeration processes using space-time clustering analyses," *The Annals of Regional Science*, vol. 45, no. 2, pp. 291–311, 2010.

[13] E.-L. Li, "Evolution and innovation and development mechanism of agricultural industry cluster in China-taking Shouguang model vegetable industry cluster as an example," *Geoscience*, vol. 40, no. 4, pp. 617–627, 2020.

[14] J.-W. Zhao, Y.-F. Chen, L. Yu, and C.-B. Yin, "Spatial and temporal characteristics and influencing factors of pig breeding industry in China," *Economic Geography*, vol. 39, no. 2, pp. 180–189, 2019.

[15] C.-F. He and F.-H. Pan, "The trends of geographical agglomeration of manufacturing sectors in China and the explanations," *South China Journal of Economics*, vol. 6, pp. 38–52, 2011.

[16] C.-F. He and X.-Z. Xie, "Geographical concentration and provincial specialization of Chinese manufacturing industries," *Journal of Geography*, vol. 61, no. 2, pp. 212–222, 2006.

[17] X.-C. Bi, R.-X. Gu, Q. Su, and B.-Q. Lin, "The agglomeration of manufacturing industry and the evolution of geographical pattern in Jiangsu province," *East China Economic Management*, vol. 32, no. 7, pp. 12–21, 2018.

[18] B.-Z. Yi and C.-Y. Zhang, "A study of spatio-temporal pattern evolution of Chinese high-tech industry cluster," *Study & Exploration*, vol. 297, no. 4, pp. 115–121, 2020.

[19] X. Zhang, E.-S. Zhong, X.-H. Zhang, and S. H. Wang, "Spatial distribution and clustering of commercial network in Beijing during 2004-2008," *Progress in Geography*, vol. 32, no. 8, pp. 1207–1215, 2013.

[20] G. Li, W.-Y. Chen, L. Yang, Q. Liu, and X. Chen, "Spatial pattern and agglomeration mode of parcel collection and delivery points in Wuhan city," *Progress in Geography*, vol. 38, no. 3, pp. 407–416, 2019.

[21] X. Tong, T. Zou, and Y.-P. Tong, "The study of Jiangxi province's construction-spatial-cluster based on the analysis of Gini coefficient," *Science and Technology Management Research*, vol. 7, no. 32, pp. 187–190, 2012.

[22] J.-D. Sun, X. Chen, and Q. Nie, "Analysis of industrial agglomeration effect in Shanghai construction industry," *Shanghai Management Science*, vol. 3, pp. 87–91, 2015.

[23] L.-K. Zhao, Y.-Q. Liu, and J.-S. Tian, "Spatiotemporal evolution law and output prediction of construction waste in the people's republic of China," *Waste Management & Research*, Article ID 985605, 2021.

[24] M. Xie, H.-Y. Zhao, N.-N. Zhu et al., "An analysis on the agglomeration of tertiary industries in Zhejiang province," *Economic Geography*, vol. 35, no. 9, pp. 96–102, 2015.

[25] C.-V. Egeraat and D. Curran, "Spatial concentration in the Irish pharmaceutical industry: the role of spatial planning and agglomeration economies," *Tijdschrift voor Economische en Sociale Geografie*, vol. 104, no. 3, pp. 338–358, 2013.

[26] G. Bruckmann, "A generalized concept of concentration and its measurement," *Statistische Hefte*, vol. 12, no. 3-4, pp. 204–223, 1971.

[27] C. Dagum, "A new approach to the decomposition of the Gini income inequality ratio," *Empirical Economics*, vol. 22, no. 4, pp. 515–531, 1997.

[28] G. Ellison and E.-L. Glaeser, "Geographic concentration in U.S. Manufacturing industries: a dartboard approach," *Journal of Political Economy*, vol. 105, no. 5, pp. 889–927, 1999.

[29] Y.-F. Liu and H.-S. Su, "Containment control of second-order multi-agent systems via intermittent sampled position data communication," *Applied Mathematics and Computation*, vol. 362, Article ID 124522, 2019.

[30] Y. Liu and H. Su, "Some necessary and sufficient conditions for containment of second-order multi-agent systems with sampled position data," *Neurocomputing*, vol. 378, pp. 228–237, 2020.

[31] B. D. Ripley, "Modelling spatial patterns," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 2, pp. 172–192, 1977.

[32] H.-H. Yuan, H. Zhang, and H.-Y. Zeng, "Measuring localization of manufacturing industries and its dynamics-using Beijing firm-level data," *China Industrial Economics*, vol. 9, pp. 38–50, 2014.

