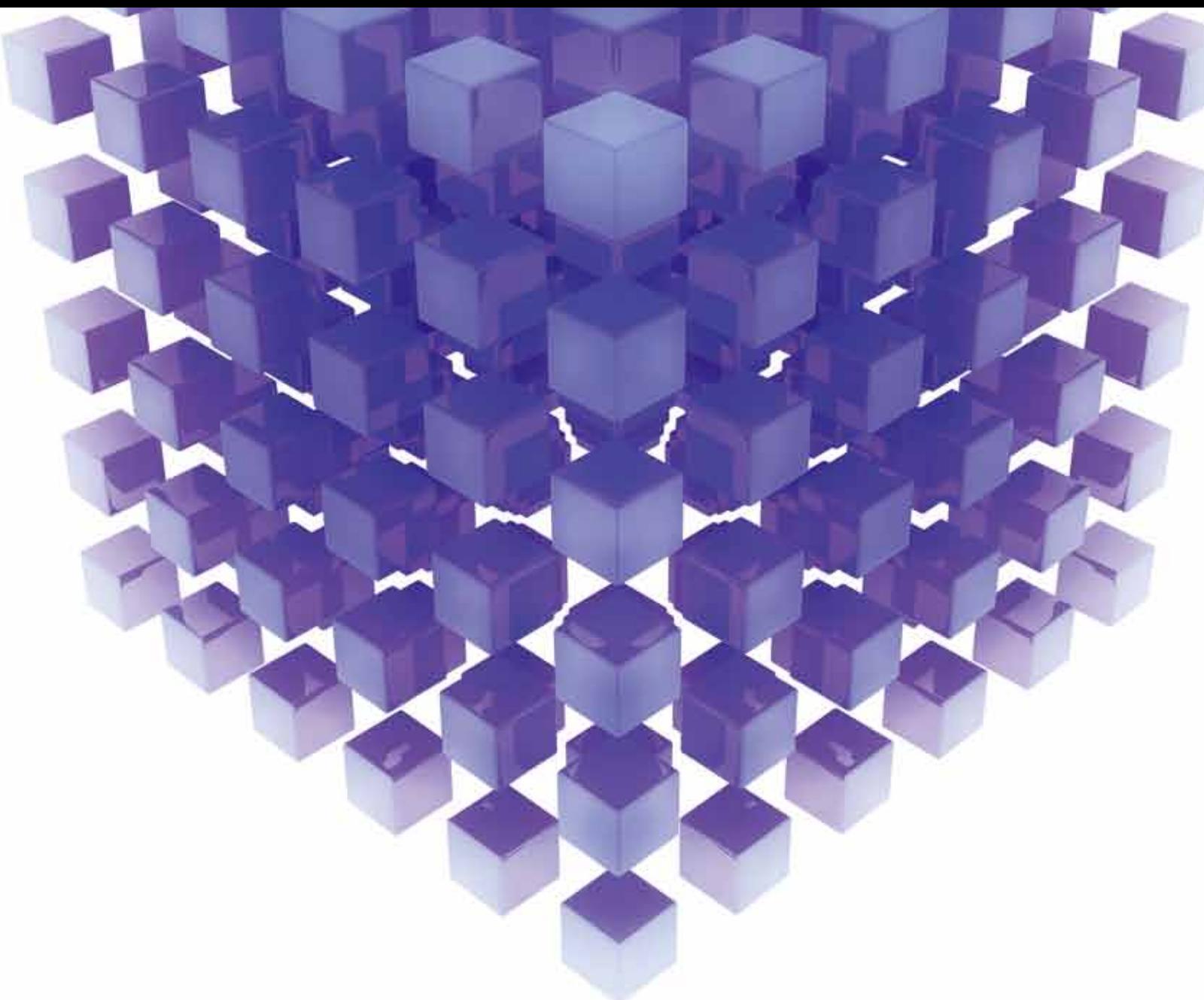


ADVANCED TOPICS IN DYNAMICS OF COMPLEX SYSTEMS

GUEST EDITORS: DUMITRU BALEANU, BASHIR AHMAD, JUAN J. NIETO,
AND J. A. TENREIRO MACHADO





Advanced Topics in Dynamics of Complex Systems

Mathematical Problems in Engineering

Advanced Topics in Dynamics of Complex Systems

Guest Editors: Dumitru Baleanu, Bashir Ahmad,
Juan J. Nieto, and J. A. Tenreiro Machado



Copyright © 2014 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in "Mathematical Problems in Engineering." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

- M. Abd El Aziz, Egypt
E. M. Abdel-Rahman, Canada
R. K. Abu Al-Rub, USA
Sarp Adali, South Africa
Salvatore Alfonzetti, Italy
Igor Andrianov, Germany
Sebastian Anita, Romania
W. Assawinchaichote, Thailand
Erwei Bai, USA
Ezzat G. Bakhoum, USA
José M. Balthazar, Brazil
R. K. Bera, India
C. Bérenguer, France
Jonathan N. Blakely, USA
Stefano Boccaletti, Spain
Stephane P.A. Bordas, USA
Daniela Bosco, Italy
M. Boutayeb, France
Michael J. Brennan, UK
Salvatore Caddemi, Italy
Piermarco Cannarsa, Italy
Jose E. Capilla, Spain
Carlo Cattani, Italy
Marcelo M. Cavalcanti, Brazil
Diego J. Celentano, Chile
Mohammed Chadli, France
Arindam Chakraborty, USA
Yong-Kui Chang, China
Michael J. Chappell, UK
Kui Fu Chen, China
Xinkai Chen, Japan
Kue-Hong Chen, Taiwan
Jyh-Horng Chou, Taiwan
Slim Choura, Tunisia
Cesar Cruz-Hernandez, Mexico
Swagatam Das, India
Filippo de Monte, Italy
Antonio Desimone, Italy
Yannis Dimakopoulos, Greece
Baocang Ding, China
Joao B. R. Do Val, Brazil
Daoyi Dong, Australia
B. Dubey, India
Horst Ecker, Austria
M. Onder Efe, Turkey
Elmetwally Elabbasy, Egypt
A. Elías-Zúñiga, Mexico
Anders Eriksson, Sweden
Vedat S. Erturk, Turkey
Moez Feki, Tunisia
Ricardo Femat, Mexico
Roberto A. Valente, Portugal
C. Fuerte-Esquivel, Mexico
Zoran Gajic, USA
Ugo Galvanetto, Italy
Furong Gao, Hong Kong
Xin-Lin Gao, USA
Behrouz Gatmiri, Iran
Oleg V. Gendelman, Israel
Didier Georges, France
Paulo B. Gonçalves, Brazil
Oded Gottlieb, Israel
Fabrizio Greco, Italy
Quang Phuc Ha, Australia
M. R. Hajj, USA
Tony S. W. Hann, Taiwan
Thomas Hanne, Switzerland
K. R. (Stevanovic) Hedrih, Serbia
M.I. Herreros, Spain
Wei-Chiang Hong, Taiwan
Jaromir Horacek, Czech Republic
Huabing Huang, China
Chuangxia Huang, China
Gordon Huang, Canada
Yi Feng Hung, Taiwan
Hai-Feng Huo, China
Asier Ibeas, Spain
Anuar Ishak, Malaysia
Reza Jazar, Australia
Zhijian Ji, China
Jun Jiang, China
J. J. Judice, Portugal
Tadeusz Kaczorek, Poland
Tamas Kalmar-Nagy, USA
Tomasz Kapitaniak, Poland
Hamid Reza Karimi, Norway
Metin O. Kaya, Turkey
Nikolaos Kazantzis, USA
Farzad Khani, Iran
K. Krabbenhoft, Australia
Ren-Jieh Kuo, Taiwan
Jurgen Kurths, Germany
Claude Lamarque, France
Usik Lee, Korea
Marek Leflik, Poland
Stefano Lenci, Italy
Roman Lewandowski, Poland
Shanling Li, Canada
Ming Li, China
Jian Li, China
Shihua Li, China
Teh-Lu Liao, Taiwan
Panos Liatsis, UK
Shueei M. Lin, Taiwan
Yi-Kuei Lin, Taiwan
Jui-Sheng Lin, Taiwan
Yuji Liu, China
Wanquan Liu, Australia
Bin Liu, Australia
Paolo Lonetti, Italy
V. C. Loukopoulos, Greece
Junguo Lu, China
Chien-Yu Lu, Taiwan
Alexei Mailybaev, Brazil
Manoranjan K. Maiti, India
O. D. Makinde, South Africa
R. Martinez-Guerra, Mexico
Driss Mehdi, France
Roderick Melnik, Canada
Xinzhu Meng, China
Yuri V. Mikhlin, Ukraine
G. Milovanovic, Serbia
Ebrahim Momoniat, South Africa
Trung Nguyen Thoi, Vietnam
Hung Nguyen-Xuan, Vietnam
Ben T. Nohara, Japan
Sotiris K. Ntouyas, Greece
Gerard Olivar, Colombia
Claudio Padra, Argentina
Bijaya Ketan Panigrahi, India
Francesco Pellicano, Italy
Matjaz Perc, Slovenia
Vu Ngoc Phat, Vietnam
M. do Rosário Pinho, Portugal
A. Pogromsky, The Netherlands

Seppo Pohjolainen, Finland	Cheng Shao, China	Moran Wang, China
Stanislav Potapenko, Canada	Bo Shen, Germany	Yijing Wang, China
Sergio Preidikman, USA	Jian-Jun Shu, Singapore	Gerhard-Wilhelm Weber, Turkey
Carsten Proppe, Germany	Zhan Shu, UK	J. A. S. Witteveen, The Netherlands
Hector Puebla, Mexico	Dan Simon, USA	Kwok-Wo Wong, Hong Kong
Justo Puerto, Spain	Luciano Simoni, Italy	Ligang Wu, China
Dane Quinn, USA	Grigori M. Sisoev, UK	Zhengguang Wu, China
K. R. Rajagopal, USA	Christos H. Skiadas, Greece	Gongnan Xie, China
Gianluca Ranzi, Australia	Davide Spinello, Canada	Wang Xing-yuan, China
Sivaguru Ravindran, USA	Sri Sridharan, USA	Xi Frank Xu, USA
G. Rega, Italy	Rolf Stenberg, Finland	Xuping Xu, USA
Pedro Ribeiro, Portugal	Changyin Sun, China	Jun-Juh Yan, Taiwan
J. Rodellar, Spain	Jitao Sun, China	Xing-Gang Yan, UK
R. Rodriguez-Lopez, Spain	Xi-Ming Sun, China	Suh-Yuh Yang, Taiwan
A. J. Rodriguez-Luis, Spain	Andrzej Swierniak, Poland	Mahmoud T. Yassen, Egypt
Ignacio Romero, Spain	Yang Tang, Germany	Mohammad I. Younis, USA
Hamid Ronagh, Australia	Allen Tannenbaum, USA	Bo Yu, China
Carla Roque, Portugal	Cristian Toma, Romania	Huang Yuan, Germany
Rubén R. García, Spain	Irina N. Trendafilova, UK	S.P. Yung, Hong Kong
Manouchehr Salehi, Iran	Alberto Trevisani, Italy	Ion Zaballa, Spain
Miguel A. Sanjuán, Spain	Jung-Fa Tsai, Taiwan	Ashraf M. Zenkour, Saudi Arabia
Ilmar F. Santos, Denmark	K. Vajravelu, USA	Jianming Zhan, China
Nickolas S. Sapidis, Greece	Victoria Vampa, Argentina	Xu Zhang, China
E. J. Sapountzakis, Greece	Josep Vehi, Spain	Yingwei Zhang, China
Bozidar Sarler, Slovenia	Stefano Vidoli, Italy	Lu Zhen, China
Andrey V. Savkin, Australia	Xiaojun Wang, China	Liancun Zheng, China
Massimo Scalia, Italy	Dan Wang, China	Jian Guo Zhou, UK
Mohamed A. Seddeek, Egypt	Youqing Wang, China	Zexuan Zhu, China
A. P. Seyranian, Russia	Yongqi Wang, Germany	Mustapha Zidi, France
Leonid Shaikh, Ukraine	Cheng C. Wang, Taiwan	

Contents

Advanced Topics in Dynamics of Complex Systems, Dumitru Baleanu, Bashir Ahmad, Juan J. Nieto, and J. A. Tenreiro Machado
Volume 2014, Article ID 354712, 1 page

Bidirectional Dynamic Diversity Evolutionary Algorithm for Constrained Optimization, Weishang Gao, Cheng Shao, and Yi An
Volume 2013, Article ID 762372, 13 pages

A New Hybrid Chaotic Map and Its Application on Image Encryption and Hiding, Yang Cao
Volume 2013, Article ID 728375, 13 pages

Modeling of an EDLC with Fractional Transfer Functions Using Mittag-Leffler Equations, J. J. Quintana, A. Ramos, and I. Nuez
Volume 2013, Article ID 807034, 7 pages

Dynamic Analysis of a Two-Language Competitive Model with Control Strategies, Lin-Fei Nie, Zhi-Dong Teng, Juan J. Nieto, and Il Hyo Jung
Volume 2013, Article ID 654619, 13 pages

Multiscale Characterization of Sea Clutter by Scale-Dependent Lyapunov Exponent, Jing Hu and Jianbo Gao
Volume 2013, Article ID 584252, 9 pages

Restoration of Natural Frequency of Cracked Cantilever Beam Using CNT Composite Patch: A Finite Element Study, Mahmoud Nadim Nahas and Mahmoud Ali Alzahrani
Volume 2013, Article ID 678412, 6 pages

Modeling and Flocking Consensus Analysis for Large-Scale UAV Swarms, Li Bing, Li Jie, and Huang KeWei
Volume 2013, Article ID 368369, 9 pages

A Microbial Fuel Cell Modified with Carbon Nanomaterials for Organic Removal and Denitrification, Njud S. Alharbi
Volume 2013, Article ID 410916, 9 pages

Study of Intelligent Photovoltaic System Fault Diagnostic Scheme Based on Chaotic Signal Synchronization, Chin-Tsung Hsieh and Jen Shiu
Volume 2013, Article ID 816296, 6 pages

Epidemic Modelling by Ripple-Spreading Network and Genetic Algorithm, Jian-Qin Liao, Xiao-Bing Hu, Ming Wang, and Mark S. Leeson
Volume 2013, Article ID 506240, 11 pages

On Classification of PDZ Domains: A Computational Study, Wasim Aftab, Adnan Memic, and Dumitru Baleanu
Volume 2013, Article ID 419485, 9 pages

Vibration Suppression of an Axially Moving String with Transverse Wind Loadings by a Nonlinear Energy Sink, Ye-Wei Zhang, Jian Zang, Tian-Zhi Yang, Bo Fang, and Xin Wen
Volume 2013, Article ID 348042, 7 pages

Modeling and Analyzing Operational Decision-Making Synchronization of C2 Organization in Complex Environment, Zou Zhigang, Liu Fuxian, Zhang Bo, and Li Xiang
Volume 2013, Article ID 645651, 17 pages

Windowed Least Square Algorithm Based PMSM Parameters Estimation, Song Wang
Volume 2013, Article ID 131268, 11 pages

θ -Metric Space: A Generalization, Farshid Khojasteh, Erdal Karapınar, and Stojan Radenović
Volume 2013, Article ID 504609, 7 pages

Synchronization Analysis of Four Symbolic Complex Dynamical Systems for Future Biology Research, Yagang Zhang, Zheng Zhao, and Yi Sun
Volume 2013, Article ID 685725, 7 pages

Swarm Stability Analysis of High-Order Linear Time-Invariant Singular Multiagent Systems, Mingyu Fu, Shimin Wang, and Yirui Cong
Volume 2013, Article ID 469747, 12 pages

A New Neumann Series Method for Solving a Family of Local Fractional Fredholm and Volterra Integral Equations, Xiao-Jing Ma, H. M. Srivastava, Dumitru Baleanu, and Xiao-Jun Yang
Volume 2013, Article ID 325121, 6 pages

Multidimensional Scaling for Orthodontic Root Resorption, Cristina Teodora Preoteasa, Ecaterina Ionescu, Elena Preoteasa, J. A. Tenreiro Machado, Mihaela Cristina Baleanu, and Dumitru Baleanu
Volume 2013, Article ID 383698, 6 pages

Systems of Navier-Stokes Equations on Cantor Sets, Xiao-Jun Yang, Dumitru Baleanu, and J. A. Tenreiro Machado
Volume 2013, Article ID 769724, 8 pages

The Stochastic Dynamics for Ecological Tourism System with Visitor Educational Intervention, Dongping Wei, Shouwen Wen, Yizeng Chen, Yong Xu, and Bing Liang
Volume 2013, Article ID 617608, 9 pages

Editorial

Advanced Topics in Dynamics of Complex Systems

Dumitru Baleanu,¹ Bashir Ahmad,² Juan J. Nieto,³ and J. A. Tenreiro Machado⁴

¹ Department of Mathematics and Computer Sciences, Cankaya University, Balgat, 06530 Ankara, Turkey

² Department of Mathematics, Faculty of Science, King Abdulaziz University, P.O. Box 80203, Jeddah 21589, Saudi Arabia

³ Institute of Mathematics, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain

⁴ Department of Electrical Engineering, Institute of Engineering, Polytechnic of Porto (ISEP), Rua Dr. Antonio Bernardino de Almeida, 431, 4200-072 Porto, Portugal

Correspondence should be addressed to Dumitru Baleanu; dumitru.baleanu@gmail.com

Received 23 February 2014; Accepted 23 February 2014; Published 23 March 2014

Copyright © 2014 Dumitru Baleanu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The investigation of complex systems in a unified manner is recognized during the last few years as a new and very promising scientific discipline. Therefore, a wide range of interdisciplinary methods and techniques should be applied and, at the same time, new tools should be invented to extract hidden and valuable information from the dynamics of complex systems.

The main aim of this special issue is to present algorithms and modelling techniques for analysing complex phenomena and dynamical systems with nonlinear and memory effects.

Fractional calculus, in view of its practical applications in several disciplines such as physics and engineering problems, is found to be capable of describing phenomena possessing long range memory effects that are difficult to handle with classical integer-order calculus. There has been a growing interest in the adaptation of fractional calculus as an effective modeling tool for complex systems, giving rise to new perspectives in their dynamical analysis and control. This development in the scientific knowledge is demonstrated by a large number of books and events emerging on the topic during the past years. However, many phenomena still pose considerable challenges to the concerned community and fractional calculus seems to be the probable candidate to encompass broader paradigms by shedding light into the description of complexity.

This special issue contains the papers addressing the recent theoretical advances and experimental results on the topics such as chaos, fractals, fractional dynamics, impulsive and neural systems, information and computation theory, signal processing (filtering, restoration, reconstruction, and

analysis of fractal noise), image encryption and hiding, mechanics, Navier-Stokes equations on Cantor sets, mechatronics (viscoelasticity, vibration insulation, and control), biology (orthodontic root resorption, classifications of some structural domains of proteins, and genetic algorithms), biophysics (signal and models of biological systems, control of language of competitive models, tourism systems, and viscoelasticity in biology), bioengineering, nanotechnology, anomalous diffusion, and complex adaptive systems.

The Guest Editors believe that the set of selected advanced studies offer readers a timely update of important research topics and may serve as a platform for motivating further contributions in this fast evolving arena.

Acknowledgments

The Guest Editors of this special issue would like to express their gratitude to the authors who have submitted papers for consideration. Thanks also to the many individuals who served as referees of the submitted papers. All the participants have made it possible to have a very stimulating interchange of ideas. We hope that these topics will stimulate future research. We would also like to thank the editorial board members of this journal for their support and help throughout the preparation of this special issue.

Dumitru Baleanu
Bashir Ahmad
Juan J. Nieto
J. A. Tenreiro Machado

Research Article

Bidirectional Dynamic Diversity Evolutionary Algorithm for Constrained Optimization

Weishang Gao,¹ Cheng Shao,^{1,2} and Yi An¹

¹ Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, Liaoning 116024, China

² Institute of Advanced Control Technology, Dalian University of Technology, Dalian, Liaoning 116024, China

Correspondence should be addressed to Weishang Gao; gao.weishang@hotmail.com

Received 1 September 2013; Accepted 13 November 2013

Academic Editor: D. Baleanu

Copyright © 2013 Weishang Gao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Evolutionary algorithms (EAs) were shown to be effective for complex constrained optimization problems. However, inflexible exploration-exploitation and improper penalty in EAs with penalty function would lead to losing the global optimum nearby or on the constrained boundary. To determine an appropriate penalty coefficient is also difficult in most studies. In this paper, we propose a bidirectional dynamic diversity evolutionary algorithm (Bi-DDEA) with multiagents guiding exploration-exploitation through local extrema to the global optimum in suitable steps. In Bi-DDEA potential advantage is detected by three kinds of agents. The scale and the density of agents will change dynamically according to the emerging of potential optimal area, which play an important role of flexible exploration-exploitation. Meanwhile, a novel double optimum estimation strategy with objective fitness and penalty fitness is suggested to compute, respectively, the dominance trend of agents in feasible region and forbidden region. This bidirectional evolving with multiagents can not only effectively avoid the problem of determining penalty coefficient but also quickly converge to the global optimum nearby or on the constrained boundary. By examining the rapidity and veracity of Bi-DDEA across benchmark functions, the proposed method is shown to be effective.

1. Introduction

A great deal of engineering design problems can be formulated as constrained nonconvex optimization problems which are difficult to be solved with classic mathematical theory. Since swarm intelligence (SI) [1, 2] was introduced by Wang and Beni [3] in 1989, many evolutionary algorithms (EAs) describing the collective behavior of decentralized, self-organized systems, natural or artificial, have provided robust evidence for dealing with incremental nonconvex or other complex optimization [4–6]. Many successful applications of EAs [7–9] have been reported to solve engineering problems such as industrial design [10–12], transportation [13], commerce [14], and bioinformation [15]. There are also many improvements [16, 17] on the original algorithms. To extend this application to constrained optimization, penalty functions are usually used to handle these multiple constraint [18, 19], which will transform the problems into unconstrained ones but meanwhile make original objective function more complex. Previous researches [20–22] have suggested that

evolutionary algorithms (EAs) can be widely used to tackle such problems. Many successful applications of EAs have been reported to solve engineering problems such as industrial design [23, 24] and military management [25]. To further expand the application of EAs to more difficult but important problems, Vural et al. carried out further study in analog filter design with evolutionary algorithms [26], Li and Yao presented a new cooperatively coevolving particle swarms for large scale optimization [27], Blackwell provided further study of collapse in bare bones particle swarm optimization [28], Pehlivanoglu enhanced particle swarm optimization with a periodic mutation strategy and neural networks [29], Chen et al. proposed particle swarm optimization with an aging leader and challengers [30], and Naznin et al. suggested a progressive alignment method using genetic algorithm for multiple sequence alignment [31]. Constrained optimization is an important kind of problems solved by EAs, such as the methods proposed by Wang et al. [32, 33], Cai and Wang [34], Krohling and Dos Santos Coelho [35], and Tessema and Yen [36]. Recently Daneshyari and Yen [37] have noted

that genetic-based algorithms and swarm-based paradigms are two popular population-based heuristics introduced as EAs for solving constrained optimization problems [38–40]. These algorithms have better global search abilities by sharing the principle of natural evolution or swarm intelligence in nature.

Genetic algorithm (GA) is a search heuristic which mimics the process of natural evolution by parallel computing. It became popular through the work of John Holland in the early 1970s, and particularly his book *Adaptation in Natural and Artificial Systems* (1975). The strong exploration of GA makes it fit for intricate optimization problems [41, 42], but GA's inborn disadvantages such as slow convergence caused by mutation operator [22] have limited the promotion of this algorithm in actual application. Furthermore, particle swarm optimization (PSO) developed by Kennedy and Eberhart [43, 44] has received more and more attention regarding its potential as a faster global optimization technique [22]. However, it might be caught in the trap of local optimization caused by premature convergence. Thus, special trade-off between exploration and exploitation is required to achieve proper balance between optimization reliability and convergence speed. A feasible method is to combine a variety of EAs then form a hybrid algorithm with strong exploration and exploitation. It is because mutually reinforcing will make hybrid algorithm fit for difficult optimization problems where acceptable solutions can be achieved [45–47]. Recently, coevolutionary algorithms (CoEAs) have been extensively studied in solving complex constrained optimization problems [20]. It can be considered as a new form of hybrid algorithm with high efficiency in exchanging information between agents. Although these algorithms can perform better than the standard EAs, inflexible exploration-exploitation and improper penalty in general EAs with penalty function would lead to losing the global optimum nearby or on the constrained boundary.

Usually the fitness function, especially around the constrained boundary where there is a global optimum, becomes much more complex because of the mixed constraints and penalty functions. Over-penalty on the forbidden agents which is nearby the global optimum will lose important samples with better guiding function. Under-penalty on the forbidden agents which is far away from the global optimum will weaken strength of converging toward the global optimum. Then inflexible exploration-exploitation will probably miss the global optimum and lead to ill-convergence. In this paper, we propose a bidirectional dynamic diversity evolutionary algorithm (Bi-DDEA) with multiagents guiding adaptive exploration-exploitation to the global optimum. In Bi-DDEA potential advantage is detected by three kinds of agents. The scale and the density of agents will change dynamically according to the emerging of potential optimal area, which play an important role of flexible exploration-exploitation. Meanwhile, a novel double optimum estimation strategy with objective fitness and penalty fitness is suggested to compute, respectively, the dominance trend of agents in feasible region and forbidden region. Penalty fitness will guide agents in forbidden region to nearby constrained boundary with optimal objective fitness, and objective fitness will guide agents in

feasible region to global optimum. This bidirectional evolving with multiagents can not only effectively avoid the problem of determining penalty coefficient but also quickly converge to the global optimum nearby or on the constrained boundary. Unlike other penalty methods, the penalty fitness suggested in this paper is not added to the objective function as a penalty term but constitutes the Bi-DDEA as a evaluation term to select optimal forbidden agents independently, which will make a convergence towards the global optimum from another direction in the forbidden region.

This paper is organized as follows. Section 2.1 presents the key elements of bidirectional information processing. Section 2.2 explains how to carry out adaptive exploration-exploitation with multiagents. Section 2.3 describes the outline of Bi-DDEA. Section 3 provides the numerical experimentation with benchmark problems and comparisons with other swarm optimization algorithms. A detailed analysis and conclusion of the algorithm are presented in Section 4.

2. Bidirectional Dynamic Diversity Evolutionary Framework

2.1. Bidirectional Information Processing. Although penalty function method is one of the most common methods to solve constrained optimization problems, the design of penalty item should be fit for the algorithm with which we will deal with practical optimization problems. In order to avoid the complication caused by adding penalty term to the objective function, objective fitness (F) and penalty fitness (P) are suggested in this paper to deal with both the agents distributed in feasible region and forbidden region respectively. The objective fitness of all the agents is computed according to the objective function. When an agent emerges in the feasible region, the penalty fitness will be set to zero. When an agent emerges in the forbidden region, the penalty fitness will be computed according to the constraint conditions and sum all the excess value as the penalty fitness.

The general constrained problem formulation that is also called the primal problem can be stated as follows:

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ \text{s.t.} & g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m \\ & h_i(\mathbf{x}) = 0, \quad i = m + 1, \dots, m + l \\ & \mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbf{R}^n, \end{array} \quad (1)$$

where \mathbf{x} represents a vector of n real variables subject to a set of m inequality constraints $g(\mathbf{x})$ and a set of l equality constraints $h(\mathbf{x})$. The penalty fitness $p(\mathbf{x})$ proposed here can be defined as

$$p(\mathbf{x}) = \sum_{i=1}^m \phi(g_i(\mathbf{x})) + \sum_{i=m+1}^{m+l} \varphi(h_i(\mathbf{x})), \quad (2)$$

where

$$\phi(x) = \begin{cases} x, & \text{if } x > 0, \\ 0, & \text{if } x \leq 0, \end{cases}$$

$$\varphi(x) = \begin{cases} x, & \text{if } x > 0, \\ 0, & \text{if } x = 0, \\ -x, & \text{if } x < 0. \end{cases} \quad (3)$$

Penalty fitness $p(\mathbf{x})$ will be used in Bi-DDEA to judge whether an agent is in the feasible region. Once an agent is updated by Bi-DDEA, $p(\mathbf{x})$ will be called to check the decision variable represented by this agent. If the penalty fitness is zero, a feasible decision variable is gotten. Otherwise, the penalty fitness will replace the objective fitness and be regarded as the fitness of the forbidden agent. The fitness function $\hat{f}(\mathbf{x})$ can be expressed as what is shown in (4):

$$\hat{f}(\mathbf{x}) = \begin{cases} F, & \text{if } p(\mathbf{x}) = 0, \\ P, & \text{if } p(\mathbf{x}) > 0, \end{cases} \quad (4)$$

$$= \begin{cases} f(\mathbf{x}), & \text{if } p(\mathbf{x}) = 0, \\ p(\mathbf{x}), & \text{if } p(\mathbf{x}) > 0, \end{cases} \quad (5)$$

where $F = f(\mathbf{x})$ and $P = p(\mathbf{x})$. As the fitness function is divided into two departments to deal with the feasible agents and the forbidden agents, respectively, bidirectional evolution can be carried out on these two kinds of agents. Thus, the fitness growth (G) of an agent can be defined as

$$G(t+1) = \begin{cases} F(t+1) - F(t), & \text{if } P(t) = 0, \\ & P(t+1) = 0, \\ F(t+1) - F_{\min}, & \text{if } P(t) > 0, \\ & P(t+1) = 0, \\ P(t+1) - 0, & \text{if } P(t) = 0, \\ & P(t+1) > 0, \\ P(t+1) - P(t), & \text{if } P(t) > 0, \\ & P(t+1) > 0, \end{cases} \quad (6)$$

where t refers to the iteration number and F_{\min} refers to the minimum fitness of all the feasible agents.

\widehat{AD} , \widehat{RD} , and \widehat{DC} are introduced in this paper as potential superiority factors to analyze the sampling information and estimate the situation of both feasible agents and forbidden agents, which will guide bidirectional evolution by distributing newborn agents with adaptive density. The potential superiority factors can be defined as below:

$$\widehat{AD} = \begin{cases} \exp\left(\frac{F - (F_{\max} + F_{\min})/2}{(F_{\max} - F_{\min})/2}\right), & \text{if Penalty} = 0, \\ \exp\left(-\frac{P - (P_{\max} + P_{\min})/2}{(P_{\max} - P_{\min})/2}\right), & \text{if Penalty} > 0, \end{cases} \quad (7)$$

$$= \begin{cases} \exp\left(\frac{2F - F_{\max} - F_{\min}}{F_{\max} - F_{\min}}\right), & \text{if Penalty} = 0, \\ \exp\left(-\frac{2P - P_{\max} - P_{\min}}{P_{\max} - P_{\min}}\right), & \text{if Penalty} > 0, \end{cases} \quad (8)$$

$$\widehat{RD} = \begin{cases} \exp\left(\frac{G - (\widetilde{G}_{\max} + \widetilde{G}_{\min})/2}{(\widetilde{G}_{\max} - \widetilde{G}_{\min})/2}\right), & \text{if Penalty} = 0, \\ \exp\left(-\frac{G - (\widetilde{G}_{\max} + \widetilde{G}_{\min})/2}{(\widetilde{G}_{\max} - \widetilde{G}_{\min})/2}\right), & \text{if Penalty} > 0, \end{cases} \quad (9)$$

$$= \begin{cases} \exp\left(\frac{2G - \widetilde{G}_{\max} - \widetilde{G}_{\min}}{\widetilde{G}_{\max} - \widetilde{G}_{\min}}\right), & \text{if Penalty} = 0, \\ \exp\left(-\frac{2G - \widetilde{G}_{\max} - \widetilde{G}_{\min}}{\widetilde{G}_{\max} - \widetilde{G}_{\min}}\right), & \text{if Penalty} > 0, \end{cases} \quad (10)$$

$$\widehat{DC} = (\widehat{RE}) \times \widehat{RD} + \widehat{RE} \times \widehat{AD}, \quad (11)$$

where F_{\max} and F_{\min} refer to the minimum and the maximum objective fitness in feasible region, P_{\max} and P_{\min} refer to the minimum and the maximum penalty fitness in forbidden region, \widetilde{G}_{\max} refers to the maximum fitness growth of feasible agents in previous generations, \widetilde{G}_{\min} refers to the minimum fitness growth of feasible agents in current generation, \widetilde{G}_{\max} refers to the maximum fitness growth of forbidden agents in current generation, \widetilde{G}_{\min} refers to the minimum fitness growth of forbidden agents in previous generations, and \widehat{RE} is a reliability factor which is defined as below:

$$\widehat{RE} = \frac{t}{2T} + \frac{D-d}{2D}, \quad (12)$$

where \widehat{RE} will increase gradually from zero to one in the process of optimization.

2.2. Dynamic Diversity Evolution. Three types of agents, partition agents (PAs), basic agents (BAs), and creative agents (CAs), are combined together in Bi-DDEA to carry out adaptive exploration-exploitation according to the guiding of \widehat{AD} , \widehat{RD} , and \widehat{DC} . Partition agents are distributed in feasible region uniformly and increase gradually, which plays a role of brute search and ensures exploration throughout the course of optimization. Basic agents are distributed in different partitions according to the property of partition agents, which plays a role of transition media between partition agents and creative agents and also between exploration and exploitation. Creative agents are distributed around basic agents according to the property of basic agents, playing a role of exploitation in global area and also exploration in local partitions.

The evolutionary process of Bi-DDEA is mainly carried out by the rebirth of new agents around the senior ones. one PA will generate some BAs in a local region, and one BA will generate some CAs in a local region around the BA. Then some CAs with higher potential superiority will be selected as the next BAs. BAs alternate with CAs, which is also affected by PAs. The focus in Bi-DDEA is no longer on the position updating of previous agents but on the density distribution of newborn agents. The density of newborn agents is controlled by regulating the range and the scale of newborn agents according to the feedback of sampling information such as \widehat{AD} , \widehat{RD} , and \widehat{DC} .

The range of newborn agents is determined by two items in Bi-DDEA. The first one is narrowing range with 2/3 ratio, which means that the child agents will inherit $(2/3)(r_{\text{parent}})$ from their parent. The second one is regulated range with \widehat{AD} , which means that the child agents will study from the sampling information and adjust their range toward $(2/3)(r_{\text{parent}}/\widehat{AD})$. Then, the range of those agents of which the fitness is lower will be set to a larger value and vice versa. To make information fusion, α is introduced as study factor to integrate the two items in the form of $(1 - \alpha) \times \text{item}_1 + \alpha \times \text{item}_2$. It can be described specifically as follows:

$$r_{i,j}^B(t+1) = \frac{2}{3} \times \left[(1 - \alpha) + \frac{\alpha}{\widehat{AD}_{i,j}^B} \right] \times r_i^P(t), \quad (13)$$

$$\tilde{r}_{i,j,k}^C(t+1) = \frac{2}{3} \times \left[(1 - \alpha) + \frac{\alpha}{\widehat{AD}_{i,j,k}^C} \right] \times r_{i,j}^B(t), \quad (14)$$

where $r_i^P(t)$, $r_{i,j}^B(t)$, and $\tilde{r}_{i,j,k}^C(t+1)$ refer to the range of PA_i, BA_{i,j}, and CA_{i,j,k} in each iteration. $\widehat{AD}_{i,j}^B$ and $\widehat{AD}_{i,j,k}^C$ refer to the correction factor \widehat{AD} of BA_{i,j} and CA_{i,j,k}. The indexes, i, j, and k, indicate that the current parameter attaches to PA_i, BA_j, and CA_k.

When newborn CAs are distributed, the sampling information will also affect the property, range, of PAs and the changing of Pas' property will also affect the property, range, of CAs'. This process is called self-correcting in range which can be described as follows:

$$r_i^P(t+1) = (1 - \alpha) \times r_i^P(t) + \alpha \times \frac{r_i^P(t)}{\widehat{AD}_i^P}, \quad (15)$$

$$\tilde{r}_{i,j,k}^C(t+1) = \left[(1 - \alpha) + \frac{\alpha}{\widehat{AD}_i^P} \right] \times \tilde{r}_{i,j,k}^C(t+1), \quad (16)$$

where $\tilde{r}_{i,j,k}^C(t+1)$ and $r_{i,j,k}^C(t+1)$ refer to the range of CA_{i,j,k} before and after the self-correction. The equation set (13)–(16) is also called the range-dividing principle in Bi-DDEA. Larger range caused by smaller \widehat{AD} will make weak agents explore outward to discover more dominant position. On the contrary, smaller range caused by larger \widehat{AD} will converge agents to exploit the emergent region fastly.

The scale of newborn agents is also determined by two items. The first one is the scale which is the same with their parent, which means the child agents will inherit s_{parent} from their parent. The second one is regulated scale with \widehat{RD} or \widehat{DC} , which means the child agents will study from the sampling information and adjust their scale toward $\widehat{RD} \times r_{\text{parent}}$ or $\widehat{DC} \times r_{\text{parent}}$. Then, the scale of those agents of which the fitness growth is larger will be set to a larger value and vice versa. To make information fusion, α is also introduced as study factor to integrate the two items in the form of $(1 - \alpha) \times \text{item}_1 + \alpha \times \text{item}_2$. It can be described specifically as follows:

$$s_{i,j}^B(t+1) = \left[(1 - \alpha) + \alpha \times \widehat{RD}_{i,j}^B \right] \times s_i^P(t), \quad (17)$$

$$\tilde{s}_{i,j,k}^C(t+1) = \left[(1 - \alpha) + \alpha \times \widehat{RD}_{i,j,k}^C \right] \times s_{i,j}^B(t), \quad (18)$$

where $s_i^P(t)$, $s_{i,j}^B(t)$, and $\tilde{s}_{i,j,k}^C(t+1)$ refer to the scale of PA_i, BA_{i,j}, and CA_{i,j,k} in each iteration. $\widehat{RD}_{i,j}^B$ and $\widehat{RD}_{i,j,k}^C$ refer to the correction factor \widehat{RD} of BA_{i,j} and CA_{i,j,k}. The indexes, i, j, and k, indicate that the current parameter attaches to PA_i, BA_j, and CA_k.

When newborn CAs are distributed, the sampling information will also affect the property, scale, of PAs and the changing of Pas' property will also affect the property, scale, of CAs'. This process is called self-correcting in scale which can be described as follows:

$$s_i^P(t+1) = \left[(1 - \alpha) + \alpha \times \widehat{DC}_i^P \right] \times s_i^P(t), \quad (19)$$

$$s_{i,j,k}^C(t+1) = \left[(1 - \alpha) + \alpha \times \widehat{DC}_i^P \right] \times \tilde{s}_{i,j,k}^C(t+1),$$

where $\tilde{s}_{i,j,k}^C(t+1)$ and $s_{i,j,k}^C(t+1)$ refer to the scale of CA_{i,j,k} before and after the self-correction. The reason to select \widehat{DC}_i^P as a self-correcting factor is that the focus of \widehat{DC} will transition from \widehat{RD} to \widehat{AD} and large number of agents should be used to exploit the area where there always is the optimal agent in the later phase of optimization. While the algorithm runs, the focus of larger scale agents will translate from complex area to dominance area. Thus, Bi-DDEA can carry out specific exploration at the earlier stage of optimization and develop fast targeted exploitation at the later phase of optimization. In addition to this, a swarm scale floor (SSF) is introduced to reduce quickly the number of agents in the area where there is little growth in fitness. It not only saves sample resource but also maintains global exploration with fewer agents, which can be described as follows:

$$s(t+1) = \widehat{SSF} \quad (\text{if } \widehat{RD} \leq 0). \quad (20)$$

Swarm scale ceil (SSC) is also introduced to prevent agents from excessive generating. Considering the forbidden agents will not be the global optimal one, the scale floor of forbidden agents will not be confined by SSF but limited to

one or more. Thus, the scale of agents will be corrected in the following form:

$$s = \begin{cases} \widehat{\text{SSF}}, & (\text{if } s < \widehat{\text{SSF}} \text{ and Penalty} = 0); \\ 1, & (\text{if } s < 1 \text{ and Penalty} > 0); \\ \widehat{\text{SSC}}, & (\text{if } s > \widehat{\text{SSC}}). \end{cases} \quad (21)$$

The equation set (17)–(21) is also called the scale-setting principle in Bi-DDEA. Larger scale caused by larger \widehat{RD} or \widehat{DC} will not only reduce the loss of information in volatile region but also strengthen the ability of both exploration and exploitation towards potential advantage areas. On the contrary, smaller scale caused by smaller \widehat{RD} or \widehat{DC} will save much time that might be wasted in poor areas.

2.3. The Proposed Bi-DDEA. Bi-DDEA mainly consists of bidirectional information processing suggested in Section 2.1 and dynamic diversity evolution suggested in Section 2.2. The algorithm flow is described as follow:

- (1) to divide searching space into a number of partitions with partition agents (PAs) and initialize the properties of each PA;
- (2) to distribute basic agents (BAs) in each partition according to the properties of PA;
- (3) to divide BAs into two types with the constrained conditions and, respectively, compute their objective fitness, penalty fitness, and fitness growth according to formulas (4), (2), and (6);
- (4) to estimate the potential advantage of BAs both in feasible region and forbidden region according to formulas (8), (10), and (11);
- (5) to assign the attributes of BAs according to formulas (13) and (17);
- (6) to distribute creative agents (CAs) around each BA according to the range and scale of the BA;
- (7) to divide CAs into two types with the constrained conditions and, respectively, compute their objective fitness, penalty fitness, and fitness growth according to formulas (4), (2), and (6);
- (8) to estimate the potential advantage of CAs both in feasible region and forbidden region according to formulas (8), (10), and (11);
- (9) to assign the attributes of CAs and PAs according to formulas (14), (15), (18), and (19);
- (10) to correct the scale of each agents according to formulas (20) and (21);
- (11) to select better CAs as next BAs according to the scale of PA in each partition;
- (12) to judge whether the breaking conditions are met. If so, then it continue to next step; otherwise it jumps to step (3);
- (13) to select the best BA as present global optimal agents and translate it into a PA;

(14) to judge whether the breaking conditions are met. If so, then it continue to next step; otherwise it jumps to step (1);

(15) output the results.

Obviously, all the agents that are divided into two types in Bi-DDEA will evolve along two directions with the dynamic diversity evolution method under the guidance of bidirectional information analysis proposed in this paper.

3. Benchmark Problems

The performance of Bi-DDEA is evaluated by comparing its rapidity, accuracy, and universality with those of PSO algorithm and GA. To be meaningful, some benchmark functions that are commonly used in randomized population-based algorithms testing are selected as objective functions. This gives a fair comparison being as far as possible free from biases that favor one style of algorithm over another. These functions include Rastrigin [48–53], Rosenbrock [48–53], Griewank [49–52], Needle-in-a-Haystack [54–56], Shubert [48, 55, 57, 58], and Schaffer [49, 50, 54, 55, 57], as shown in Figure 1. As one of the most difficult but important issues in designing optimization algorithms is managing the balance between exploration and exploitation, these test functions of which the peak number and the peak distribution are different will really challenge both the exploration and the exploitation behavior of the algorithms. As Bi-DDEA is used to find the maximum value, we select the opposite value of objective function as fitness in the following evaluation process. In addition, the benchmark problems will be modified into constrained ones with constrained condition as shown in the following formula:

$$D \times 20^2 - \sum_{d=1}^D (x_d - 20)^2 \leq 0, \quad (22)$$

where D refers to the dimension of decision variable.

For GA, the initial parameters are set to their default values as described below:

- (1) population size: 100;
- (2) probability of mutation: 0.05.

For PSO, the initial parameters are set to their default values as described below:

- (1) population size: 40;
- (2) inertia weight (ω): 1, then decreases from 1 to 0.1 during the iteration progress.

Dynamic equation as follows:

$$\begin{aligned} \vec{v}_{i+1} &= \omega \vec{v}_i + \vec{U}(0, \phi_1) \otimes (\vec{p}_i - \vec{x}_i) \\ &\quad + \vec{U}(0, \phi_2) \otimes (\vec{p}_g - \vec{x}_i), \\ \vec{x}_{i+1} &= \vec{x}_i + \vec{v}_{i+1}. \end{aligned} \quad (23)$$

For Bi-DDEA, the initial parameters are set to their default values as described below:

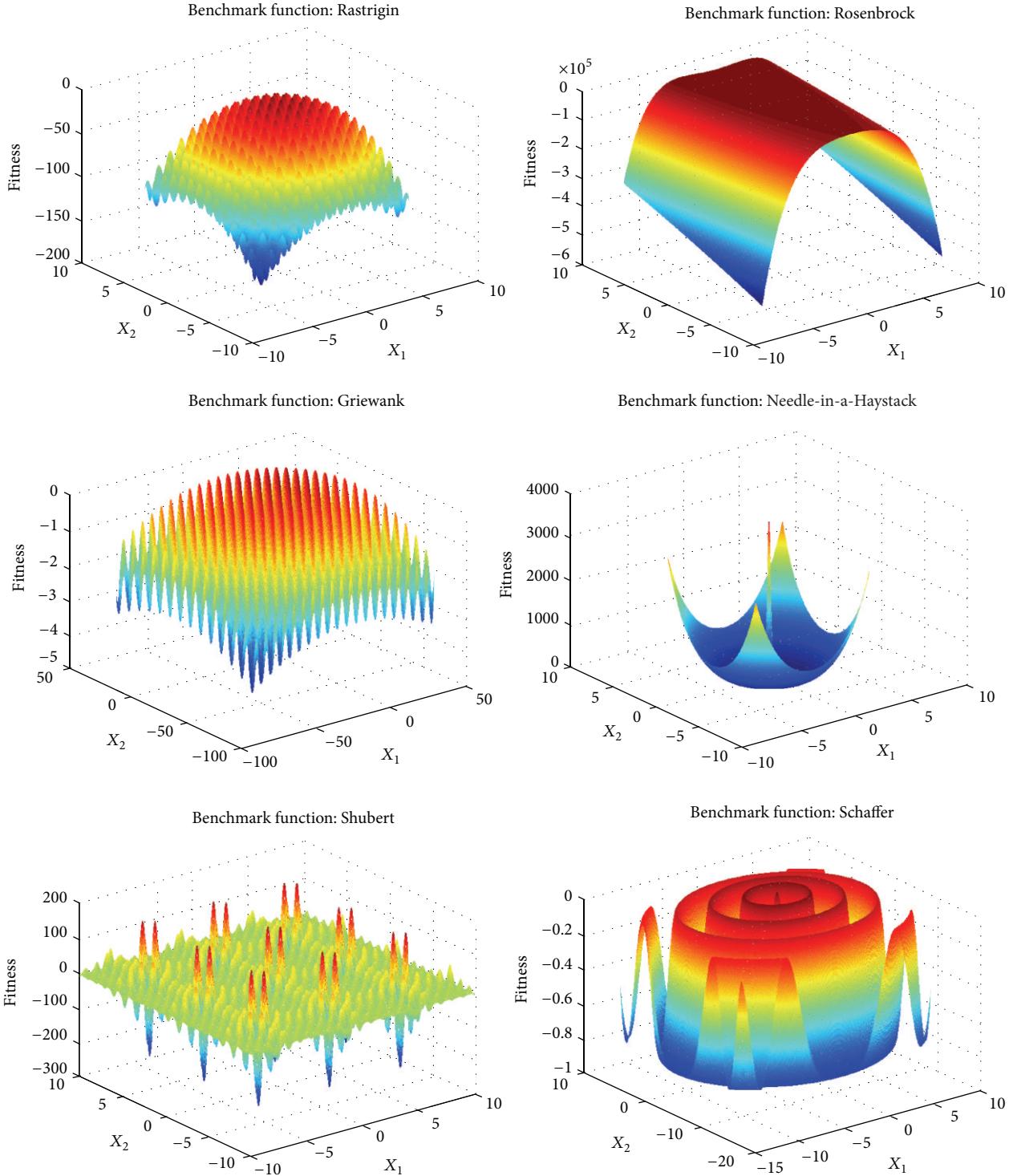


FIGURE 1: Benchmark functions.