[33] P. A. P. Moran, "Notes on continuous stochastic phenomena," *Biometrika*, vol. 37, no. 1-2, pp. 17–23, 1950.

[34] D. W. S. Wong, "Several fundamentals in implementing spatial statistics in GIS:using centrographic measures as examples," *Annals of GIS*, vol. 5, no. 2, pp. 163–174, 1999.

[35] G. Duranton and H. G. Overman, "Testing for localization using micro-geographic data," *The Review of Economic Studies*, vol. 72, no. 4, pp. 1077–1106, 2005.

[36] G. Duranton and H. G. Overman, "Exploring the detailed location patterns of U.K. manufacturing industries using microgeographic data∗," *Journal of Regional Science*, vol. 48, no. 1, pp. 213–243, 2008.

[37] E. Marcon and F. Puech, "Measures of the geographic concentration of industries: improving distance-based methods," *Journal of Economic Geography*, vol. 10, no. 5, pp. 745–762, 2009.

[38] G. Lang, E. Marcon, and F. Puech, "Distance-based measures of spatial concentration: introducing a relative density function," *The Annals of Regional Science*, vol. 64, no. 2, p. 243, 2020.

[39] E. Howard, C. Newman, and F. Tarp, "Measuring industry coagglomeration and identifying the driving forces," *Journal of Economic Geography*, vol. 16, no. 5, pp. 1055–1078, 2016.

[40] K. Kopczewska, "Comment to XCL co-agglomeration index: distance-weighted improved DCL index," *Papers in Regional Science*, vol. 95, no. 4, pp. 903–910, 2016.

[41] W. Zhen, "Spatial agglomeration level measurement and evaluation of manufacturing in Fujian Province," *Economic Geography*, vol. 32, no. 7, pp. 74–80, 2012.

[42] L. Qiu and C.-L. Fang, "Comprehensive assessment on the spatial agglomeration of producer services in Beijing," *Economic Geography*, vol. 32, no. 1, pp. 99–110, 2013.

[43] Q.-B. Guo and Z.-H. Zhang, "Spatiotemporal evolution of factors aggregating ability in urban agglomeration in the middle reaches of the Yangtze River," *Acta Geographica Sinica*, vol. 72, no. 10, pp. 1746–1761, 2017.

[44] X.-W. Liu, "Measurement on China's service industry agglomeration and urban productivity: taking 233 prefectural and above level cities for example," *Urban Problems*, vol. 11, pp. 59–68, 2019.

[45] G.-F. Jenks, "The data model concept in statistical mapping," *International Yearbook of Cartography*, vol. 7, pp. 186–190, 1967.

[46] L. Anselin and R.-J. Florax, "Small sample properties of tests for spatial dependence in regression models: some further results," in *New Directions in Spatial Econometrics*, pp. 21–74, Springer-Verlag, Berlin, Germany, 1995.

[47] L. Anselin, "Local indicators of spatial association-lisa," *Geographical Analysis*, vol. 27, no. 2, pp. 93–115, 2010.

[48] S. J. Rey, "Spatial empirics for economic growth and convergence," *Geographical Analysis*, vol. 33, no. 3, pp. 195–214, 2010.

[49] X. Zheng and Z.-N. Liu, "Considerations on the gini coefficient and industrial agglomeration development of the construction industry in the eight Southeastern Provinces," in *Proceedings of the 2010 International Conference On Information Technology & Scientific Management*, IEEE, Tianjin, China, December 2010.

[50] B.-S. Liu and B. Xue, "Analysis of the three-dimensional system of industry linkages with spatial difference for the Chinese regional construction industry," *Journal of Chongqing University (Social Science Edition)*, vol. 21, no. 1, pp. 16–22, 2015.

[51] L. Xiao and Y.-C. Guo, "Spatial agglomeration analysis of total output value of construction industry in East China," *Co-Operative Economy & Science*, vol. 3, pp. 24-25, 2013.

[52] M. Zhang, D.-Y. Qiao, and Y.-A. Dai, "A spatial econometric analysis of the construction's differences in China's provincial regions," *Scientific Decision Making*, vol. 3, pp. 87–93, 2010.

[53] Y.-A. Dai and C. Chen, "Spatial econometric analysis of regional differences in China's construction industry development," *Statistics & Information Forum*, vol. 5, pp. 53–58, 2010.

[54] J.-K. Hong, Z.-J. Huang, J.-D. Yao, and C.-Y. Wang, "Study on the construction agglomeration of the Yangtze River economic belt and its spillover effect," *Construction Economy*, vol. 41, no. 12, pp. 21–25, 2020.