- (1) partition dividing: from 5×5 to 10×10 ;
- (2) learning factor (α): 0.5;
- (3) SSF: D ;
- (4) SSC: $D \times 5$;

- (5) outside loop number: from 3 to 5;
- (6) inner loop number: from 5 to 8.

Dynamic equations are given as what formula (13)–formula (21) shown to compute r_i^P , $r_{i,j}^B$, $r_{i,j,k}^C$, s_i^P , $s_{i,j}^B$, and $s_{i,j,k}^C$ in Section 2.2. Although the dynamic equations of Bi-DDEA are

more complex than those of PSO, it is time-saving because it can determine the next distribution of many sampling points once.

3.1. Rapidity. This section compares the rapidity of Bi-DDEA with the results obtained by GA (genetic algorithm) and PSO (Particle swarm optimization) when they are applied to the six benchmark functions mentioned above. Sampling time is used to represent the optimization time logically. It refers to the cumulative time spent by every sample to calculate their fitness according to the coordinates. It can eliminate the program editor, environment, and other external factors affecting the evaluation of the optimization speed, which allows us to compare the rate of these algorithms more explicitly. An algorithm with different program design or in different computer environments will have a different performance. For example, a good programmer can implement an algorithm process with streamlined code to improve operational efficiency, but it is difficult to reflect its advantages when running in CPU with low frequency. In the experimental comparison, we would probably get wrong result due to these external factors. So, we use sampling time as reference to analyze their speed characteristics of the most original. Since the time used for each sample is almost equal, we use the number of samples to replace sampling time in this paper.

Figures 2, 3, 4, 5, 6, and 7 show some of the algorithms' qualified run-length distributions (RLDs, for short) based on sampling time when solving some benchmark functions. The solid line, dashed line, dash-dot line, and dotted line represent, respectively, the emerging of Bi-DDEA, LP-DDEA, PSO, and GA. On the whole, the optimization speeds of both PSO and Bi-DDEA are similar to each other and higher than that of GA. LP-DDEA is better sometimes than PSO and Bi-DDEA across functions Rosenbrock and Shubert, but slower than Bi-DDEA across the other problems. From this point of view, the algorithm designed in this paper is quite advantageous. Although PSO's convergence is mostly faster than that of Bi-DDEA in the early stage sometimes, the smallest first hitting times across different benchmark functions are mainly obtained by Bi-DDEA. It is because Bi-DDEA can dynamically give divers and sufficient exploration to obtain more accurate evolution in two directions. When a more fit area is found by exploration, fast exploitation will be carried out rapidly in this area. But the exploitation will not take up all the system resources. There are still many agents continuing to explore in both feasible region and forbidden region, and more and more exploitation agents will be changed into exploration agents quickly if exploitation can not find much more fit results. Thus, more and more samples will emerge at the constrained boundary and better local optimal area which probably include the global optimum. When the global optimum locates at the constrained boundary, the convergence speed will be much more quick. There is no algorithm comparison without exception. PSO can show an amazing speed when some particles drop into a better space in the experiments. Generally speaking, exploration and exploitation in Bi-DDEA are almost synchronic and exploration in Bi-DDEA is more intensive than that of PSO.

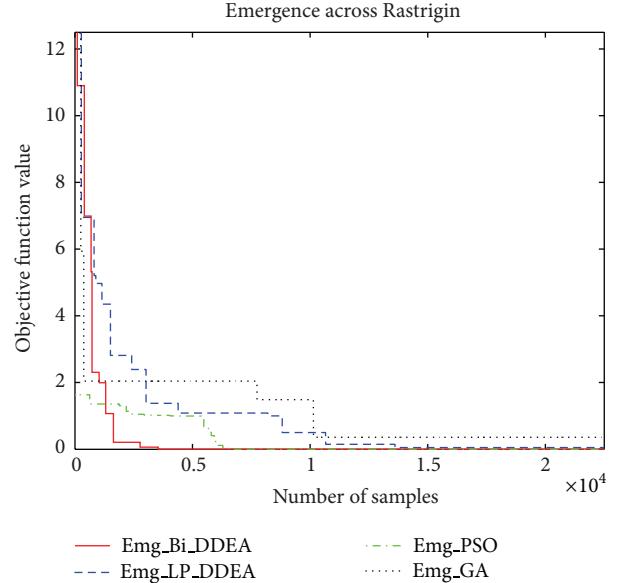


FIGURE 2: Comparison of speed across Rastrigin.

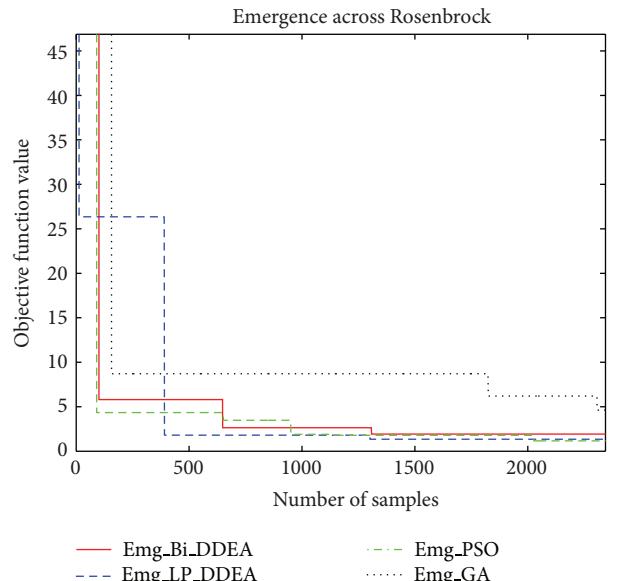


FIGURE 3: Comparison of speed across Rosenbrock.

So Bi-DDEA can come across global optimal result earlier in these experiments.

3.2. Veracity. High accuracy and global superiority are the key indicators of good algorithm in quality evaluation. From Tables 1, 2, 3, 4, 5, and 6 we also show that some optimal fitness get by four algorithms across six benchmark functions. Obviously, we can see from the results that Bi-DDEA and PL-DDEA find better fitness than what GA found, and PSO is sometimes trapped in local optimum. But on the contrary, PSO can also find more precise results than DDEA if PSO is not trapped in local optimal area. In fact, PSO and DDEA both have higher accuracy because of the rapid convergence

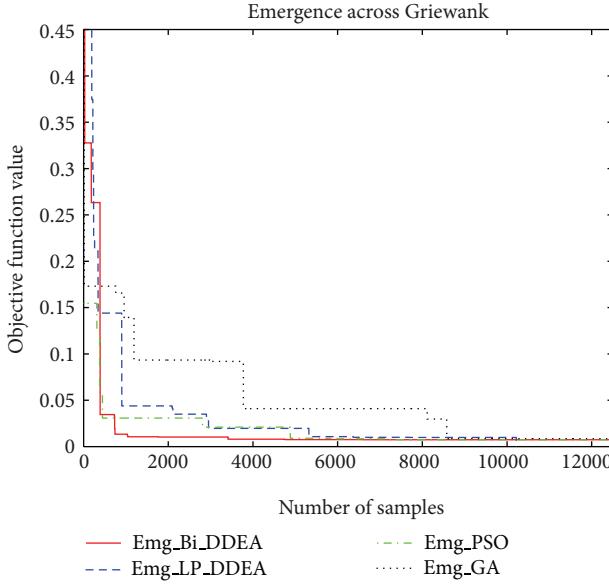


FIGURE 4: Comparison of speed across Griewank.

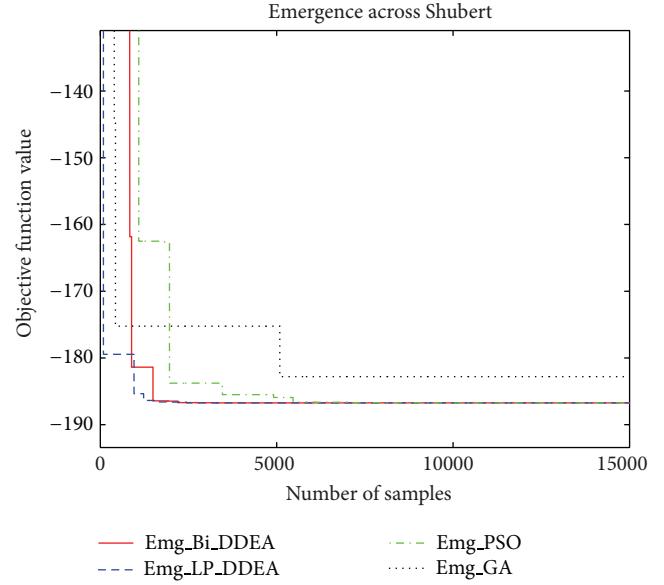


FIGURE 6: Comparison of speed across Shubert.

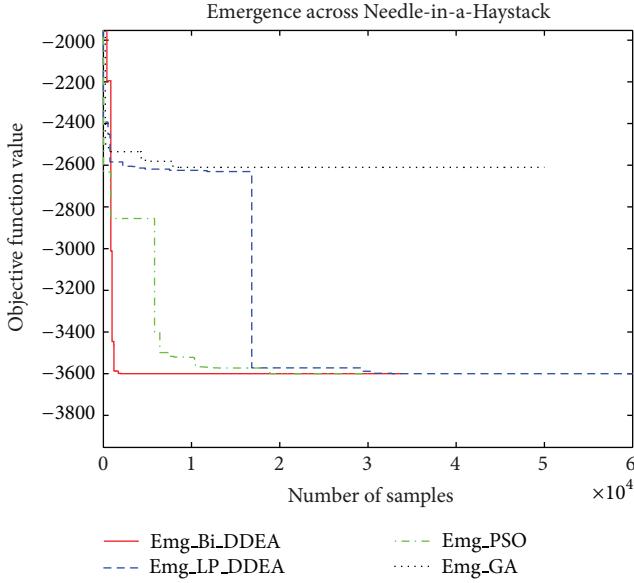


FIGURE 5: Comparison of speed across Needle-in-a-Haystack.

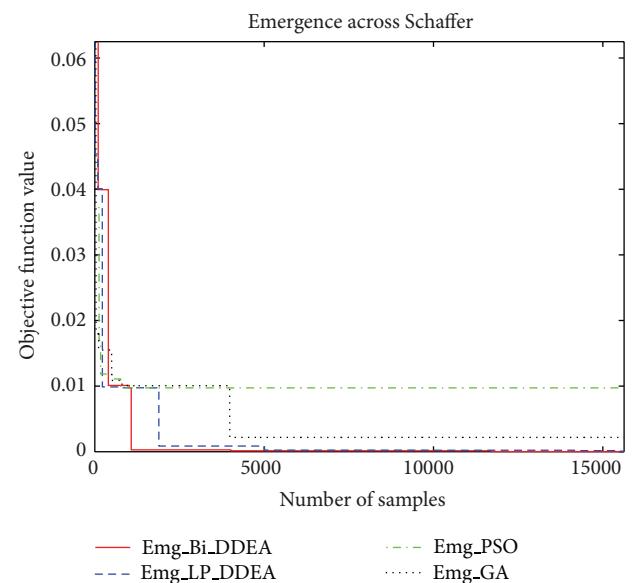


FIGURE 7: Comparison of speed across Schaffer.

of sample swarm, an exploitative behavior, in optimal area. The reason DDEA can mostly find that the minimum fitness is that DDEA's global search capability is stronger than that of PSO, and PSO sometimes did not avoid getting into local minimum across Schaffer function, Needle function, and Griewank function in these experiments.

3.3. Samples Distribution. To avoid falling into local minimum and find better optimum or constrained boundary, a good algorithm should have the ability to explore the unknown feasible region. As a better cases analysis, we now consider the distribution of all the samples through the whole

optimization process, as shown in Figure 8. Sample distribution shown in the figure of the four columns is produced, respectively, by GA, PSO, and LP-DDEA and Bi-DDEA from left to right. The figures of sample distribution in a row are produced by these algorithms across the same benchmark function, and the benchmark functions used in each row are different from each other. These benchmark functions from top to bottom are Rastrigin, Rosenbrock, Griewank, Needle-in-a-Haystack, Shubert, and Schaffer. From these figures, it can be intuitively found that GA has the strongest global exploring ability, followed by LP-DDEA and Bi-DDEA, and finally PSO. This is why GA's performance is better than

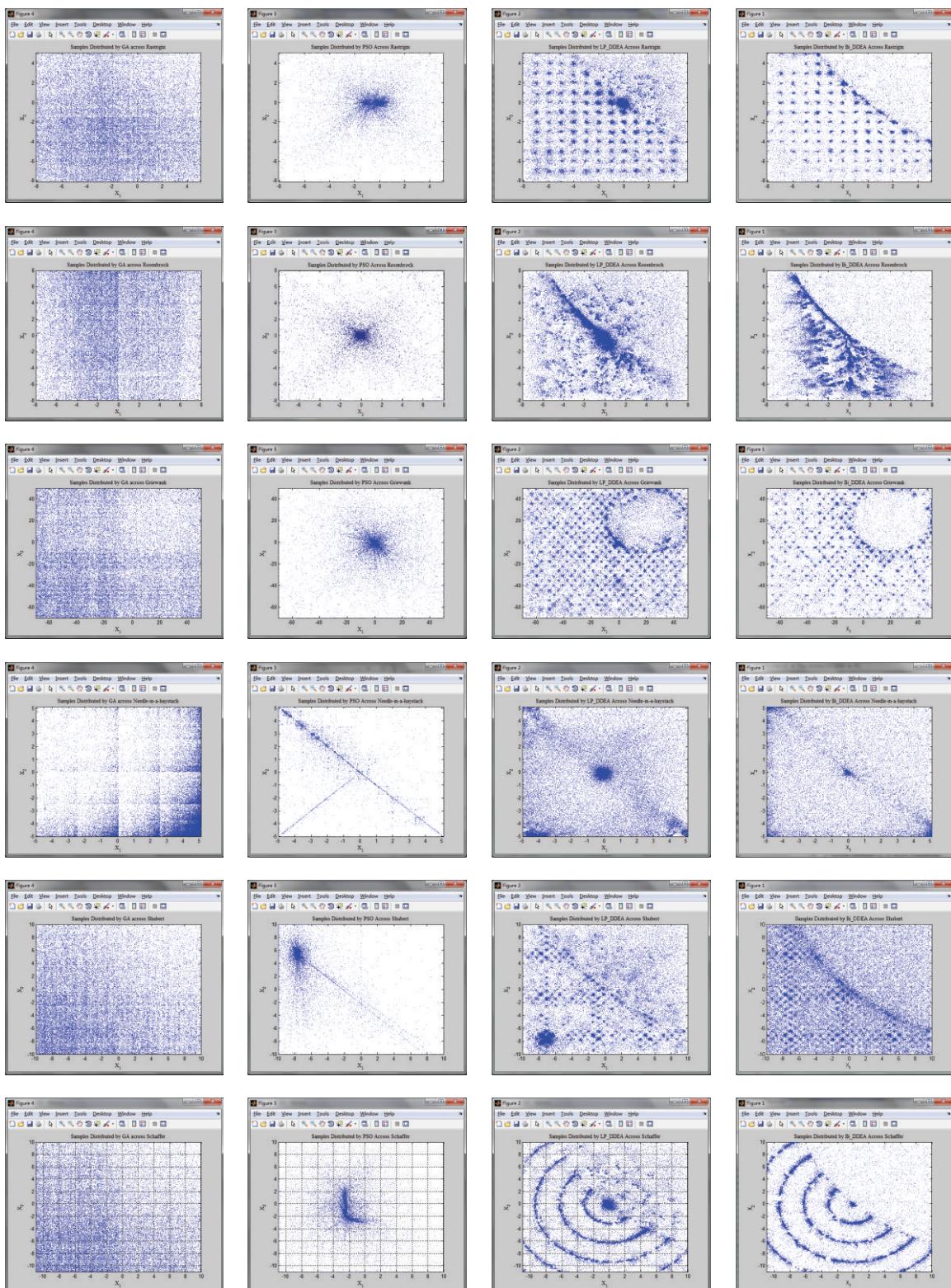


FIGURE 8: Samples distribution across benchmark functions.

TABLE 1: Comparison of optimal results across Rastrigin.

Function	Times	GA	PSO	PL-DDEA	Bi-DDEA
Rastrigin	1	0.1833	0	0.0000	0.0000
	2	0.0093	0	0.0000	0.0000
	3	0.0136	0	0.0000	0.0000
	4	0.0216	0	0.0000	0.0000
	5	0.0411	0	0.0000	0.0000
	6	0.0008	0	0.0000	0.0000
	7	0.0025	0	0.0000	0.0000
	8	0.0747	0	0.0000	0.0000
	9	0.0450	0	0.0000	0.0000
	10	0.0394	0	0.0000	0.0000
Mean		0.0431	0	0.0000	0.0000

TABLE 2: Comparison of optimal results across Rosenbrock.

Function	Times	GA	PSO	PL-DDEA	Bi-DDEA
Rosenbrock	1	1.1084	0.9903	0.9903	0.9903
	2	1.1828	0.9903	0.9903	0.9903
	3	1.1587	0.9903	0.9904	0.9903
	4	1.1089	0.9903	0.9903	0.9903
	5	1.1362	0.9903	0.9903	0.9903
	6	1.1643	0.9903	0.9903	0.9903
	7	1.0680	0.9903	0.9903	0.9903
	8	1.1026	0.9903	0.9903	0.9903
	9	1.1626	0.9903	0.9903	0.9903
	10	1.1017	0.9903	0.9903	0.9903
Mean		1.0219	0.9903	0.9903	0.9903

TABLE 3: Comparison of optimal results across Griewank.

Function	Times	GA	PSO	PL-DDEA	Bi-DDEA
Griewank	1	0.0024	0	0	0
	2	0.0000	0	0	0
	3	0.0076	0.0074	0	0
	4	0.0107	0.0074	0	0
	5	0.0138	0.0099	0	0
	6	0.0094	0.0074	0	0
	7	0.0099	0.0074	0	0
	8	0.0055	0.0099	0	0
	9	0.0008	0.0074	0	0
	10	0.0096	0.0074	0	0
Mean		0.0070	0.0064	0	0

PSO when searching the extreme point in function Needle-in-a-Haystack. But it also takes GA a long time to explore everywhere without any convergence exploitation in feasible region, which not only reduces optimization speed but also affects optimization accuracy to some extent.

In addition, Figure 8 shows that Bi-DDEA and LP-DDEA both have the ability of finding the constrained boundary where there is the global optimum in all probability. But Bi-DDEA can find a more clear boundary with more effective sample distribution. It is because bidirection evolution

TABLE 4: Comparison of optimal results across Needle-in-a-haystack.

Function	Times	GA	PSO	PL-DDEA	Bi-DDEA
Needle /(1.0e + 003)	1	-3.4857	-2.7488	-3.6000	-3.6000
	2	-3.4282	-3.6000	-3.6000	-3.6000
	3	-3.5981	-3.6000	-3.6000	-3.6000
	4	-3.5342	-3.6000	-3.6000	-3.6000
	5	-3.5363	-3.6000	-3.6000	-3.6000
	6	-3.4579	-2.7488	-3.6000	-3.6000
	7	-3.5867	-2.7488	-3.6000	-3.6000
	8	-3.5637	-3.6000	-3.6000	-3.6000
	9	-3.5728	-3.6000	-3.6000	-3.6000
	10	-3.5824	-2.7488	-3.6000	-3.6000
Mean		-3.5346	-3.2595	-3.6000	-3.6000

TABLE 5: Comparison of optimal results across Shubert.

Function	Times	GA	PSO	LP-DDEA	Bi-DDEA
Shubert	1	-186.6995	-186.7309	-186.7309	-186.7309
	2	-186.6186	-186.7309	-186.7309	-186.7309
	3	-186.6912	-186.7309	-186.7309	-186.7309
	4	-186.6519	-186.7309	-186.7309	-186.7309
	5	-186.6551	-186.7309	-186.7309	-186.7309
	6	-186.7122	-186.7309	-186.7309	-186.7309
	7	-186.7244	-186.7309	-186.7308	-186.7309
	8	-186.0812	-186.7309	-186.7309	-186.7309
	9	-186.2785	-186.7309	-186.7309	-186.7309
	10	-186.4776	-186.7309	-186.7309	-186.7309
Mean		-186.5590	-186.7309	-186.7309	-186.7309

TABLE 6: Comparison of optimal results across Schaffer.

Function	Times	GA	PSO	LP-DDEA	Bi-DDEA
Schaffer	1	0.0051	0.0000	0.0000	0.0000
	2	0.0011	0	0.0000	0.0000
	3	0.0030	0	0.0000	0.0000
	4	0.0027	0.0000	0.0000	0.0000
	5	0.0020	0.0000	0.0000	0.0000
	6	0.0006	0.0097	0.0000	0.0000
	7	0.0086	0.0097	0.0000	0.0000
	8	0.0001	0.0000	0.0000	0.0000
	9	0.0070	0.0000	0.0000	0.0000
	10	0.0024	0.0000	0.0000	0.0000
Mean		0.0033	0.0019	0.0000	0.0000

suggested in Bi-DDEA carries out special convergence from forbidden region to better constrained boundary. As the global optimum always on the constrained boundary in most constrained optimization, Bi-DDEA is faster than the LP-DDEA when dealing with constrained problems in engineering application as shown in Section 3.1.

4. Analysis and Conclusion

Experimental results have shown notably that Bi-DDEA with bidirection evolutions is more successful in solving non-convex constrained optimization problems. The wonderful performance profits from the bidirection information processing and the dynamic diversity evolution carried out by multi-agents consisting of PAs, Bas, and CAs both in feasible region and forbidden region. Adaptive exploration-exploitation in feasible region makes Bi-DDEA find out most of the local optimum which include the global optimum. It indicates that the suggested algorithm has the ability of discovering the character of function distribution, which provides robust evidence for the global searching ability of Bi-DDEA. Another evolution direction from the forbidden region is detected specially by agents which are distributed in infeasible region inevitably. The mechanism of Bi-DDEA does not penalize these infeasible agents with large penalty to eliminate their guidance, but to make full use of their guidance to the constrained boundary. When a feasible agent is generated by a forbidden agent, it will get much higher fitness growth according to its present fitness. Thus, more and more samples will be distributed on the boundary with higher fitness as shown in Figure 8.

Bi-DDEA is not sensitive to initial population in most cases. From the mechanism analysis of Bi-DDEA, we can find that sample size in each iteration is regulated according to the feedback sampling information. Even though the setting population is different from the need greatly in the initial phase, Bi-DDEA will adjust the population scale automatically to a needed degree through a few iterations. It is worth noting that some special kinds initial partitions are able to improve efficiency of optimization algorithm, and these kinds of initial partitions are called sensitive initial conditions. How to determine the sensitive initial conditions will also be one of future research aspects for improving the performance of Bi-DDEA in practical application.

To let Bi-DDEA adapt to diversified and comprehensive constrained environment easily is another important goal of this paper. Although the six benchmark functions modified with the constrained condition can not give full instructions of Bi-DDEA's generalization capability, the samples distribution has provided robust evidence for the predominance of bidirection dynamic diversity evolution. Many more experiments and practices are needed for verifying the validity of this optimization method. In fact, to find exactly the types for which an optimization method fit will be more important and useful than to improve hardly the generalization capability of the method. In future study, surveying and concluding the best parameters of Bi-DDEA to match different types of optimization problems will be done to expand the scope in which Bi-DDEA can give better results. There are also other special application fields in which Bi-DDEA can be attempted to apply, such as discrete optimization problems [59], dynamic optimization problems [60], and multiobjective optimization problems [37, 61, 62]. So much more potential applications of Bi-DDEA are waiting for exploring and exploiting in future studies.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant no. 61074020 and the Fundamental Research Funds for the Central Universities.

References

- [1] A. Engelbrecht, X. Li, M. Middendorf, and L. M. Gambardella, "Editorial: special issue: swarm intelligence," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 4, pp. 677–680, 2009.
- [2] A. E. Smith, "Swarm intelligence: from natural to artificial systems [book reviews]," *IEEE Transactions on Evolutionary Computation*, vol. 4, no. 2, pp. 192–193, 2000.
- [3] J. Wang and G. Beni, "Swarm intelligence in cellular robotic systems," in *Proceedings of the NATO Advanced Workshop on Robots and Biological Systems*, Tuscany, Italy, June 1989.
- [4] H. Bai and B. Zhao, "A survey on application of swarm intelligence computation to electric power system," in *Proceedings of the 6th World Congress on Intelligent Control and Automation (WCICA '06)*, pp. 7587–7591, IEEE, Dalian, China, June 2006.
- [5] K. Kim, P. Rao, and J. Burnworth, "Application of swarm intelligence to a digital excitation control system," in *Proceedings of the IEEE Swarm Intelligence Symposium (SIS '08)*, pp. 1–8, IEEE Conference Proceedings, St. Louis, Mo, USA, September 2008.
- [6] E. Chapman and F. Sahin, "Application of swarm intelligence to the mine detection problem," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC '04)*, vol. 6, pp. 5429–5434, IEEE Conference Proceedings, October 2004.
- [7] G. Wang, W. Y. Zhang, Q. Ning, and H. L. Chen, "A novel framework based on aco and pso for rna secondary structure prediction," *Mathematical Problems in Engineering*, vol. 2013, Article ID 796304, 8 pages, 2013.
- [8] X. Fu, W. Liu, B. Zhang, and H. Deng, "Quantum behaved particle swarm optimization with neighborhood search for numerical optimization," *Mathematical Problems in Engineering*, vol. 2013, Article ID 469723, 10 pages, 2013.
- [9] W. Gao, C. Shao, and Q. Gao, "Pseudo-collision in swarm optimization algorithm and solution: rain forest algorithm," *Acta Physica Sinica*, vol. 62, no. 19, Article ID 190202, 2013.
- [10] X. Jiang, H. Ling, J. Yan, B. Li, and Z. Li, "Forecasting electrical energy consumption of equipment maintenance using neural network and particle swarm optimization," *Mathematical Problems in Engineering*, vol. 2013, Article ID 194730, 8 pages, 2013.
- [11] D. Sendrescu, "Parameter identification of anaerobic wastewater treatment bioprocesses using particle swarm optimization," *Mathematical Problems in Engineering*, vol. 2013, Article ID 103748, 8 pages, 2013.
- [12] J. Chen, Y. Ding, and K. Hao, "The bidirectional optimization of carbon fiber production by neural network with a gaipso hybrid algorithm," *Mathematical Problems in Engineering*, vol. 2013, Article ID 768756, 16 pages, 2013.
- [13] Y. Zhang, L. Wu, and S. Wang, "UCAV path planning by fitness-scaling adaptive chaotic particle swarm optimization," *Mathematical Problems in Engineering*, vol. 2013, Article ID 705238, 9 pages, 2013.
- [14] C. Liu, O. Yacine, N. Antoine, B. Abdelaziz, and J. Zhou, "The reputation evaluation based on optimized hidden markov model in ecommerce," *Mathematical Problems in Engineering*, vol. 2013, Article ID 391720, 11 pages, 2013.

- [15] X. Su, W. Fang, Q. Shen, and X. Hao, "An image enhancement method using the quantum-behaved particle swarm optimization with an adaptive strategy," *Mathematical Problems in Engineering*, vol. 2013, Article ID 824787, 14 pages, 2013.
- [16] P. Liu, W. Leng, and W. Fang, "Training anfis model with an improved quantum-behaved particle swarm optimization algorithm," *Mathematical Problems in Engineering*, vol. 2013, Article ID 595639, 10 pages, 2013.
- [17] J.-Y. Wu, "Solving unconstrained global optimization problems via hybrid swarm intelligence approaches," *Mathematical Problems in Engineering*, vol. 2013, Article ID 256180, 15 pages, 2013.
- [18] K. Deb, "An efficient constraint handling method for genetic algorithms," *Computer Methods in Applied Mechanics and Engineering*, vol. 186, no. 2–4, pp. 311–338, 2000.
- [19] R. Farmani and J. A. Wright, "Self-adaptive fitness formulation for constrained optimization," *IEEE Transactions on Evolutionary Computation*, vol. 7, no. 5, pp. 445–455, 2003.
- [20] S. Nema, J. Y. Goulermas, G. Sparrow, and P. Helman, "A hybrid cooperative search algorithm for constrained optimization," *Structural and Multidisciplinary Optimization*, vol. 43, no. 1, pp. 107–119, 2011.
- [21] S. He, E. Prempain, and Q. H. Wu, "An improved particle swarm optimizer for mechanical design optimization problems," *Engineering Optimization*, vol. 36, no. 5, pp. 585–605, 2004.
- [22] B. Yang, Y. Chen, Z. Zhao, and Q. Han, "A master-slave particle swarm optimization algorithm for solving constrained optimization problems," in *Proceedings of the 6th World Congress on Intelligent Control and Automation (WCICA '06)*, vol. 1, pp. 3208–3212, Dalian, China, June 2006.
- [23] X. Huang and T. Chai, "Particle swarm optimization for raw material purchasing plan in large scale ore dressing plant," *Acta Automatica Sinica*, vol. 35, no. 5, pp. 632–636, 2009.
- [24] C. Wang, M. Wu, W. Cao, and Y. He, "Intelligent integrated modeling and synthetic optimization for blending process in Lead-Zinc sintering," *Acta Automatica Sinica*, vol. 35, no. 5, pp. 605–612, 2009.
- [25] G. Liu, S. Lao, C. Yuan, L. Hou, and D. Tan, "Oacrrpso algorithm for anti-ship missile path planning," *Acta Automatica Sinica*, vol. 38, no. 9, pp. 1528–1537, 2012.
- [26] R. A. Vural, T. Yildirim, T. Kadioglu, and A. Basargin, "Performance evaluation of evolutionary algorithms for optimal filter design," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 1, pp. 135–147, 2012.
- [27] X. Li and X. Yao, "Cooperatively coevolving particle swarms for large scale optimization," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 2, pp. 210–224, 2012.
- [28] T. Blackwell, "A study of collapse in bare bones particle swarm optimization," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 3, pp. 354–372, 2012.
- [29] Y. Pehlivanoglu, "A new particle swarm optimization method enhanced with a periodic mutation strategy and neural networks," *IEEE Transactions on Evolutionary Computation*, vol. 17, no. 3, pp. 436–452, 2012.
- [30] W. N. Chen, J. Zhang, Y. Lin et al., "Particle swarm optimization with an aging leader and challengers," *IEEE Transactions on Evolutionary Computation*, vol. 17, no. 2, pp. 241–258, 2012.
- [31] F. Naznin, R. Sarker, and D. Essam, "Progressive alignment method using genetic algorithm for multiple sequence alignment," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 5, pp. 615–631, 2012.
- [32] Y. Wang and Z. Cai, "Combining multiobjective optimization with differential evolution to solve constrained optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 1, pp. 117–134, 2012.
- [33] Y. Wang, Z. Cai, Y. Zhou, and W. Zeng, "An adaptive tradeoff model for constrained evolutionary optimization," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 1, pp. 80–92, 2008.
- [34] Z. Cai and Y. Wang, "A multiobjective optimization-based evolutionary algorithm for constrained optimization," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 6, pp. 658–675, 2006.
- [35] R. A. Krohling and L. dos Santos Coelho, "Coevolutionary particle swarm optimization using gaussian distribution for solving constrained optimization problems," *IEEE Transactions on Systems, Man, and Cybernetics B*, vol. 36, no. 6, pp. 1407–1416, 2006.
- [36] B. Tessema and G. G. Yen, "An adaptive penalty formulation for constrained evolutionary optimization," *IEEE Transactions on Systems, Man, and Cybernetics A*, vol. 39, no. 3, pp. 565–578, 2009.
- [37] M. Daneshyari and G. G. Yen, "Constrained multiple-swarm particle swarm optimization within a cultural framework," *IEEE Transactions on Systems, Man, and Cybernetics A*, vol. 42, no. 2, pp. 475–490, 2012.
- [38] S. Venkatraman and G. G. Yen, "A generic framework for constrained optimization using genetic algorithms," *IEEE Transactions on Evolutionary Computation*, vol. 9, no. 4, pp. 424–435, 2005.
- [39] Y. Wang, Y. Jiao, and H. Li, "An evolutionary algorithm for solving nonlinear bilevel programming based on a new constraint-handling scheme," *IEEE Transactions on Systems, Man and Cybernetics C*, vol. 35, no. 2, pp. 221–232, 2005.
- [40] J. J. Grefenstette, "Optimization of control parameters for genetic algorithms," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 16, no. 1, pp. 122–128, 1986.
- [41] D. Whitley, T. Starkweather, and C. Bogart, "Genetic algorithms and neural networks: optimizing connections and connectivity," *Parallel Computing*, vol. 14, no. 3, pp. 347–361, 1990.
- [42] D. S. Weile and E. Michielssen, "Genetic algorithm optimization applied to electromagnetics: a review," *IEEE Transactions on Antennas and Propagation*, vol. 45, no. 3, pp. 343–353, 1997.
- [43] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the 1995 IEEE International Conference on Neural Networks*, vol. 4, pp. 1942–1948, Perth, Wash, USA, December 1995.
- [44] R. Eberhart and J. Kennedy, "New optimizer using particle swarm theory," in *Proceedings of the 1995 6th International Symposium on Micro Machine and Human Science*, pp. 39–43, Nagoya, Japan, October 1995.
- [45] M. Settles and T. Soule, "Breeding swarms: a GA/PSO hybrid," in *Proceedings of the Conference on Genetic and Evolutionary Computation (GECCO '05)*, pp. 161–168, ACM, Washington, DC, USA, June 2005.
- [46] J. Robinson, S. Sinton, and Y. Rahmat-Samii, "Particle swarm, genetic algorithm, and their hybrids: optimization of a profiled corrugated horn antenna," in *Proceedings of the 2002 IEEE Antennas and Propagation Society International Symposium*, vol. 1, pp. 314–317, IEEE, June 2002.

- [47] Y. Rahmat-Samii, “Genetic algorithm (ga) and particle swarm optimization (pso) in engineering electromagnetics,” in *Proceedings of the 17th International Conference on Applied Electromagnetics and Communications (ICECom '03)*, pp. 1–5, 2003.
- [48] J. Ronkkonen, X. Li, V. Kyrki, and J. Lampinen, “A framework for generating tunable test functions for multimodal optimization,” in *Soft Computing—A Fusion of Foundations, Methodologies and Applications*, pp. 1689–1706, Springer, 2010.
- [49] M. Meissner, M. Schmuker, and G. Schneider, “Optimized particle swarm optimization (OPSO) and its application to artificial neural network training,” *BMC Bioinformatics*, vol. 7, article 125, 2006.
- [50] C. MacNish, “Towards unbiased benchmarking of evolutionary and hybrid algorithms for real-valued optimisation,” *Connection Science*, vol. 19, no. 4, pp. 361–385, 2007.
- [51] R. Salomon, “Re-evaluating genetic algorithm performance under coordinate rotation of benchmark functions. A survey of some theoretical and practical aspects of genetic algorithms,” *BioSystems*, vol. 39, no. 3, pp. 263–278, 1996.
- [52] J. G. Digalakis and K. G. Margaritis, “An experimental study of benchmarking functions for genetic algorithms,” *International Journal of Computer Mathematics*, vol. 79, no. 4, pp. 403–416, 2002.
- [53] M. A. Potter, *The design and analysis of a computational model of cooperative coevolution [Ph.D. thesis]*, George Mason University, 1997.
- [54] W. Gao, “Study on immunized ant colony optimization,” in *Proceedings of the 3rd International Conference on Natural Computation (ICNC '07)*, vol. 3, pp. 792–796, Haikou, China, August 2007.
- [55] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, Springer, Berlin, Germany, 1996.
- [56] Y. Qing and D. Sheng-Chao, “Amplitude-based coding hybrid quantuminspired evolutionary algorithm,” *Computer Engineering and Applications*, vol. 43, no. 21, pp. 80–83, 2007.
- [57] J. Zhao and C. Yan, “A bottleneck assigned binary ant system for multimodal optimization,” in *Proceedings of the 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference*, pp. 6195–6200, Shanghai, China, December 2009.
- [58] D. Cvijović and J. Klinowski, “Taboo search: an approach to the multiple minima problem,” *Science*, vol. 267, no. 5198, pp. 664–666, 1995.
- [59] W. Chen, J. Zhang, H. S. H. Chung, W. Zhong, W. Wu, and Y. Shi, “A novel set-based particle swarm optimization method for discrete optimization problems,” *IEEE Transactions on Evolutionary Computation*, vol. 14, no. 2, pp. 278–300, 2010.
- [60] S. Yang and C. Li, “A clustering particle swarm optimizer for locating and tracking multiple optima in dynamic environments,” *IEEE Transactions on Evolutionary Computation*, vol. 14, no. 6, pp. 959–974, 2010.
- [61] M. Farina, K. Deb, and P. Amato, “Dynamic multiobjective optimization problems: test cases, approximations, and applications,” *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 5, pp. 425–442, 2004.
- [62] P. A. Bosman, “On gradients and hybrid evolutionary algorithms for real-valued multiobjective optimization,” *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 1, pp. 51–69, 2012.

Research Article

A New Hybrid Chaotic Map and Its Application on Image Encryption and Hiding

Yang Cao

College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications,
Nanjing 210003, China

Correspondence should be addressed to Yang Cao; caoyeacy@njuptsast.org

Received 4 June 2013; Accepted 27 September 2013

Academic Editor: Dumitru Baleanu

Copyright © 2013 Yang Cao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A new hybrid chaotic map is constructed in this paper and applied to image encryption and hiding. MD5 is used as part of initial condition and control parameter to perturb the trajectory in order to further increase the security against plaintext-chosen attacks and differential attacks. Ciphered image is embedded into several carrier images to reduce suspiciousness and to increase robustness. The results demonstrate that the proposed scheme is of great security, robustness, and efficiency.

1. Introduction

Owing to the rapid development of the Internet, a lot of confidential information transfers across the world over public networks in all fields of the society. However, those networks are not suitable for the direct transmission of private messages [1, 2]. Various types of files, such as text, image, and video [3–5] have to face the threat of unsafety. To keep secrecy and to make use of the networks already developed simultaneously, the actual and potential applications of cryptography extended prominently with the modern technological advance [6–8]. Since that, covert communication methods aroused the interest of many researchers and are becoming the world's attentive focus. Traditional encryption methods were an effective solution for information security [6, 9]. Nonetheless, most of conventional ciphers, such as Data Encryption Standard (DES) [8], International Data Encryption Algorithm (IDEA) [10], and Advanced Encryption Standard (AES) [11], are not suitable for image encryption for three main reasons. The first reason is that the image size is almost always much greater than that of text, which needs much time to directly encrypt the image data using the software implementations of traditional cryptosystems. The second reason is laying on its sensitivity. The decrypted text must be exactly equal to the original text. However, this requirement is not necessary for image data. Because of human perception property and the high redundancy of image data,

a decrypted image containing small distortion is usually acceptable. The third reason is because digital images' contents are strongly correlated. However, traditional cryptosystems never used those attributes, which hindered the promotion of encryption efficiency.

To adapt those unique characteristics of image data and to improve efficiency and security of image encryption, numerous special image encryption and hiding schemes were proposed. Among those schemes, the chaos based schemes have attracted the interest of many researchers from science and engineering realms. As ergodic, high sensitivity dependence on initial conditions, random-like behavior, continuous broad-band power spectrum, and unpredictability yet reproducible, chaos has huge potential applications in several vital fields of cryptosystem especially the image encryption. At the same time, different approaches for the transmission of information signals using chaotic dynamics were presented, such as chaotic masking, chaotic modulation, and chaotic switching. The idea of blending chaos theory into the cryptosystem came up since 1980s when Matthews [12] used chaos in applications to cryptology. After that numerous studies were done in the field of chaotic secure communication [13–16]. Baptista [17] proposed a text encryption using the ergodicity of the simple low-dimension and chaotic logistic map. But Álvarez et al. [18] pointed out that the method was not robust and secure enough. Afterwards, many improved encryption algorithms based on Baptista-type and other new ideas

including the usage of fractional chaotic map were raised [19, 20]. On the one hand, in order to guarantee the speed, some researchers continued to use one-dimensional chaotic maps for image encryption [21, 22]. But analysis in [23, 24] showed that these schemes had some security defects due to some limitations such as small key space and weak security function in one-dimensional chaotic maps. Although chaos is infinite acyclic theoretically, in practical applications, there still exist many unpredicted problems, such as periodic degradation on account of finite precision limitation of computers [25] and low efficiency both caused by huge data and complicated chaotic behavior. The short cycle length problem makes many attacks feasible. On the other hand, some scholars adopt high-dimensional chaos systems into image encryption to enhance security for hyperchaos [26] having more complex dynamical characteristics. Although these methods like adopting high dimension maps can extend the cycle to some extent, the huge defect of computing time cost can totally offset its advantages.

In this paper, a novel image encryption and hiding scheme based on the new hybrid chaotic map will be proposed in order to keep the security and efficiency simultaneously. Primarily, a novel approach of constructing a new hybrid chaotic map will be given. Based on composition of three classic chaotic maps: logistic map, Hénon map, and Ikeda map, the new hybrid chaotic map demonstrates a more complex chaotic characteristic than most of single chaotic maps by Lyapunov exponent comparison, which means the hybrid map enjoys stronger encryption properties and is more sensitive dependence on initial conditions. The study results are of profound theoretical and practical significance because with the approach of composition, generating a brand new and stronger chaotic maps can be much simpler than before, which avoid the threat of known-system based attacks to a large extent. On account of the discrete map based properties, the cryptosystem also shows great performance on its efficiency. In the encryption scheme, the new hybrid chaotic map will be served as the key stream generator for encryption.

Meanwhile, MD5 values will be used as part of initial condition and control parameter to perturb the trajectory during iterations considering the capability to resist different attacks and the potential degradation from the computer limited precision problems. Generally, an encryption scheme is only highly connected to a fixed key which shows the weakness in plaintext-chosen and differential attack. In [27], the author's new idea that makes the initial key to be dependent on MD5 values of the plaintext made the encryption more sensitive and secure. In our scheme, MD5 values of plain image will be allowed to participate in much more place in the iteration of the new hybrid chaotic map, not only the initial keys but also the control parameters, which will further increase the sensitivity of plain image. Besides, with combination of MD5 values and fixed key, it is much more complex than the ordinary system [27]. Hence, it will guarantee the safety of this system since there are few works about the adaptive estimation of unknown secret parameters not to mention in the unknown chaotic map. According to the statistical tests involving key space, histogram analysis, information entropy,

correlation coefficient, and mean square errors, this type of cryptosystem will be proved to be theoretically unbreakable by common attacks.

Furthermore, watermark technique and packet splitting strategy will be used together to embed ciphered image into several carrier images, which could make the ciphered image less conspicuous as well as more robust and meet the need of high efficient and robustness in covert packet networks. According to our simulations, those methods can greatly improve cipher's antinoise capability. Additionally, with MD5 embedding, an extra function of plain image confirmation can be added to the system based on its one-way associated property.

The rest of this paper will be organized as follows: Section 2 will give a descriptive presentation of the hybrid chaotic map and its great property. The design concept of proposed image encryption and hiding scheme will be described in Section 3. Simulation results and security analyses will be drawn in Sections 4 and 5. Finally, conclusions about the dissertation will be summarized.

2. The New Hybrid Chaotic Map

In this section, a new approach for constructing a new hybrid chaotic map will be put forward. In our scheme, the new hybrid chaotic map, which is constructed by composition of three classic chaotic maps: logistic map, Hénon map, and Ikeda map, reveals remarkable sensitivity to the initial condition and parameters comparing with their Lyapunov exponents. The detail process of construction and comparison results of Lyapunov exponents will be illustrated as follows.

2.1. Logistic Map. Logistic map [28, 29] is a polynomial mapping of degree 2 that exhibits chaotic behavior. The logistic map equation is given by

$$x_{n+1} = \mu x_n (1 - x_{n-1}), \quad (1)$$

where $x_n \in (0, 1)$, control parameter $\mu \in (3.57, 4]$. The bifurcation diagram of the map is shown in Figure 1. When the parameter approaches 4, the chaotic systems exhibit a superb sensitivity to initial conditions.

2.2. Hénon Map. Hénon map was introduced by Hénon [4]. It is a simple two-dimensional reversible nonlinear chaotic map which iterates the point (x_n, y_n) via the equation:

$$x_{n+1} = 1 - ax_n^2 + y_n, \quad y_{n+1} = bx_n, \quad (2)$$

where $a \in (0, 1.4]$, $b \in (0.2, 0.314]$ are the control parameters on which the map depends. For $a = 1.4$, $b = 0.3$, the system can notably demonstrate chaotic behavior. Figure 2 shows the 10000 successive points obtained by iteration of the mapping.

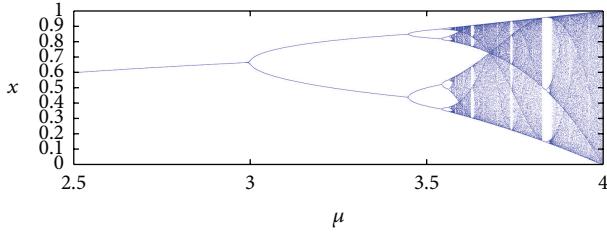


FIGURE 1: Bifurcation diagram of logistic map.

2.3. Ikeda Map. Ikeda map [30] is a discrete-time dynamical system. It is often used in a modified form, which is defined as the following equation:

$$\begin{aligned} x_{n+1} &= 1 + u(x_n \cos t_n - y_n \sin t_n), \\ y_{n+1} &= u(x_n \sin t_n - y_n \cos t_n), \end{aligned} \quad (3)$$

where u is a parameter and $t_n = c - (6/(1 + x_n^2 + y_n^2))$.

For $c = 0.4$, $u = 0.9$, $x_0 = y_0 = 0.1$, this system has a distinct chaotic attractor, which is shown in Figure 3.

2.4. The New Hybrid Chaotic Map. To generate a new compound mapping system, we composite (1)–(3) together with some necessary modification. The new hybrid mapping system is described as follows:

$$\begin{aligned} t_n &= c - \frac{6}{1 + x_n^2 + y_n^2}, \\ w_{n+1} &= 1 + u(w_n \cos t_n - s_n \sin t_n), \\ s_{n+1} &= u(w_n \sin t_n - s_n \cos t_n), \\ x_{n+1} &= 1 - a \cdot s_{n+1}^2 + b \cdot w_{n+1}, \\ y_{n+1} &= \mu \cdot w_{n+1} (1 - s_{n+1}), \end{aligned} \quad (4)$$

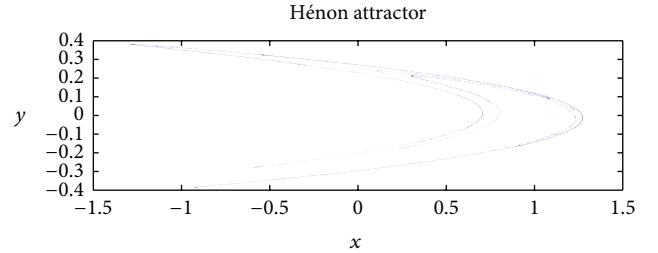
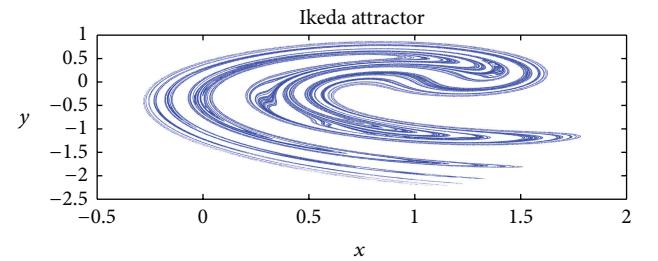
where $w_0 = s_0 \in (0, 1)$, $c \in [0.1, 0.4]$, $u \in [0.8, 0.85) \cup (0.85, 0.9]$, $b/a > 1$, $\mu > 1$, and the phase diagram of hybrid map has demonstrated a notable chaotic attractor. For $w_0 = s_0 = 0.1$, $c = 0.4$, $u = 0.9$, $a = 1.3$, $b = 4$, $\mu = 4$, the phase diagram of hybrid map is shown in Figure 4.

The value distribution over a random chosen initial key in $(0, 1]$ is shown in Figure 5.

This construction approach is of profound theoretical and practical importance because with the idea of compositing, more and more new and sensitive chaotic maps are able to generate much simpler than before, which could vastly avoid the threat of known-system based attack.

2.5. Comparison of Lyapunov Exponent. In the preceding paragraphs, the descriptions of chaos phenomenon are qualitative rather than quantitative. To make a more quantitative analysis on those maps, their Lyapunov exponents will be calculated and compared in this section.

The Lyapunov exponent [31] defined as formula (5) is a common method to study the mean exponential rate of

FIGURE 2: Phase diagram of Hénon map ($x_0 = y_0 = 0.4$, $a = 1.4$, $b = 0.3$).FIGURE 3: Phase diagram of Ikeda map ($u = 0.9$).

divergence of two initially close orbits over time, which indicates the sensitive dependence on initial conditions:

$$L(f, x_0) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln \left| \left(f^i \right)' (x_0) \right|, \quad (5)$$

where $(f^i)'$ is the derivative of the i th iterate (f^i) .

$L < 0$: negative Lyapunov exponents are of dissipative or nonconservative systems, which indicates the orbit attracts to a stable fixed point or stable periodic orbit. Such systems exhibit asymptotic stability. The more negative the exponent, the greater the stability.

$L = 0$: a Lyapunov exponent of zero shows that the system is in some sort of steady state mode, which signs that the orbit is a fixed point (or an eventually fixed point).

$L > 0$: positive Lyapunov exponents reveal that the orbit is unstable and chaotic. Nearby points, no matter how close, will eventually diverge to any arbitrary separation [32].

Because sensitive dependence can be arisen only in some portions of a system (such as the logistic map), this separation is also a function of the location of the initial value and has the form $(f^i)'(x_0)$ for an orbit of f with starting point x_0 . In a system with attracting fixed points or attracting periodic points, L diminishes asymptotically with time. If a system is unstable, like pins balanced on their points, then the orbits diverge exponentially for a while but eventually settle down. For chaotic situation, the function will behave erratically [32].

For a 2D mapping system, there are two Lyapunov exponents L_x , L_y for x and y orbits. Chaos only appears in the system which has at least one positive Lyapunov

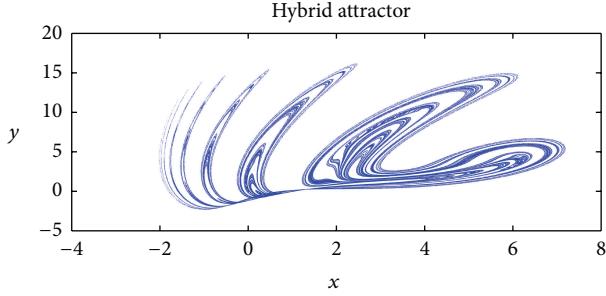


FIGURE 4: Phase diagram of the hybrid chaotic Map.

TABLE 1: Maximum Lyapunov exponent comparison.

Maps	L_x	L_y
Logistic map	0.6913	\
Hénon map	0.4222	-1.6261
Ikeda map	0.4926	-0.7298
Hybrid chaotic map	0.5077	-0.7329

exponent. Figures 6, 7, and 8 illustrate the Lyapunov exponent spectrums of the discussed maps.

Through the analysis of the Lyapunov exponent, the new hybrid chaotic map has one positive Lyapunov exponent, which can be theoretically convinced that the hybrid mapping system is in a chaotic state. Furthermore, through the comparison between the four mapping systems in Table 1, it is easy to find out that the new hybrid system has the largest positive Lyapunov exponent along with the second largest negative Lyapunov exponent over all original two dimensional maps, which indicates that it has more obvious chaotic characteristics and is much more sensitive dependence on initial conditions than most of common single chaotic maps. Thus, it also denotes much stronger encryption abilities.

3. Chaos-Based Image Processing Scheme

3.1. Overall Process. In this section, the whole image processing scheme will be presented in detail, which will be constructed by two major parts—the encryption part and the hiding part. To enhance the security of image encryption, in the encryption part, a new algorithm based on the blending of the hybrid chaotic mapping system and MD5 hash will be proposed. In the hiding part, the Least Significant Bit (LSB) based watermarking embedded algorithm with packet division scheme will be adopted to hide and to split the ciphered image into different carrier images, which make the ciphered image less conspicuous as well as more robust and meet the need in high efficient packet communication.

The basic architecture of the whole chaos-based image processing scheme is displayed in Figure 9 while detailed explanation is presented afterwards.

3.2. Encryption Scheme

3.2.1. Generation of Initial Key. Our key includes initial user-set fixed real number key—key₀, control parameter

a, b, c, u, μ in (4), and MD5 values decided by the plain image. So the 128 bit MD5 value should be computed in advance, given by a color plain image P with size of $M \times N \times 3$, where $P_{i,j,\text{dim}}$ represents pixel in i th row j th column in dimth color space (using RGB color space). After calculation, the 128 bit MD5 values of the plain image are split up into 32 4 bit blocks in decimal format—MD5(i) $1 \leq i \leq 32$. (To achieve a higher speed the block number can be reduced through an adjustable block splitting algorithm with size of plain image.) Then each MD5 block MD5(i) and the user-set key key₀ will act on the initial values w_0, s_0 and the control parameter μ of the hybrid chaotic map which is given by the following equations for i th iteration:

$$\begin{aligned} \text{key}_1 &= \frac{\text{key}_0 + 1}{10^{\text{length}(\text{key}_0)} + 1}, \\ w_0 = s_0 &= \frac{[\text{MD5}(i) + 1]^{\text{key}_1}}{[\text{MD5}(i)_{\max} + 1]^{\text{key}_1}}, \\ \mu &= 4 - \frac{w_0}{10}, \end{aligned} \quad (6)$$

where $\text{MD5}(i)_{\max} = 16$, for splitting up to 32 parts.

Hence, the key not only disturbs the initial value of the new hybrid chaotic map but also adds perturbations to the control parameter to the chaos trajectory, which signifies that each pseudorandom sequence is generated from switching the trajectory from different hybrid chaotic map. Meanwhile, with the on-way property of MD5 hash, the decrypted image is able to audit with MD5. Since there is seldom work about the adaptive estimation of the unknown secret parameters in the unknown chaotic mapping systems even if the cryptanalyst obtains the whole trajectories, it not only guarantees the safety but also makes verification available.

3.2.2. Encryption Algorithm. As previously discussed, for i th iteration, the initial condition w_0, s_0, μ is given by (6) and the basic idea of the algorithm is shown in Figure 9. The detailed encryption flow is reflected in the next several steps. After that, the ciphered image will be created.

Step 1 (generation of key stream). Subiterative calculation of the chaotic sequence X, Y by (4) as the algorithm described below. The time of subiterations is equal to $M \times N - 1$. Then a chaotic key stream is generated. Consider

$$\text{for } n = 0 : 1 : M \times N - 1$$

$$t = c - \frac{6}{1 + w^2(n) + s^2(n)},$$

$$w(n+1) = 1 + u [w(n) \cos t - s(n) \sin t],$$

$$s(n+1) = u [w(n) \sin t - s(n) \cos t], \quad (7)$$

$$x(n+1)_{n+1} = 1 - a \times s^2(n+1) + b \times w(n+1),$$

$$y(n+1)_{n+1} = \mu \times w(n+1) \times [1 - s(n+1)],$$

end.

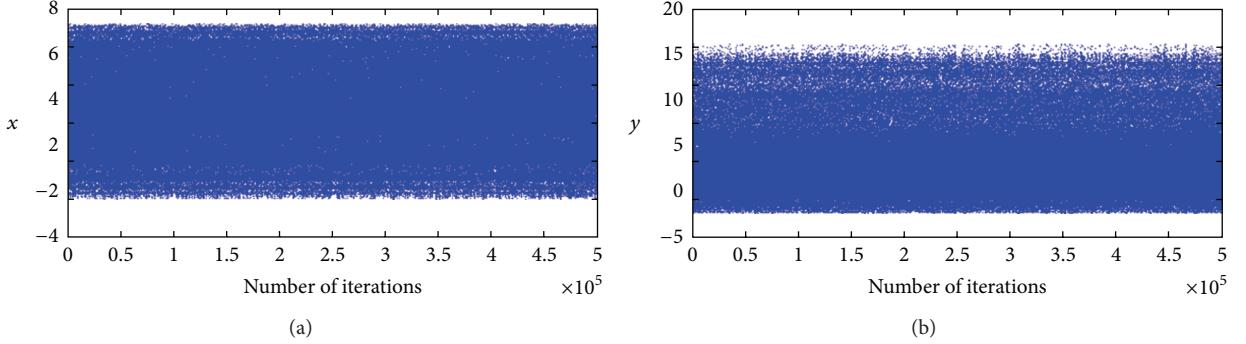


FIGURE 5: x_n (a), y_n (b) values distribution over iterations.

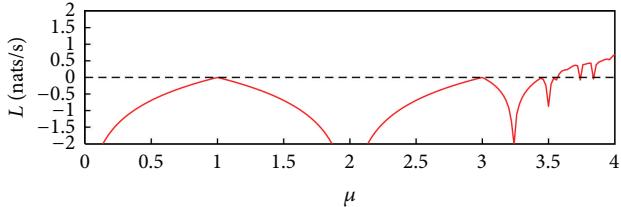


FIGURE 6: Lyapunov spectrum of logistic map.

Step 2 (XOR encryption). Via the algorithm described by (8) below, the XOR password sequences MX, MY are constructed with the chaotic key stream obtained from Step 1. Consider

$$\begin{aligned} mx(n) &= \text{mod}\left(10^{13+\text{key}_1^{-1}} \cdot x(n), 256\right), \\ my(n) &= \text{mod}\left(10^{14+\text{key}_1^{-1}} \cdot y(n), 256\right), \end{aligned} \quad (8)$$

$\text{mod}(\dots, 256)$ for making the XOR passwords of an approximate uniform distribution since there are 256 gray levels of each pixel.

After that, to further encrypt the pixel, we XOR this pseudorandom passwords sequences with the gray level in plain image bit by bit through Algorithm 1.

After that, with the switching method employed, the preciphered image C' is obtained, which is able to avoid detection of the chaotic trajectories.

Step 3 (chaotic scrambling). Firstly, remove part of the backward elements in key stream X and Y to make the length of X' and Y' equal to M and N . Map $X'(i)X'(j)$ with $C'_{i,j,\dim}$ $1 \leq \dim \leq 3$ and save the mapping relationship.

Secondly, range X' and Y' from the smallest to the largest. Scramble C' with the mapping relation and then the final ciphered image C is developed.

Apparently, after dual-perturbing effect of Steps 2 and 3 is much better to resist the influence caused by computer precision and can lengthen the period.

3.3. Hiding Scheme and Packet Splitting Scheme. In former process, the plain image is encrypted into the ciphered image. Despite being secure enough, the ciphered image is too notable for attacker due to its strikingly chaotic attribute. In

order to make the ciphered image less conspicuous and to increase the robustness and efficiency in real communication, it is essential to use the ciphered image as a watermark which is embedded into several carrier images. The detailed explanation is as follows.

To simplify, the following embedding algorithm is only built with the situation with two carrier images offered. Some feasible adjustments are also declared for the case of more carrier images offered.

To maximize the channel utilization rate for two carrier images' case, the spatial domain based Least Significant Bit (LSB) embedding algorithm will be adopted, because it only needs the carrier image as the same size as the plain image and will not increase the size of the embedded image. For case of more carrier images, transform domain based embedding algorithms are also available and will be much easier to neglect by human perceptual system [33]. However, as a consequence of the limited space, only the LSB algorithm is introduced in detail. Here is the algorithm in steps.

Step 1. Given two color carrier images \mathbf{Ca} with size of $\widetilde{M} \times \widetilde{N} \times 3$ and then scale them to the same size of the plain image $M \times N \times 3$. $\mathbf{Ca}_{i,j,\text{dim}}$ represents pixel in i th row j th column in dimth color space after adjustment. For case of more carrier images, the adjustment is not necessary.

Step 2. For a uint8 format color image, each pixel has 256 grey levels in each RGB dimension, which can be represented by 8 bits binary codes for each dimension. So each pixel in each dimension can be divided into 8 bit planes, which can be represented by a binary image. Because human vision is not sensitive to the least bit plane. So for having two carrier images, the least 4 bits for each of two carrier images are available for hiding the ciphered image. Thus the least 4 bits of each of two carrier images are set to zero from operations, as shown in Algorithm 2.

For case of more carrier images, we can even use part of the least important bit in each carrier image to make the watermark become much more difficult to notice through LSB algorithm.

Step 3. Embed the 32 MD5 parts $\text{MD5}(i)$ of plain images into the embedded image with the user set key key_0 for watermark detection, decryption, and validation; see Algorithm 3.

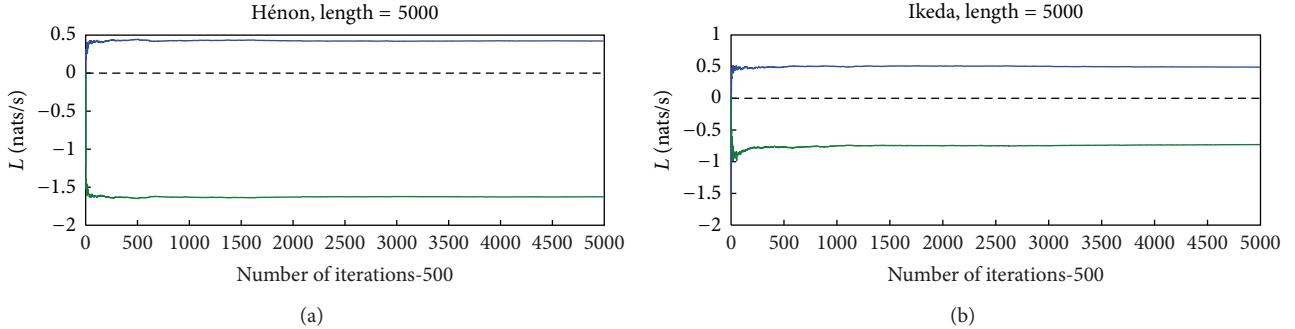


FIGURE 7: Lyapunov spectrum of Hénon map (a) and Ikeda map (b) (the first 500 iterations are ignored to avoid the influence of initial state).

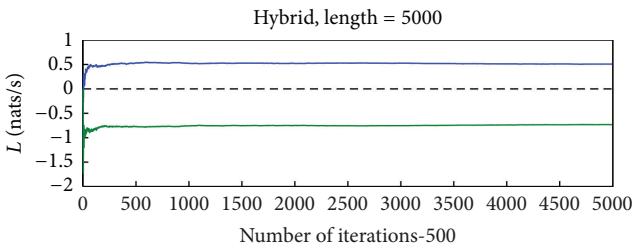


FIGURE 8: Lyapunov spectrum of the new hybrid chaotic map (the first 500 iterations are ignored to avoid the influence of initial state).

Consider

$$\begin{aligned} m_{\text{MD5}} &= \left\lfloor (M - 16) \cdot \frac{\text{key}_0 + 1}{10^{\text{length}(\text{key}_0)} + 1} \right\rfloor, \\ n_{\text{MD5}} &= \left\lfloor (N - 16) \cdot \frac{\text{key}_0 + 1}{10^{\text{length}(\text{key}_0)} + 1} \right\rfloor. \end{aligned} \quad (9)$$

Step 4. To detect the order of carrier image, a gray pixel should be selected, which does not cover the MD5 value and add the order value into its high bit part for both the transmitter and receiver. This is not necessary for case of two carrier images, but it is necessary for case of more carrier images.

Step 5. Split the ciphered images into two parts. To simplify, the two part is divided into the high 4-bit part and least 4-bit part. Then move the 4 bits with information to the least 4 bits and set the high 4 bits to zero, as shown in Algorithm 4.

For case of having more carrier images, the remaining chaotic maps can be used for random pixel selection in the ciphered image so as to embed the watermark in the carrier images' locations. This method will further increase the robustness of the watermark and further enhance the security.

Step 6. Embed the two parts of ciphered images into the carrier images from operations, as shown in Algorithm 5.

The switching operation is aimed at well-distributing the two parts into each carrier image and increasing the robustness. For case of more carrier images, the robustness will greatly increase because each embedded image carries

less information. With transform domain based embedding algorithms for more carrier images, the robustness can also be increased rapidly. Meanwhile, with channel coding technique, the broken mark can even be recovered.

3.4. Watermark Extraction and Decryption Scheme. The extraction and decryption are the inverse process of the embedding and encryption. Here is the watermark extraction and decryption scheme in five brief steps.

Step 1. Extract the order value and sort the embedded images in right order.

Step 2. Extract the MD5 value of plain image from the embedded image.

Step 3. Extract ciphered image information from several embedded image and join together to recover the ciphered image.

Step 4. Use the MD5 value and fixed key key_0 to generate chaotic sequence to realign and decrypt the ciphered image.

Step 5. Calculate the MD5 value of decrypted image and validate with comparison whether the watermark is modified.

4. Experimental Results

The scheme is run on MATLAB 8.0 in a personal computer. The 550×366 color image (Figure 10(a)) sized 590 KB is used as the plain image. To simplify, two carrier images are chosen as the same 550×366 color image (Figure 10(c)) sized 590 KB. The secret keys and parameters are set as follows:

$$\begin{aligned} \text{key}_0 &= 100, & w_0 = s_0 &= 0.1, & u &= 0.9, \\ a &= 1.3, & b &= 4, & c &= 0.4. \end{aligned} \quad (10)$$

The order values are chosen as binary code "0000," "1111" to make a distinction of order between embedded images 1 and 2 in space of $M1(1, 1, 1)$ and $M2(1, 1, 1)$.

The ciphered image is shown in Figure 10(b). Figure 10(d) illustrates one of embedded images.

The watermark extraction and decryption results are presented in Figure 10.

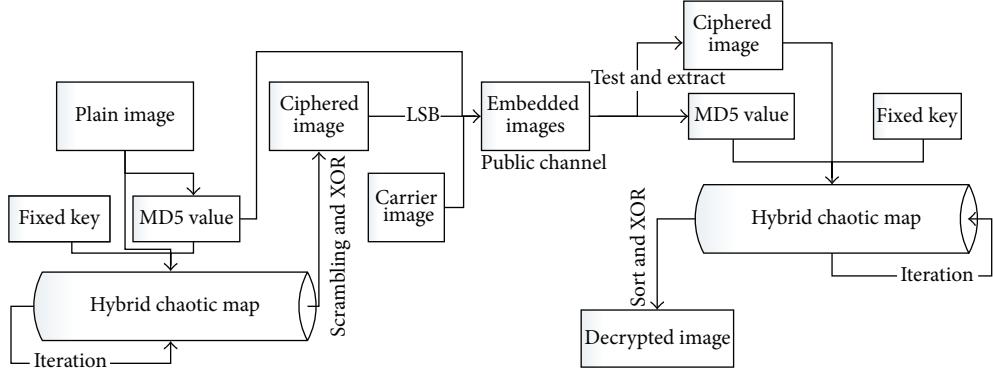
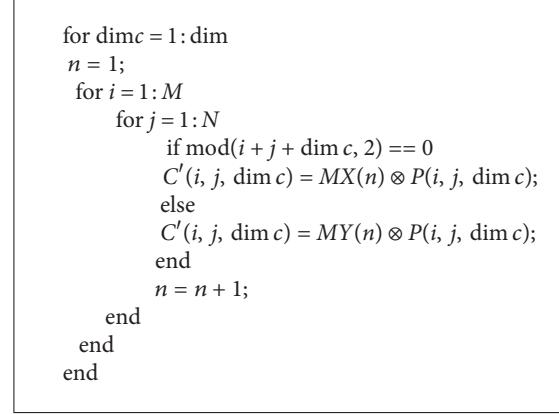


FIGURE 9: Architecture of chaos-based image processing scheme.



ALGORITHM 1

In Figure 11(a), we recover plain image when we use the correct key. And Figure 11(b) shows that uses the correct control parameter except $\text{key}_0 = 100 + 10^{-16}$. The same phenomenon like Figure 11(b) also appears when any control parameters change over 10^{-13} and the extraction order is incorrect.

Hence, it can be summarized that even a very tiny difference in keys, control parameters, and the extraction order can cause complete decryption failure. And Figure 12(b) shows the differences between two ciphers, whose corresponding plain images have only one gray pixel different. From the results, it can be concluded that the proposed scheme has high sensitivity to keys and plain image.

5. Security and Performance Analysis

Several tests are performed to check the security and robustness of the proposed cryptosystem. Statistical tests are performed, including key space, histogram analysis, information entropy, correlation coefficient, mean square error (MSE), and peak signal-to-noise ratio (PSNR). Likewise, security tests and analysis, such as chosen-plaintext attack and differential attack, are also presented on the scheme. The results are revealed below.

5.1. Key Space. For our proposed cryptosystem, the following keys are included in the key space:

(1) seven control parameters of the hybrid chaotic map; (2) initial fixed key; (3) MD5 values of plain image; (4) two order values; (5) iteration times. To ensure speed, assume the iteration times to be limited less than 10^3 times. Because of the computer precision problem, chaotic state value cannot be infinite. Based on our test results, the chaos behavior works well when accuracy is under 10^{-16} for initial fixed key and at least 10^{-13} for control parameters. Accordingly, the total key space is $10^3 \times 10^{13 \times 7} \times 10^{16} \times 10^6 \times 2^{4 \times 2} \times 2^{128} \gg 2^{128}$, which is sufficiently large to brute-force attack.

5.2. Histogram Analysis. As is known to all, well-behaved encryption algorithms should resist all attacks based on statistical analysis, which requires that the distribution of ciphered image's elements should be close to uniform distribution under any circumstances of plain image's element distribution. Figure 13 is the gray level histogram of both the plain and the ciphered image. From the histogram, it is clear to see that the gray level distribution in ciphered image tends to be uniform distribution, which covers the statistical characteristic of the plain image.

```

for dim c = 1:dim
    for i = 1:M
        for j = 1:N
            Ca(i, j, dim c) = Ca(i, j, dim c)  $\wedge$  240; % (240)Dec = (11110000)Bin
        end
    end
end

```

ALGORITHM 2

```

for i = 1:16
    dim1 = mod(mMD5 + nMD5 + i, 3); dim2 = mod(mMD5 + nMD5 + i + 16, 3);
    Cal(mMD5 + i, N - nMD5 - i, dim1) = Cal(mMD5 + i, N - nMD5 - i, dim1)  $\wedge$  15;
    Cal(mMD5 + i, N - nMD5 - i, dim1) = Cal(mMD5 + i, N - nMD5 - i, dim1)  $\vee$  [MD5(i)  $\cdot$  16];
    Ca2(M - mMD5 - i, nMD5 + i, dim2) = Ca1(M - mMD5 - i, nMD5 + i, dim2)  $\wedge$  15;
    Ca2(M - mMD5 - i, nMD5 + i, dim2) = Ca2(M - mMD5 - i, nMD5 + i, dim2)  $\vee$  [MD5(i + 16)  $\cdot$  16];
end

```

ALGORITHM 3

5.3. Information Entropy. According to the Shannon theory [34], information entropy is the most essential expressive feature of randomness. We can calculate information entropy $H(m)$ over (11):

$$H(m) = -\sum_i p(m_i) \log_2 \frac{1}{p(m_i)}, \quad (11)$$

where m is the message, m_i represents the probability of symbol m_i , and the entropy is expressed in bits. So to a pixel of which a symbol is encoded by $8 \times 3 = 24$ bits (3 means 3 color planes), when ideally random, entropy should be 24; in general, the entropy value of the source is smaller than 24 but ought to be close to be ideal. From our calculation, the entropy obtained from the ciphered image equals to 23.9973, which is very close to the ideal value 24, while the entropy obtained from the plain image is only 21.8747.

5.4. Correlation Coefficient. Correlation coefficient reflects the connection between pixels. According to the following formula, we obtain the correlation coefficients of two adjacent pixels and of corresponding pixels between two different images

$$r_{xy} = \frac{\text{cov}(x, y)}{\sqrt{D(x)} \sqrt{D(y)}}, \quad (12)$$

where $\text{cov}(x, y) = (1/N) \sum_{i=1}^N (x_i - E(x))(y_i - E(y))$, $E(x) = (1/N) \sum_{i=1}^N x_i$, $D(x) = (1/N) \sum_{i=1}^N (x_i - E(x))^2$.

Table 2 displays the results of correlation coefficients, and Figure 14 shows the correlation distribution. Results point out that the correlation between two adjacent pixels in ciphered image and between corresponding pixels in two different images is tremendously small, which displays its great encryption performance.

5.5. Noise Interference and Mean Square Error (MSE). Mean square error (MSE) is key value to quantify the difference between two images in statistics, which can demonstrate the average difference between different ciphered images and between plain and ciphered images. It also reflects the robustness in watermarks when the channel is not ideal. The definition of mean square error is according to the following formula:

$$\text{MSE} = \frac{1}{M \cdot N \cdot \text{dim}} \sum (P_{i,j,\text{dim}} - \widehat{D}_{i,j,\text{dim}})^2, \quad (13)$$

where $\widehat{D}_{i,j,\text{dim}}$ means another image's corresponding pixel.

Table 3 shows the results of MSE. Results indicate that the difference between different ciphered images and between plain and ciphered image is hugely immense, while the difference between decrypted and plain images both with and without the noise interference is relatively small, which is illustrated in Figure 15 and shows both the great encryption performance and the great robustness of the scheme.

5.6. Peak Signal-to-Noise Ratio (PSNR). Peak signal-to-noise ratio (PSNR) is a ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Because many signals have a very wide dynamic range, PSNR is usually expressed in terms of the logarithmic decibel scale [35]. The PSNR is defined as

$$\text{PSNR} = 10 \log_{10} \left(\frac{\text{MAX}_I^2}{\text{MSE}} \right). \quad (14)$$

Here, MAX_I is the maximum possible pixel value of the image. When the pixels are represented using 8 bits per sample, $\text{MAX}_I = 255$. Typical values for the PSNR in image and video compression are between 30 and 50 dB, where higher is better [35]. For previous simulation of noise interference, PSNR of the decrypted image with 0.001 Gaussian

```

for dim c = 1 : dim
    for i = 1 : M
        for j = 1 : N
            HC(i, j, dim c) = C(i, j, dim c)  $\wedge$  240;
            LC(i, j, dim c) = C(i, j, dim c)  $\wedge$  15;
        end
    end
end
HC(i, j, dim c) = HC(i, j, dim c)/16;

```

ALGORITHM 4

```

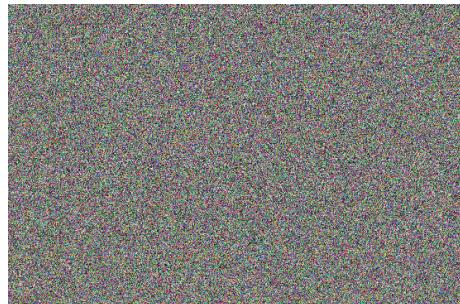
for dimw = 1 : dim
    for i = 1 : M
        for j = 1 : N
            if mod(i + j, 2) == 0
                M1(i, j, dimw) = Cal(i, j, dimw)  $\vee$  HC(i, j, dimw);
                M2(i, j, dimw) = Ca2(i, j, dimw)  $\vee$  LC(i, j, dimw);
            else
                M2(i, j, dimw) = Ca2(i, j, dimw)  $\vee$  HC(i, j, dimw);
                M1(i, j, dimw) = Cal(i, j, dimw)  $\vee$  LC(I, j, dimw);
            end
        end
    end
end

```

ALGORITHM 5



(a) Plain image



(b) Cipher image

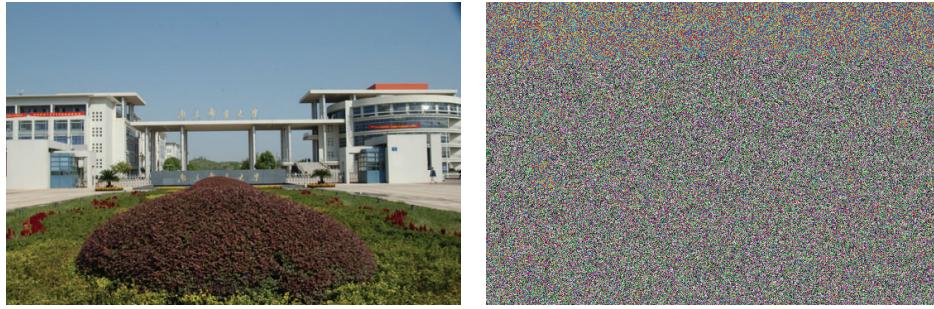


(c) Carrier image



(d) Watermark embedded

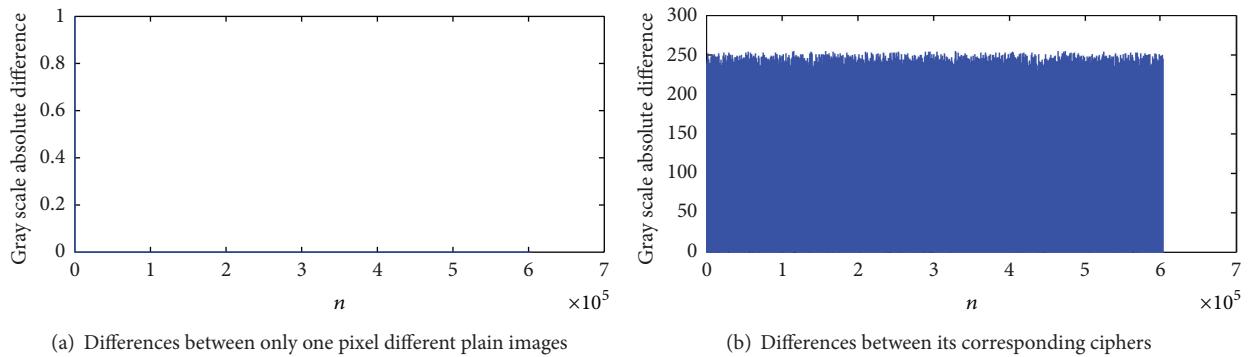
FIGURE 10: Encryption and watermark embedding results.



(a) Success of decryption

(b) Failure of decryption

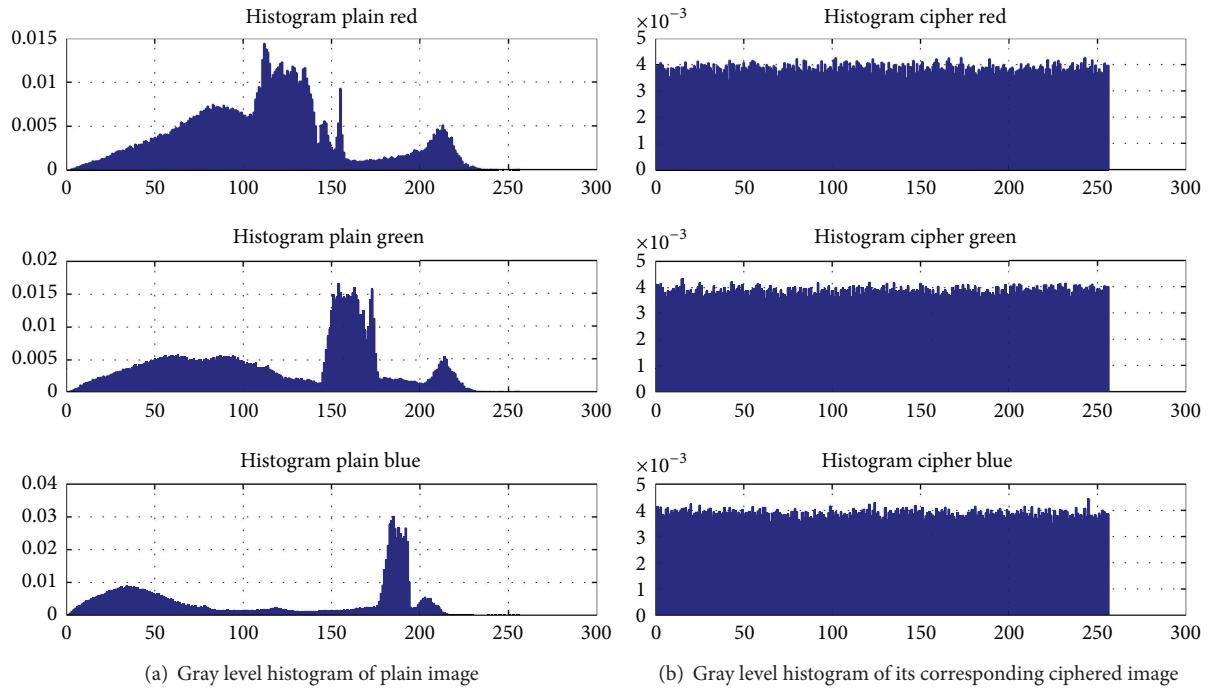
FIGURE 11: Watermark extraction and decryption results.



(a) Differences between only one pixel different plain images

(b) Differences between its corresponding ciphers

FIGURE 12: Differences comparison.



(a) Gray level histogram of plain image

(b) Gray level histogram of its corresponding ciphered image

FIGURE 13: Gray level histogram comparisons.

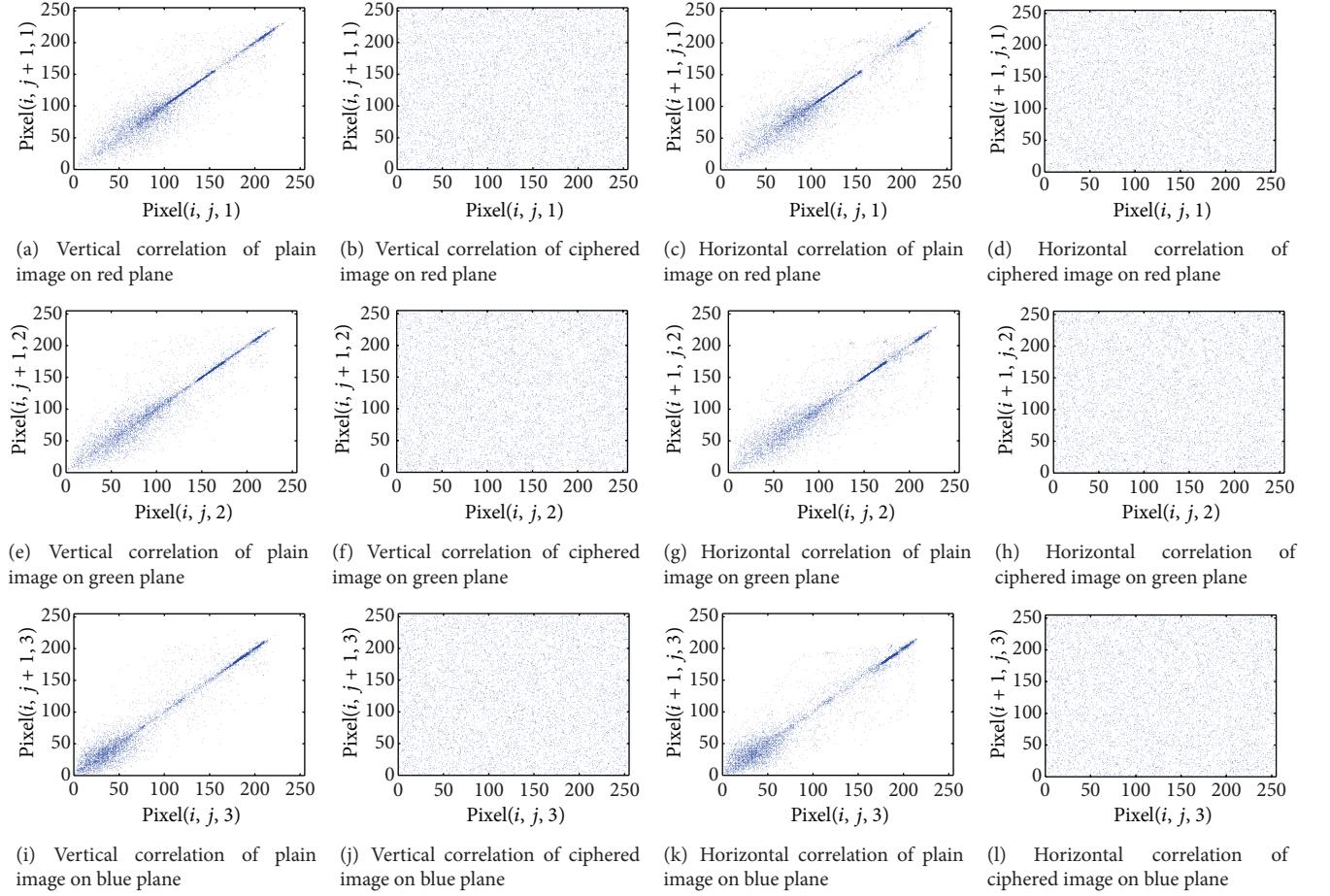


FIGURE 14: Vertical (left) and horizontal (right) correlation of plain image and ciphered image.



FIGURE 15: Noise interference comparisons.

TABLE 2: Correlation coefficient comparison.

		Between two adjacent pixels		Between corresponding pixels in two different images	
		Plain image	Ciphered image	Different ciphered images ²	Plain-ciphered
Correlation coefficient	Vertical	0.96684049	$7.16211445 \times 10^{-5}$		
	Horizontal	0.954339329	-0.0010516	$7.29256967 \times 10^{-4}$	-0.00135729535
	Diagonal	0.935083629	0.00151382		

²The two difference ciphered images' corresponding plain images only have one gray pixel in difference.

TABLE 3: MSE comparison.

MSE	Corresponding ciphered image	Corresponding decrypted image
Plain image	127.7904	0 (channel without any noise) 36.9984 (channel with 0.001 Gaussian noise)
Different ciphered image	117.2472	106.0239

noise interference is 32.4490 dB, which denotes the decent robustness of the system.

5.7. Plaintext-Chosen Attacks Analysis. Many encryption algorithms are faced with the risk of being broken by the chosen-plaintext analysis. But the proposed scheme in this paper combining the chaotic map's initial key with the plain image's MD5 value, which is one-way dependent on the plain image's information, makes the chosen-plaintext attack unavailable. Because even though much progress has been made in the field of attacking the MD5 algorithm in recent years [36], the price of finding a group of pictures with the same MD5 value and the great characteristic for attacking analysis, such as plaintext-chosen attack and differential attack, is still extremely expensive. Meanwhile, by using the block cipher of MD5 like the 32 blocks in Section 3.2.1, the difficulty of MD5 collide is increased. As a result, just like the simulation results in Section 4, for different plain images (even a tiny little change), the corresponding sequences are totally different on account of the property of MD5 hash and chaos. By these means, our proposed algorithm is robust to plaintext-chosen attack.

5.8. Differential Attacks Analysis. There are two most common magnitudes—Number of Pixels Change Rate (NPCR) and Unified Average Changing Intensity (UACI), which are used to evaluate the strength of image encryption in regard to differential attacks. Conventionally, a high NPCR and UACI score usually signifies a high resistance to differential attacks. The definitions of NPCR and UACI are stated by (15) [37]. To do the test, let us assume two ciphered images (C, C') whose corresponding plain images only have one pixel difference initially:

$$\begin{aligned} \text{NPCR} &= \frac{\sum_{i,j,\text{dim}} D(i, j, \text{dim})}{W \times H \times \text{dim}} \times 100\%, \\ \text{UACI} &= \frac{1}{W \times H \times \text{dim}} \\ &\quad \times \left[\sum_{i,j,\text{dim}} \frac{|C(i, j, \text{dim}) - C'(i, j, \text{dim})|}{255} \right] \\ &\quad \times 100\%, \end{aligned} \quad (15)$$

where W and H represent the width and height of the image, respectively. $C(i, j, \text{dim})$ and $C'(i, j, \text{dim})$ are the ciphered images before and after one gray level of one pixel

in the plain image is changed. For position (i, j, dim) , if $C(i, j, \text{dim}) \neq C'(i, j, \text{dim})$, then $D(i, j, \text{dim}) = 1$; else $D(i, j, \text{dim}) = 0$.

Based on our test results, the NPCR is 99.67% and the UACI is 33.45%, which proves that the proposed algorithm is very sensitive to small changes in the plain image and can survive to differential attack.

6. Conclusion

In this paper, an innovative hybrid chaotic mapping system has been crafted through compositing with three classic chaotic maps. Additionally, its application on image encryption and hiding has also been discussed with a demonstration scheme. Notably, from our test results, the new map has the largest positive Lyapunov exponent and the second largest negative Lyapunov exponent over the original 2D chaotic maps. What is more, this constructing method is of profound theoretical and practical prominence because with the plan of composition, more new and sensitive chaotic maps will come to the fore much easily, which could extremely avoid the threat of known-system based attack. In the image encryption scheme, the new hybrid chaotic map has been functioned as the key stream generator for encryption in view of its impressive efficiency and pseudorandom properties. On top of that, to improve the facilities of resisting plaintext-chosen and differential attacks and to make pseudorandom keys much more unpredictable, the MD5 hash of plain image has been used as part of initial condition and control parameter to perturb the trajectory during iterations. Simulation results and security analysis have disclosed that the proposed scheme has marvelous efficiency and robustness and is proved to be theoretically unbreakable by conventional attacks.

References

- [1] N. Doraswamy and D. Harkins, *Ipsec: the New Security Standard For the Internet, Intranet and Virtual Private Network*, Prentice Hall, New York, NY, USA, 2003.
- [2] B. Schneier, *Applied Cryptography: Protocols, Algorithms, and Source Code in C*, John Wiley & Sons, New York, NY, USA, 2007.
- [3] J. K. Su, F. Hartung, and B. Girod, "Digital watermarking of text, image, and video documents," *Computers and Graphics*, vol. 22, no. 6, pp. 687–695, 1998.
- [4] M. Hénon, "A two-dimensional mapping with a strange attractor," *Communications in Mathematical Physics*, vol. 50, no. 1, pp. 69–77, 1976.
- [5] A. Uhl and A. Pommer, *Image and Video Encryption: From Digital Rights Management to Secured Personal Communication*, vol. 15, Springer, New York, NY, USA, 2004.
- [6] L. Kocarev and S. Lian, *Chaos-Based Cryptography: Theory, Algorithms and Applications*, vol. 354, Springer, New York, NY, USA, 2011.
- [7] P. Stavroulakis and M. Stamp, *Handbook of Information and Communication Security*, Springer, New York, NY, USA, 2010.
- [8] M. E. Smid and D. K. Branstad, "Data encryption standard: past and future," *Proceedings of the IEEE*, vol. 76, pp. 550–559, 1988.
- [9] S. William, *Cryptography and Network Security*, 4/E, Pearson Education, New Delhi, India, 2006.

- [10] J. L. Freitas and R. R. Oliveira, “Implementação, comparação e análise dos algoritmos IDEA e DES,” in *I Encontro Regional em Modelagem e Análise Computacional de Sistemas (ERMACS '04)*, Goiânia, Brazil, 2004.
- [11] J. Daemen and V. Rijmen, *The Design of Rijndael: AES-the Advanced Encryption Standard*, Springer, New York, NY, USA, 2002.
- [12] R. Matthews, “On the derivation of a “chaotic” encryption algorithm,” *Cryptologia*, vol. 13, no. 1, pp. 29–42, 1989.
- [13] K.-W. Wong, B. S.-H. Kwok, and W.-S. Law, “A fast image encryption scheme based on chaotic standard map,” *Physics Letters A*, vol. 372, no. 15, pp. 2645–2652, 2008.
- [14] C. K. Huang and H. H. Nien, “Multi chaotic systems based pixel shuffle for image encryption,” *Optics Communications*, vol. 282, no. 11, pp. 2123–2127, 2009.
- [15] S. Behnia, A. Akhshani, S. Ahadpour, H. Mahmodi, and A. Akhavan, “A fast chaotic encryption scheme based on piecewise nonlinear chaotic maps,” *Physics Letters A*, vol. 366, no. 4–5, pp. 391–396, 2007.
- [16] X.-Y. Wang, L. Yang, R. Liu, and A. Kadir, “A chaotic image encryption algorithm based on perceptron model,” *Nonlinear Dynamics*, vol. 62, no. 3, pp. 615–621, 2010.
- [17] M. S. Baptista, “Cryptography with chaos,” *Physics Letters A*, vol. 240, no. 1–2, pp. 50–54, 1998.
- [18] G. Álvarez, F. Montoya, M. Romera, and G. Pastor, “Cryptanalysis of an ergodic chaotic cipher,” *Physics Letters A*, vol. 311, no. 2–3, pp. 172–179, 2003.
- [19] A. Kiani-B, K. Fallahi, N. Pariz, and H. Leung, “A chaotic secure communication scheme using fractional chaotic systems based on an extended fractional Kalman filter,” *Communications in Nonlinear Science and Numerical Simulation*, vol. 14, no. 3, pp. 863–879, 2009.
- [20] G. C. Wu and D. Baleanu, “Discrete fractional logistic map and its chaos,” *Nonlinear Dynamics*, 2013.
- [21] N. K. Pareek, V. Patidar, and K. K. Sud, “Cryptography using multiple one-dimensional chaotic maps,” *Communications in Nonlinear Science and Numerical Simulation*, vol. 10, no. 7, pp. 715–723, 2005.
- [22] N. K. Pareek, V. Patidar, and K. K. Sud, “Image encryption using chaotic logistic map,” *Image and Vision Computing*, vol. 24, no. 9, pp. 926–934, 2006.
- [23] G. Álvarez, F. Montoya, M. Romera, and G. Pastor, “Cryptanalysis of a discrete chaotic cryptosystem using external key,” *Physics Letters A*, vol. 319, no. 3–4, pp. 334–339, 2003.
- [24] C. Li, S. Li, M. Asim, J. Nunez, G. Alvarez, and G. Chen, “On the security defects of an image encryption scheme,” *Image and Vision Computing*, vol. 27, no. 9, pp. 1371–1381, 2009.
- [25] S. Li, X. Mou, Y. Cai, Z. Ji, and J. Zhang, “On the security of a chaotic encryption scheme: problems with computerized chaos in finite computing precision,” *Computer Physics Communications*, vol. 153, no. 1, pp. 52–58, 2003.
- [26] T. Gao and Z. Chen, “A new image encryption algorithm based on hyper-chaos,” *Physics Letters A*, vol. 372, no. 4, pp. 394–400, 2008.
- [27] H. Liu and X. Wang, “Color image encryption based on one-time keys and robust chaotic maps,” *Computers & Mathematics with Applications*, vol. 59, no. 10, pp. 3320–3327, 2010.
- [28] P. F. Verhulst, “Recherches mathématiques sur la loi d'accroissement de la population,” *Nouveaux Mémoires de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles*, vol. 18, pp. 3–38, 1845.
- [29] R. M. May, “Biological populations with nonoverlapping generations: stable points, stable cycles, and chaos,” *Science*, vol. 186, no. 4164, pp. 645–647, 1974.
- [30] K. Ikeda, “Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system,” *Optics Communications*, vol. 30, no. 2, pp. 257–261, 1979.
- [31] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, “Determining Lyapunov exponents from a time series,” *Physica D*, vol. 16, no. 3, pp. 285–317, 1985.
- [32] G. Elert, *The Chaos Hypertextbook: Measuring Chaos*, vol. 10, 2003.
- [33] K. Mahmoud, S. Datta, and J. Flint, *Frequency Domain Watermarking: An Overview*, 2003.
- [34] C. E. Shannon, “A mathematical theory of communication,” *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, pp. 3–55, 2001.
- [35] Q. Huynh-Thu and M. Ghanbari, “Scope of validity of PSNR in image/video quality assessment,” *Electronics Letters*, vol. 44, no. 13, pp. 800–801, 2008.
- [36] A. Cilardo, L. Coppolino, N. Mazzocca, and L. Romano, “Elliptic curve cryptography engineering,” *Proceedings of the IEEE*, vol. 94, no. 2, pp. 395–405, 2006.
- [37] Y. Wu, J. P. Noonan, and S. Agaian, “NPCR and UACI randomness tests for image encryption,” *Cyber Journals: Multidisciplinary Journals in Science and Technology, Journal of Selected Areas in Telecommunications*, pp. 31–38, 2011.

Research Article

Modeling of an EDLC with Fractional Transfer Functions Using Mittag-Leffler Equations

J. J. Quintana,¹ A. Ramos,² and I. Nuez¹

¹ Electronic and Automatic Engineering Department, University of Las Palmas de G.C., 35017 Las Palmas, Spain

² Process Engineering Department, University of Las Palmas de G.C., 35017 Las Palmas, Spain

Correspondence should be addressed to J. J. Quintana; jjquintana@diea.ulpgc.es

Received 31 July 2013; Revised 30 October 2013; Accepted 31 October 2013

Academic Editor: Bashir Ahmad

Copyright © 2013 J. J. Quintana et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Electrochemical double-layer capacitors (EDLC), also known as supercapacitors or ultracapacitors, are devices in which diffusion phenomena play an important role. For this reason, their modeling using integer-order differential equations does not yield satisfactory results. The higher the temporal intervals are, the more problems and errors there will be when using integer-order differential equations. In this paper, a simple model of a real capacitor formed by an ideal capacitor and two parasitic resistors, one in series and the second in parallel, is used. The proposed model is based on the ideal capacitor, adding a fractional behavior to its capacity. The transfer function obtained is simple but contains elements in fractional derivatives, which makes its resolution in the time domain difficult. The temporal response has been obtained through the Mittag-Leffler equations being adapted to any EDLC input signal. Different charge and discharge signals have been tested on the EDLC allowing modeling of this device in the charge, rest, and discharge stages. The obtained parameters are few but identify with high precision the charge, rest, and discharge processes in these devices.

1. Introduction

In recent years, the growing demand for new electrical energy storing systems has led to a remarkable development of electrochemical double-layer capacitors (EDLC). The EDLC are devices capable of storing energy and are characterized by their very rapid response during charge and discharge cycles, which allows them to provide high power and to hold a high number of charge and discharge cycles.

The complementary qualities of EDLC and batteries have allowed the generation of numerous hybrid applications for energy recovery or storage systems. Energy storage in EDLC is not supported by chemical processes. Moreover, supercapacitors have a high life cycle and do not need any maintenance. From the above it can be concluded that because of their characteristics these devices have raised great interesting expectations [1, 2].

It is a usual procedure that the dynamics of real systems are modeled by differential equations. In most cases, the differential equations are based on conventional derivatives,

yielding sufficiently accurate mathematical models. However, there are a variety of systems or phenomena in which mathematical models based on ordinary differential equations do not provide satisfactory solutions. One of these elements is the EDLC. It is this case, when the application of models based on fractional derivatives, studied in the field of fractional calculus, is very useful. These equations are called fractional differential equations.

In multiple fields of physics, there are numerous examples in which a differential equation is used as fractional modeling tool [3–5]. This is very common when considering geometry conditions with fractal dimension and distributed parameter systems [6–8]. In the field of electric power, the current through a capacitor is proportional to the noninteger-order integral for electric current [9, 10]. An electric network composed of infinite RC elements can be modeled through fractional differential equations [11, 12].

In this paper, a dynamic fractional model for EDLC is developed, based on a differential equation with fractional

derivatives. This equation has been solved by means of the Laplace transform, and in order to obtain the time domain solution used the Mittag-Leffler function $E_{\alpha,\beta}(z)$. The solution has been obtained for a unitary step input and for a generic input. This model has been applied to an EDLC for its identification in long time periods, obtaining very satisfactory results.

This paper is organized as follows: Section 2 shows a brief review of fractional calculus; in Section 3 the fractional model of the EDLC is deduced; Section 4 expounds the experimental data and the discussion; finally in Section 5 the conclusions are presented.

2. Brief Review of Fractional Calculus

In this section is exposed a brief mathematical background of the fractional calculus [3], for understanding methods, results, and conclusions presented in this paper. The following are some of the most popular functions, definitions, and properties of the fractional calculus.

2.1. Riemann-Liouville Derivative Definition. Riemann-Liouville definition expresses the fractional derivative as a time convolution integral

$${}_aD_t^\alpha f(t) = \frac{1}{\Gamma(n-\alpha)} \frac{d^n}{dt^n} \int_a^t \frac{f(\tau)}{(t-\tau)^{\alpha-n+1}} d\tau \quad (n-1 < \alpha < n), \quad (1)$$

where n is an integer and α is a real number and the fractional derivative order.

2.2. Caputo Derivative Definition. Caputo definition of a fractional derivative of a function is

$${}_a^C D_t^\alpha f(t) = \frac{1}{\Gamma(n-\alpha)} \int_a^t \frac{f^{(n)}(\tau)}{(t-\tau)^{\alpha+1-n}} d\tau \quad (n-1 \leq \alpha < n), \quad (2)$$

where n is an integer and α is a real number and the fractional derivative order.

2.3. Grunwald-Letnikov Derivative Definition. Grunwald-Letnikov definition is a numerical form of the fractional derivative

$${}_aD_t^\alpha f(t) = \lim_{h \rightarrow 0} \frac{1}{h^\alpha} \sum_{j=0}^{[(t-a)/h]} (-1)^j \binom{\alpha}{j} f(t-jh). \quad (3)$$

2.4. Laplace Transform of the Fractional Derivative. An interesting property of the fractional derivative operator is its Laplace transform

$$\begin{aligned} L[{}_0D_t^\alpha f(t)] &= \int_0^\infty e^{-st} {}_0D_t^\alpha f(t) dt \\ &= s^\alpha F(s) - \sum_{k=0}^{n-1} s^k {}_0D_t^{\alpha-k-1} f(t)|_{t=0} \quad (4) \\ &\quad (n-1 < \alpha \leq n), \end{aligned}$$

where, if initial conditions are null, it yields

$$L[{}_0D_t^\alpha f(t)] = s^\alpha F(s). \quad (5)$$

2.5. Mittag-Leffler Function. The two-parameter function of the Mittag-Leffler type $E_{\alpha,\beta}(z)$ plays a very important role in the fractional calculus and is defined as

$$E_{\alpha,\beta}(z) = \sum_{k=0}^\infty \frac{z^k}{\Gamma(\alpha k + \beta)} \quad (\alpha > 0, \beta > 0). \quad (6)$$

The exponential function e^z is a particular case of the Mittag-Leffler function

$$E_{1,1}(z) = \sum_{k=0}^\infty \frac{z^k}{\Gamma(k+1)} = e^z. \quad (7)$$

One of the main uses of the Mittag-Leffler function comes from the following Laplace transform

$$\begin{aligned} L[t^{\alpha k+\beta-1} E_{\alpha,\beta}^{(k)}(\pm at^\alpha)] &= \int_0^\infty e^{-st} t^{\alpha k+\beta-1} E_{\alpha,\beta}^{(k)}(\pm at^\alpha) dt \\ &= \frac{k! s^{\alpha-\beta}}{(s^\alpha \mp a)^{k+1}}, \quad (8) \end{aligned}$$

where k indicates the order of the derivative of the Mittag-Leffler function

$$E_{\alpha,\beta}^{(k)}(z) = \frac{d^k}{dz^k} E_{\alpha,\beta}(z). \quad (9)$$

Considering the specific case for $k = 0$ yields

$$L[t^{\beta-1} E_{\alpha,\beta}(\pm at^\alpha)] = \frac{s^{\alpha-\beta}}{s^\alpha \mp a}. \quad (10)$$

And by substituting the Mittag-Leffler function, the following is obtained:

$$L\left[t^{\beta-1} \sum_{k=0}^\infty \frac{(\pm at^\alpha)^k}{\Gamma(\alpha k + \beta)}\right] = \frac{s^{\alpha-\beta}}{s^\alpha \mp a}. \quad (11)$$

Which will be the relation used for calculating the solution in the time domain.

3. Dynamic Model of EDLC

Electric modeling of supercapacitors is an active line of investigation with numerous contributions in the last few years [13–15]. There are not yet totally satisfactory models applicable to any operating mode of the EDLC, allowing the EDLC behavior to be observed and simulated. That is the reason why models based on heuristic techniques are mainly used [14]. Models based on the physics processes occurring in EDLC are also used [15]. Most of the models obtained make use of a considerable number of variables, either of a high number of passive elements or of fractional variables models [16–18]. Either way, testing these devices during long time periods has allowed observing their behavior and their applicability to energy storage systems. Charge and discharge are rapid processes occurring in a short time interval, since high currents are used.

Due to diffusion processes resulting after the charging phase. The evolution of the processes occurring inside the EDLC generates a nonreversing thermal dissipation which makes the assessment of losses in the device possible. This process would occur in both the EDLC charge and discharge phases.

The model of EDLC has been considered as single input (current), single output (voltage) system. The electric model used comes from the model applied to ideal capacitors using integer-order differential equations. Subsequently, a model based on fractional order differential equations is proposed.

3.1. Electric Model of a Conventional Capacitor. The mathematical equation, relating the current and the voltage at the terminals of a real capacitor, can be deduced from an electric circuit made of an ideal capacitor and two parasitic resistors [9, 19].

Naming the series resistor as R_1 and the parallel one as R_2 and operating the following equation is obtained:

$$R_2 \cdot C \frac{dv(t)}{dt} + v(t) = (R_1 + R_2) \cdot i(t) + R_1 \cdot R_2 \cdot C \frac{di(t)}{dt} \quad (12)$$

with $v(t)$ being the voltage at the terminals of the EDLC and $i(t)$ its current. The transfer function or impedance in the Laplace domain will be calculated as the quotient between voltage and current

$$G(s) = \frac{V(s)}{I(s)} = R_1 \left(\frac{s}{s + (1/(R_2 \cdot C))} \right) + \frac{R_1 + R_2}{R_2 \cdot C} \left(\frac{1}{s + (1/(R_2 \cdot C))} \right). \quad (13)$$

Gathering the constants of the transfer function the following is obtained:

$$G(s) = \frac{V(s)}{I(s)} = \left(k_1 \frac{s}{s + a} + k_2 \frac{1}{s + a} \right), \quad (14)$$

where

$$k_1 = R_1, \quad k_2 = \frac{R_1 + R_2}{R_2 \cdot C}, \quad a = \frac{1}{R_2 \cdot C}. \quad (15)$$

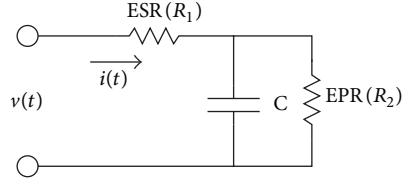


FIGURE 1: Equivalent circuit of a real capacitor.

The differential equation resulting from the transfer function can be solved for simple and complex signals using the existing simulation programs.

3.2. Electric Model of EDLC. The model obtained for real capacitor (12) fits well to the behavior of real capacitors in a wide range of frequencies. However, if applied to an EDLC, the adjustment is unsatisfactory, especially when long time periods are analyzed.

Many papers [3, 10, 20] exposed the fractional behavior of EDLC. Thus, by changing the ideal capacitor in the model in Figure 1 for a fractional one, the voltage and the current across such fractional capacitor are given by

$$C \cdot D^\alpha v(t) = i(t). \quad (16)$$

Applying this relation to the circuit of Figure 1, it gives a differential equation of the form

$${}_0 D_t^\alpha v(t) + a \cdot v(t) = k_1 \cdot {}_0 D_t^\alpha i(t) + k_2 \cdot i(t). \quad (17)$$

Applying Laplace transform (5) and assuming that the initial conditions are zero, the following is obtained:

$$G(s) = \frac{V(s)}{I(s)} = \left(k_1 \frac{s^\alpha}{s^\alpha + a} + k_2 \frac{1}{s^\alpha + a} \right). \quad (18)$$

This transfer function is similar to (14), differing only in the fractional aspect of the capacitor. In this case differential equation (17) is of fractional order, being defined by variable α . This equation can be solved for simple and complex signals input, using Mittag-Leffler functions [3, 21].

3.3. Solution to the Fractional Equation of EDLC. In this section the solution to (18) will be deduced in the time domain for a variable input over time.

3.3.1. Solution to a Unitary Step Input. Multiplying transfer function (18) by the unit step input $i(s) = 1/s$, the response of the system to this signal is obtained:

$$\begin{aligned} v(s) &= k_1 \frac{s^{\alpha-1}}{s^\alpha + a} + k_2 \frac{s^{-1}}{s^\alpha + a} \\ \implies v(t) &= k_1 L^{-1} \left[\frac{s^{\alpha-1}}{s^\alpha + a} \right] + k_2 L^{-1} \left[\frac{s^{-1}}{s^\alpha + a} \right]. \end{aligned} \quad (19)$$

Applying relation (11),

$$v(t) = k_1 \sum_{j=0}^{\infty} \frac{(-a \cdot t^\alpha)^j}{\Gamma(\alpha \cdot j + 1)} + k_2 \cdot t^\alpha \cdot \sum_{j=0}^{\infty} \frac{(-a \cdot t^\alpha)^j}{\Gamma(\alpha \cdot j + \alpha + 1)}. \quad (20)$$

The response in the time domain to a unit step input is obtained.

3.3.2. Solution to a Generic Input. The response can be extended to any input. For this purpose, the superposition of steps displaced by the sampling period of the input signal has been used. Thus, for the first input signal, the temporal response matches the response obtained by (20). So, for the first time interval,

$$\begin{aligned} v(t) = & i(0) \cdot \left(k_1 \sum_{j=0}^{\infty} \frac{(-a \cdot t^\alpha)^j}{\Gamma(\alpha \cdot j + 1)} \right. \\ & \left. + k_2 \cdot t^\alpha \cdot \sum_{j=0}^{\infty} \frac{(-a \cdot t^\alpha)^j}{\Gamma(\alpha \cdot j + \alpha + 1)} \right), \quad (21) \\ & 0 \leq t \leq T. \end{aligned}$$

As the input signal changes in the second period, its response can be obtained by superposing to the previous response a second step, delayed in time, a period T , resulting in

$$\begin{aligned} v(t) = & i(0) \cdot \left(k_1 \sum_{j=0}^{\infty} \frac{(-a \cdot t^\alpha)^j}{\Gamma(\alpha \cdot j + 1)} \right. \\ & \left. + k_2 \cdot t^\alpha \cdot \sum_{j=0}^{\infty} \frac{(-a \cdot t^\alpha)^j}{\Gamma(\alpha \cdot j + \alpha + 1)} \right) \\ & + (i(T) - i(0)) \cdot \left(k_1 \sum_{j=0}^{\infty} \frac{(-a(t-T)^\alpha)^j}{\Gamma(\alpha \cdot j + 1)} \right. \\ & \left. + k_2(t-T)^\alpha \sum_{j=0}^{\infty} \frac{(-a(t-T)^\alpha)^j}{\Gamma(\alpha \cdot j + \alpha + 1)} \right) \\ & T \leq t \leq 2T. \quad (22) \end{aligned}$$

Generalizing (22) for a generic time interval N , the following is obtained:

$$\begin{aligned} v(t) = & \sum_{k=0}^N (i(kT) - i((k-1)T)) \\ & \cdot \left[k_1 \sum_{j=0}^{\infty} \frac{(-a \cdot (t-kT)^\alpha)^j}{\Gamma(\alpha \cdot j + 1)} \right. \\ & \left. + k_2 (t-kT)^\alpha \cdot \sum_{j=0}^{\infty} \frac{(-a \cdot (t-kT)^\alpha)^j}{\Gamma(\alpha \cdot j + \alpha + 1)} \right], \\ & NT \leq t \leq (N+1)T, \quad (23) \end{aligned}$$

where the initial conditions are zero.

4. Experimental Results, Identification, and Discussion

4.1. Instruments. The experimental data have been obtained in the laboratory making a test circuit and recording the

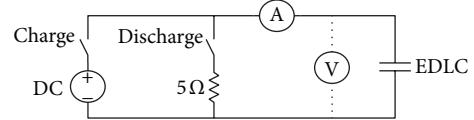


FIGURE 2: Schematic diagram of the experimental circuit.

electrical variables of the EDLC. The analyzed EDLC, manufactured by ELNA, has a capacity of 4.7 F and a voltage of 2.5 V.

Figure 2 shows the electric circuit for the charge-discharge control. A Crouzet Millenium III PLC has controlled the charge and discharge switches. Voltage and current data have been recorded through a NI USB-6009 data logger by National Instruments.

4.2. Experimental Results. Several tests have been performed to the EDLC in order to monitor voltage and current signals. Because EDLC have memory effect, just before each test they have been short-circuited for 24 hours to ensure zero initial conditions.

The tests were performed as follows. In a first phase lasting a few minutes, was charged at a constant current between 20 mA and 0.9 A until the voltage reached 2.4 V; in the second phase, the current is cut off and left to rest for 8 hours to allow the voltage stabilization, and finally in the third phase is discharged through a resistance of 5 ohm until it was fully discharged.

Although the sampling period of voltage and current has been recorded at 0.1 s; for identification purposes it has been considered a 5 s sampling time. Figure 3 shows the experimental data obtained. The evolution of voltage in the EDLC throughout the process can be appreciated in Figure 3(a), while the current supplied at Figure 3(b). They are observed perfectly the three stages described above.

4.3. Identification. The coefficients for the proposed model have been defined according to transfer function (18) and implemented using (23).

For identification the MATLAB simulation software was used. The function created has as input parameters α , a , k_1 , and k_2 . Using (23) the voltage response is generated taking all the current samples as input. The voltage is compared with the experimental data through standard deviation (24), generating the output of the function. Using this function in the command fminsearch of MATLAB, sought parameters are obtained. For implementing (23), the terms j of the sum may be truncated without causing significant errors.

The index chosen in this paper is the standard deviation defined as

$$\sigma_d = \sqrt{\frac{\sum_{i=1}^N (y_{i,\text{exp}} - y_{i,\text{cal}})^2}{N-1}}, \quad (24)$$

where N is the number of samples and subindexes which represent the values obtained experimentally and calculated according to the model.

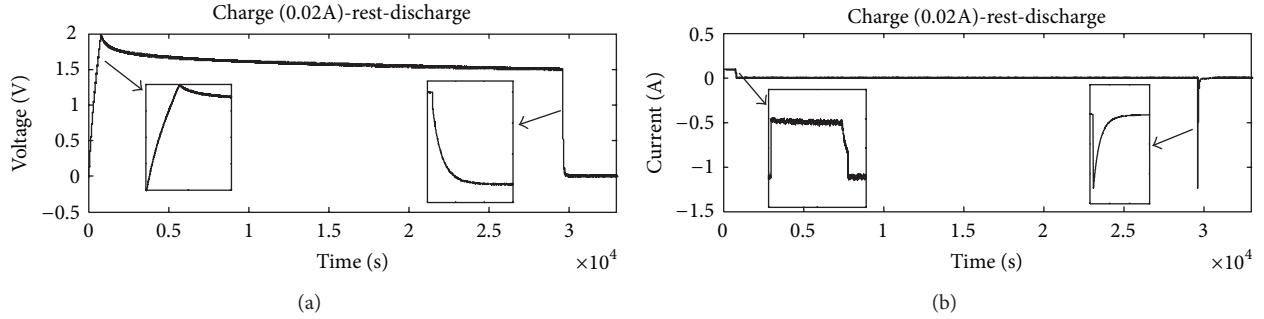


FIGURE 3: Charge-rest-discharge at 20 mA.

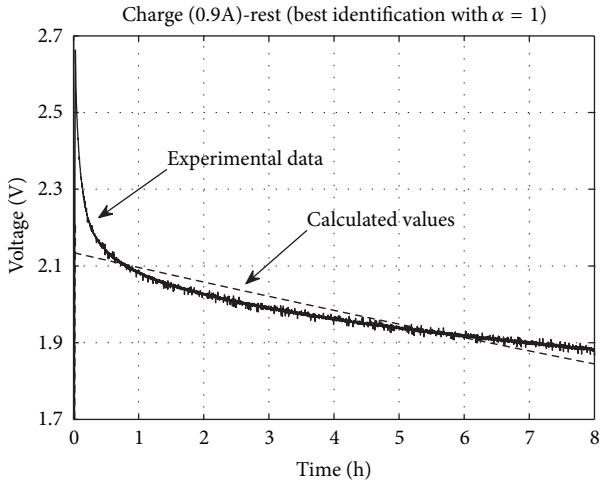


FIGURE 4: Fitting in the resting phase using integer derivatives.

TABLE 1: Charge + rest identification parameters.

Test	I (A)	α	a	k_1	k_2	σ_d
1	0.02 A	0.941	$5.974E - 08$	0.239	0.202	$2.079E - 02$
2	0.1 A	0.952	$8.671E - 08$	0.261	0.187	$1.297E - 02$
3	0.5 A	0.950	$1.02E - 07$	0.267	0.170	$1.354E - 02$
4	0.5 A	0.950	$1.020E - 07$	0.250	0.169	$1.338E - 02$
5	0.9 A	0.953	$9.917E - 08$	0.228	0.164	$1.096E - 02$

The EDLC identification has been carried out calculating the parameters α , a , k_1 , and k_2 in (18) and has consisted of two parts: first, from the beginning until just before discharged and second, from the discharge onwards. This is because the internal processes in the EDLC are different [22, 23].

The data obtained for the charge and rest phases are shown in Table 1.

Considering relation (15) between the parameters calculated with the resistors and the capacity in Figure 1, the parameters obtained for the charge are in Table 2.

In order to compare the adjustment with traditional model (14) in which the fractional index is an integer, Table 3 and Figure 4 show the best fit to the data of test number 5 setting $\alpha = 1$.

TABLE 2: Charge + rest electrical parameters.

Test	I (A)	α	C	R_1	R_2
1	0.02 A	0.941	4.95	0.239	$3.382E + 06$
2	0.1 A	0.952	5.35	0.261	$2.155E + 06$
3	0.5 A	0.950	5.89	0.267	$1.660E + 06$
4	0.5 A	0.950	5.91	0.250	$1.660E + 06$
5	0.9 A	0.953	6.08	0.228	$1.658E + 06$

TABLE 3: Parameters calculated for test number 5 setting $\alpha = 1$.

Test	α	C	R_1	R_2
1	1.000	9.00	1.262	$2.191E + 04$

TABLE 4: Discharge identification parameters.

Test	α	a	k_1	k_2	σ_d
1	0.975	$9.600E - 08$	0.235	0.187	$1.391E - 02$
2	0.981	$9.282E - 08$	0.244	0.185	$1.281E - 02$
3	0.978	$8.427E - 08$	0.267	0.192	$2.068E - 02$
4	0.976	$1.556E - 07$	0.251	0.202	$2.338E - 02$
5	0.982	$1.032E - 07$	0.240	0.171	$1.285E - 02$

TABLE 5: Discharge electrical parameters.

Test	α	C	R_1	R_2
1	0.975	5.34	0.235	$1.950E + 06$
2	0.981	5.40	0.244	$1.996E + 06$
3	0.978	5.21	0.267	$2.278E + 06$
4	0.976	4.95	0.251	$1.297E + 06$
5	0.982	5.85	0.240	$1.656E + 06$

Operating in a similar way, the parameters for the discharge through the 5Ω resistor are in Table 4.

Of which the following electric parameters are obtained in Table 5.

The following figures show graphically the fit of the proposed model with experimental data. Figure 5 shows that the model has a good precision all over the time range tested. Figures 6 and 8 display the model and the experimental data in the charging and discharging phases. Figures 4 and 7 show the fit when using transfer functions with integer and

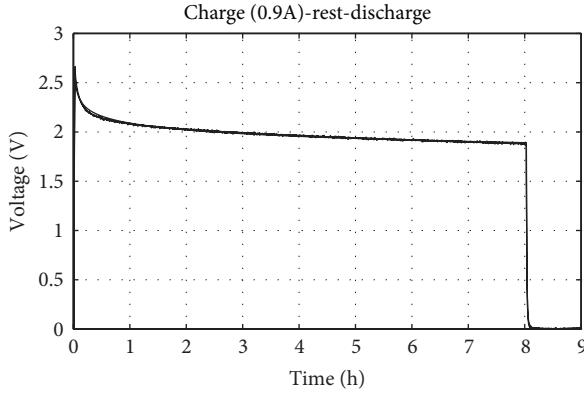


FIGURE 5: Experimental data of test number 5 versus proposed model.

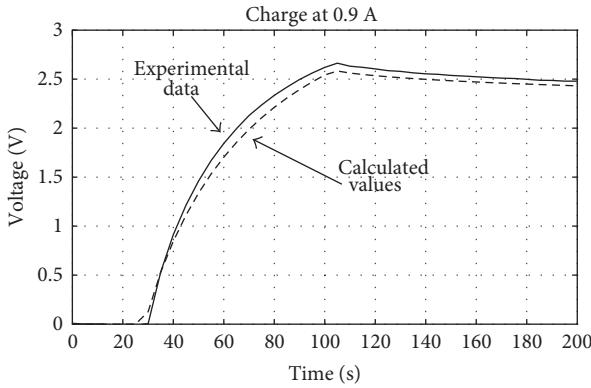


FIGURE 6: Detail of fitting in charging phase.

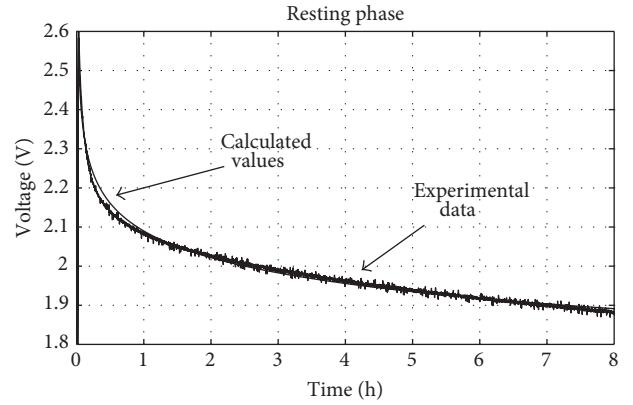


FIGURE 7: Fitting in the resting phase using the proposed model.

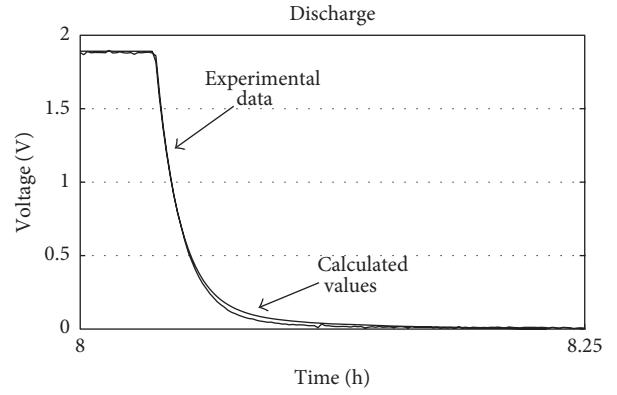


FIGURE 8: Fitting in the discharging phase.

fractional indices, showing that the fit with the traditional model gives unsatisfactory results. In summary, it can be observed that the voltage deduced through the fractional model adjusts very close to the real data even in long time periods.

4.4. Discussion. The self-discharge adjustment of EDLC after charging and after long time resting phases is a complex phenomenon which is not easily applicable to modeling through integer-order transfer functions. In this work the EDLC has been modeled making use of the fractional derivative. Taking the electric circuit in Figure 1 as starting point and assuming that the capacitor has a fractional index α , fractional transfer function (18) has been deduced. This function has been solved in the time domain using Mittag-Leffler functions and implemented in a simulation program through (23).

First of all, considering Figure 5, it can be observed that the proposed transfer function adjusts well to the real data, modeling satisfactorily the self-discharge phenomenon.

Unlike the traditional capacitor model, which is defined by the value of two resistors and the capacity, the proposed model adds a new parameter, that is, the fractional index α of the fractional capacitor. This new parameter models to a greater extent the self-discharging phenomenon.

Although constants in Tables 1 and 4 have been taken for the identification of equations, constants in Tables 2 and 5 will provide the data related to the value of the electric elements.

In all cases the adjustment has been very precise, with the typical deviation being less than 0.03.

The constancy of R_1 , R_2 , and C values for both charge and discharge is especially noteworthy.

At last, the most significant difference in terms of charge and discharge is given by the value of the fractional index α , which is around 0.95 in the charge and around 0.98 in the discharge.

5. Conclusions

This work has consisted of EDLC modeling by implementing a simple fractional model allowing a very satisfactory response for long time periods. Traditional models based on transfer functions with integer coefficients yield good results in those cases when there is no self-discharge phenomenon. Applying fractional mathematics implies an additional parameter to the already necessary parameters in order to define the basic model shown in Figure 1, which is the fractional index α . According to the tests this index has different values for the charge and the discharge, with those

values being apparently independent of the EDLC charging current.

Thanks to the proposed transfer function and the Mittag-Leffler functions, it is possible to estimate the voltage at the EDLC terminals after a long time, taking into account the charging current and the way it has been charged.

References

- [1] A. Burke, "Ultracapacitors: why, how, and where is the technology," *Journal of Power Sources*, vol. 91, no. 1, pp. 37–50, 2000.
- [2] B. Conway, *Electrochemical Supercapacitors. Scientific Fundamentals and Technological Applications*, Kluwer Academic, New York, NY, USA, 1999.
- [3] I. Podlubny, *Fractional Differential Equations*, vol. 198 of *Mathematics in Science and Engineering*, Academic Press, San Diego, Calif USA, 1999.
- [4] K. B. Oldham and J. Spanier, *The Fractional Calculus*, vol. 1047, Academic Press, New York, NY, USA, 1974.
- [5] K. S. Miller and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*, John Wiley & Sons, New York, NY, USA, 1993.
- [6] A. Oustaloup, A. Ballouk, and B. Bansard, "Partial differential equations and non integer derivation," in *Proceedings of the International Conference on Systems, Man and Cybernetics*, pp. 166–173, October 1993.
- [7] B. Mbodje and G. Montseny, "Boundary fractional derivative control of the wave equation," *IEEE Transactions on Automatic Control*, vol. 40, no. 2, pp. 378–382, 1995.
- [8] G. Montseny, J. Audounet, and B. Mbodje, "Optimal model of fractional integrators and application to systems with fading memory," in *Proceedings of the International Conference on Systems, Man and Cybernetics*, pp. 65–70, October 1993.
- [9] S. Westerlund and L. Ekstam, "Capacitor theory," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 1, no. 5, pp. 826–839, 1994.
- [10] N. Bano and R. A. Hashmi, "Electrical studies of leaves over wide frequency range," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 3, no. 2, pp. 229–232, 1996.
- [11] G. E. Carlson and C. A. Halijak, "Simulation of the fractional derivate operator and the fractional integral operator \sqrt{s} ," in *Proceedings of the Central State Simulation Council Meeting*, vol. 45, no. 7, Kansas State University, May 1961.
- [12] T. C. Haba, M. Martos, G. Ablart, and P. Bidan, "Composants électroniques à impédance fractionnaire," in *Proceedings of the ESAIM conference on Fractional Differential Systems: Models, Methods and Applications*, vol. 5, pp. 99–109, 1998.
- [13] J. J. Quintana, A. Ramos, and I. Nuez, "Identification of the fractional impedance of ultracapacitors," in *Proceedings of the IFAC Workshop (FDA'06)*, Oporto, Portugal, 2006.
- [14] L. Zubieta and R. Bonert, "Characterization of double-layer capacitors for power electronics applications," *IEEE Transactions on Industry Applications*, vol. 36, no. 1, pp. 199–205, 2000.
- [15] N. Bertrand, J. Sabatier, O. Briat, and J.-M. Vinassa, "Embedded fractional nonlinear supercapacitor model and its parametric estimation method," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 12, pp. 3991–4000, 2010.
- [16] R. Martín, J. J. Quintana, A. Ramos, and I. de la Nuez, "Modeling of electrochemical double layer capacitors by means of fractional impedance," *Journal of Computational and Nonlinear Dynamics*, vol. 3, no. 2, Article ID 021303, pp. 21303–21400, 2008.
- [17] V. Musolino, L. Piegari, and E. Tironi, "New full-frequency-range supercapacitor model with easy identification procedure," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 1, pp. 112–120, 2013.
- [18] Y. Wang, T. Hartley, C. Lorenzo, and J. Adams, *Modeling Ultracapacitors as Fractional-Order Systems*, New Trends in Nanotechnology and Fractional Calculus Applications, 2010.
- [19] R. L. Spyker, "Classical equivalent circuit parameters for a double-layer capacitor," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, no. 3, pp. 829–836, 2000.
- [20] A. Dzieliński, G. Sarwas, and D. Sierociuk, "Time domain validation of ultracapacitor fractional order model," in *Proceedings of the 49th IEEE Conference on Decision and Control (CDC '10)*, pp. 3730–3735, December 2010.
- [21] T. J. Freeborn, B. Maundy, and A. S. Elwakil, "Measurement of supercapacitor fractional-order model parameters from voltage-excited step response," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 3, no. 3, pp. 416–424, 2013.
- [22] R. de Levie, "On porous electrodes in electrolyte solutions. I. Capacitance effects," *Electrochimica Acta*, vol. 8, no. 10, pp. 751–780, 1963.
- [23] M. Kaus, J. Kowal, and D. U. Sauer, "Modelling the effects of charge redistribution during self-discharge of supercapacitors," *Electrochimica Acta*, vol. 55, no. 25, pp. 7516–7523, 2010.

Research Article

Dynamic Analysis of a Two-Language Competitive Model with Control Strategies

Lin-Fei Nie,^{1,2} Zhi-Dong Teng,¹ Juan J. Nieto,^{3,4} and Il Hyo Jung²

¹ College of Mathematics and Systems Science, Xinjiang University, Urumqi 830046, China

² Department of Mathematics, Pusan National University, Busan 609-735, Republic of Korea

³ Departamento de Análisis Matemático and Instituto de Matemáticas, Universidad de Santiago de Compostela, 15782 Santiago de Compostela, Spain

⁴ Department of Mathematics, Faculty of Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Correspondence should be addressed to Lin-Fei Nie; lfnie@163.com

Received 2 August 2013; Accepted 7 October 2013

Academic Editor: Bashir Ahmad

Copyright © 2013 Lin-Fei Nie et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The dynamic behavior of a two-language competitive model is analyzed systematically in this paper. By the linearization and the Bendixson-Dulac theorem on dynamical system, some sufficient conditions on the globally asymptotical stability of the trivial equilibria and the existence and the stability of the positive equilibrium of this model are presented. Nextly, in order to protect the endangered language, an optimal control problem relative to this model is explored. We derive some necessary conditions to solve the optimal control problem and present some numerical simulations using a Runge-Kutta fourth-order method. Finally, the languages competitive model is extended to this model assessing the impact of state-dependent pulse control strategy. Using the Poincaré map, differential inequality, and method of qualitative analysis, we prove the existence and stability of positive order-1 periodic solution for this control model. Numerical simulations are carried out to illustrate the main results and the feasibility of state-dependent impulsive control strategy.

1. Introduction

The diversity of cultures is the greatest charm of the human civilization, and languages are the most important carrier for culture. In the past decades, with the progress of the globalization, local tongues are increasingly replaced by hegemonic languages [1], this trend that has been investigated from multiple points of view, including that of physics. We refer to some of them in [2–11] and the references therein.

Perhaps the earliest and simplest mathematical model for languages shift was developed by Abrams and Strogatz [2], Patriarca et al. [12, 13], and Stauffer et al. [14]. They considered a stable population in which two languages with different statuses competed for speakers and predicted that one of the languages would inevitably die out. The theoretical results were successfully fitted to historical data on the competition between Scottish Gaelic and English, Welsh and English, and Quechua and Spanish, among other language pairings [2]. However, there was no mention of the fact that the possibility of bilingual individuals might exist, a possibility that is of

course realized in numerous multilingual societies. This is widely exists in all parts of the world. For example, in Spain, Castilian Spanish is the official language throughout the state, but in certain regions it is coofficial with another language (mainly Galician, Basque, Catalan, or Valencian); individual bilingualism is common in communities with more than one official language.

Recently, Mira et al. [8, 15, 16] proposed a modified Abrams-Strogatz model that allows for bilingual as well as monolingual speakers of the competing languages and that includes a parameter that represents the ease of bilingualism. The model is accordingly described by the following differential equations:

$$\begin{aligned} \frac{dx}{dt} &= yP_{YX} + bP_{BX} - x(P_{XY} + P_{XB}), \\ \frac{dy}{dt} &= xP_{XY} + bP_{BY} - y(P_{YX} + P_{YB}), \\ \frac{db}{dt} &= xP_{XB} + yP_{YB} - b(P_{BY} + P_{BX}), \end{aligned} \quad (1)$$

where capital letters X and Y denote the two languages spoken in population; the uppercased letter B denotes the group of bilingual speakers; and the lowercased letters x , y , and b (with $x + y + b = 1$) refer to the fraction of speakers of each of the languages in population and the fraction of bilingual speakers, respectively; parameter P_{XY} denotes the probability of a monolingual speaker of language X being replaced in the population by a monolingual speaker of language Y , with analogous notation for the other parameters P_{YX} , P_{XB} , P_{BX} , P_{YB} , and P_{BY} . The probability of a monolingual person being replaced by mono- or bilingual speaker of the other language is assumed to be proportional both to the status of the second language, that is, the social and/or economic advantages it offers, and to a power of the proportion of population that speaks it. Thus, denoting by s the relative status of language X and by $1 - s$ that of language Y ,

$$\begin{aligned} P_{XB} &= ck(1-s)(1-x)^\alpha, \\ P_{BX} &= P_{YX} = c(1-k)s(1-y)^\alpha, \\ P_{YB} &= cks(1-y)^\alpha, \\ P_{BY} &= P_{XY} = c(1-k)(1-s)(1-x)^\alpha, \end{aligned} \quad (2)$$

where c is a normalization factor related to the time scale, α is the power parameter, and k is the probability that the disappearance of a monolingual speaker of language X (resp., Y) will be compensated for by the appearance of a bilingual rather than by a monolingual speaker of language Y (resp., X). On basis of detailed analysis and extensive calculations, authors showed that both languages may coexist and survive in the long term. They pointed out that it is possible only if the competing languages are sufficiently similar, in which case its occurrence is favored by both similarity and status symmetry.

It is generally known that the disappearance of race language will bring the disappearance of race culture, even the whole disappearance of the corresponding race. The protection of endangered language has been concerned increasingly interdisciplinary in different contexts.

Very recently, the dynamical model with optimal control strategies has become a major topic in mathematical biology (see [17–23] and the references therein). Particularly, Joshi, in [18], proposed an HIV immunology model with optimal drug treatment strategies, and the existence and uniqueness results for the optimal control pair are established. Jung et al. [24] proposed a two-strain tuberculosis model with two control terms, and the optimal controls are characterized in terms of the optimality system. In addition, the state-dependent impulsive feedback control measure is also applied widely to the control of spread of infectious disease due to its economic, high-efficiency, and feasibility nature; see [25, 26] and the references therein.

Motivated by these facts, in this paper, the dynamic behavior of two-language competitive model (1) with (2) is analyzed systematically in Section 2. A set of necessary conditions that an optimal control and state must satisfy, are derived in Section 3. In Section 4, we extend model (1)

with state-dependent pulse control measure. Some sufficient conditions are presented in this section for the existence and stability of positive periodic solution. Some concluding remarks are presented in Section 5.

2. Qualitative Analysis for Model (1)

Inserting the parameters formula (2) into model (1) and taking into account $x + y + b = 1$, we get the following reduced model (here, we assume that $\alpha = 1$ throughout the rest of this paper):

$$\begin{aligned} \frac{dx}{dt} &= cs(1-k)(1-x)(1-y) - c(1-s)x(1-x) := f_1(x, y), \\ \frac{dy}{dt} &= c(1-k)(1-s)(1-x)(1-y) - csy(1-y) := f_2(x, y). \end{aligned} \quad (3)$$

By the biological background of model (3), we only consider model (3) in the biological meaning region $\Omega = \{(x, y) : x \geq 0, y \geq 0, x + y \leq 1\}$.

The following theorem is on the nonnegativity of solution of model (3).

Theorem 1. *The region Ω is positively invariant for model (3).*

The proof of Theorem 1 is simple; we, therefore, omit it here.

Now, we discussed the existence and stability of equilibria for model (3). The isocline corresponding to $dx/dt = 0$ is the line $s(1-k)(1-y) = (1-s)x$. The isocline corresponding to $dy/dt = 0$ is the line $(1-k)(1-s)(1-x) = sy$. Note that two isoclines do not have positive intersection point for $s < (1-s)(1-k)$ or $(1-s) < s(1-k)$ and have only one positive intersection point for $1-k < (1-s)/s < 1/(1-k)$. Namely, model (3) has only two trivial equilibria $(1, 0)$ and $(0, 1)$ for $s < (1-s)(1-k)$ or $(1-s) < s(1-k)$ and has two trivial equilibria $(1, 0)$; $(0, 1)$ and a positive equilibrium $E(x^*, y^*)$ for $1-k < (1-s)/s < 1/(1-k)$, where

$$\begin{aligned} x^* &= \frac{(1-k)s - (1-s)(1-k)^2}{1-s - (1-s)(1-k)^2}, \\ y^* &= \frac{(1-k)(1-s) - (1-k)^2s}{s(1-(1-k)^2)}. \end{aligned} \quad (4)$$

The locally asymptotical stabilities of equilibria are determined by the eigenvalues of Jacobian matrixes of the linearization of model (3) around equilibria. It is easy to calculate that

$$\begin{aligned} \frac{\partial f_1(x, y)}{\partial x} &= -cs(1-k)(1-y) - c(1-s)(1-x) + c(1-s)x, \\ &= -cs(1-k)(1-y) - c(1-s)(1-x) + c(1-s)x, \end{aligned}$$

$$\begin{aligned}\frac{\partial f_1(x, y)}{\partial y} &= -cs(1-k)(1-x), \\ \frac{\partial f_2(x, y)}{\partial x} &= -c(1-k)(1-s)(1-y), \\ \frac{\partial f_2(x, y)}{\partial y} &= -c(1-k)(1-s)(1-x) - cs(1-y) + csy.\end{aligned}\tag{5}$$

Computing Jacobian matrixes from model (3) around $(1, 0)$, $(0, 1)$, and (x^*, y^*) , we have that

$$\begin{aligned}J_{(1,0)} &= \begin{pmatrix} c(1-s) - cs(1-k) & 0 \\ -c(1-k)(1-s) & -cs \end{pmatrix}, \\ J_{(0,1)} &= \begin{pmatrix} -c(1-s) & -cs(1-k) \\ 0 & cs - c(1-k)(1-s) \end{pmatrix},\end{aligned}\tag{6}$$

$$J_{(x^*, y^*)} = \begin{pmatrix} -c(1-s)(1-x^*) & -cs(1-k)(1-x^*) \\ -c(1-k)(1-s)(1-y^*) & -cs(1-y^*) \end{pmatrix}.\tag{7}$$

From (6), since $(1-s) < s(1-k)$ implies that $(1-s)(1-k) < s$, it follows that the two eigenvalues of $J_{(1,0)}$ are negative and that one eigenvalue is positive and another one is negative for $J_{(0,1)}$ when $(1-s) < s(1-k)$. Therefore, model (3) has a locally asymptotically stable node $(1, 0)$ and a saddle point $(0, 1)$ for $(1-s) < s(1-k)$. Furthermore, since there is not equilibrium in interior of Ω for $(1-s) < s(1-k)$, model (3) has no periodic orbit lying entirely in the interior of Ω . Therefore, $(1, 0)$ is globally asymptotically stable in the interior of Ω for this case. Similarly, since $s < (1-s)(1-k)$ implies that $s(1-k) < (1-s)$, we also have that $(0, 1)$ is globally asymptotically stable in the interior of Ω and $(1, 0)$ is a saddle point for model (3) when $s < (1-s)(1-k)$.

Consider the eigenvalues of $J_{(x^*, y^*)}$; let $T = \text{Tr } J_{(x^*, y^*)}$ and $D = \text{Det } J_{(x^*, y^*)}$ be the trace and determinant of $J_{(x^*, y^*)}$, respectively. The eigenvalues of $J_{(x^*, y^*)}$ are provided by the following characteristic equation:

$$\lambda^2 - T\lambda + D = 0.\tag{8}$$

Note that when $(1-k) < (1-s)/s < 1/(1-k)$,

$$\begin{aligned}T &= -c(1-s)(1-x^*) - cs(1-y^*) < 0, \\ D &= c^2s(1-s)(1-x^*)(1-y^*) \\ &\quad - c^2s(1-s)(1-k)^2(1-x^*)(1-y^*) > 0.\end{aligned}\tag{9}$$

Hence, both eigenvalues have negative real part, and, hence, (x^*, y^*) is locally asymptotically stable when $(1-k) < (1-s)/s < 1/(1-k)$. Furthermore, when we mention that model (3) has no periodic orbits, refer to Lemma 3.2 of [16].

To summarize the above discussion, we give some sufficient conditions for the existence and asymptotical stability of equilibria for model (3).

Theorem 2. For any $k, s \in (0, 1)$, one of the following statements is valid.

- (a) If $(1-s) < s(1-k)$, then model (3) has only two trivial equilibria $(1, 0)$ and $(0, 1)$, where $(0, 1)$ is a saddle point and $(1, 0)$ is globally asymptotically stable in the interior of Ω .
- (b) If $s < (1-s)(1-k)$, then model (3) has only two trivial equilibria $(1, 0)$ and $(0, 1)$, where $(1, 0)$ is a saddle point and $(0, 1)$ is globally asymptotically stable in the interior of Ω .
- (c) If $(1-k) < (1-s)/s < 1/(1-k)$, then model (3) has two saddle points $(1, 0)$, $(0, 1)$ and a globally asymptotically stable positive equilibrium (x^*, y^*) .

Remark 3. From Theorem 2 and the equivalence of models (1) and (3), we obtain that model (1) has a globally asymptotically stable node $(1, 0, 0)$ and a saddle point $(0, 1, 0)$ for $(1-s) < s(1-k)$ and has a saddle point $(1, 0, 0)$ and a globally asymptotically stable node $(0, 1, 0)$ for $s < (1-s)(1-k)$. Furthermore, if $(1-k) < (1-s)/s < 1/(1-k)$, model (1) has a coexistent equilibrium $(x^*, y^*, 1-x^*-y^*)$ which is globally asymptotically stable.

Finally, we fix all parameters including c , k , s , and α and carry out numerical investigations to confirm our main results obtained in this section. Firstly, we choose $k = 0.5$, $s = 0.7$, $c = 0.5$, and $\alpha = 1$; it is easy to calculate that $1-s = 0.3 < s(1-k) = 0.7 \times 0.5 = 0.35$. So, from the first conclusion of Theorem 2, we know that model (1) has only two trivial equilibria $(1, 0, 0)$ and $(0, 1, 0)$, $(1, 0, 0)$ is a globally asymptotically stable node, and $(0, 1, 0)$ is a saddle point, which is shown in Figure 1(a). It is clear that language X is permanent and language Y will fade away in this case. Similar conclusion can be obtained from Figure 1(b) with parameters $k = 0.3$, $s = 0.4$, $c = 0.5$, and $\alpha = 1$. However, if we choose $k = s = c = 0.5$ and $\alpha = 1$, then languages X and Y are coexistent and tend to positive equilibrium as shown in Figure 1(c).

3. The Protection of an Endangered Language by a Continuous Control Strategy

With the development of human civilization, people have taken effective measures to prevent the disappearance of language. In this section and the following, therefore, model (1) with parameters (2) is extended to assess the impact of control measures. And, in general, control strategies are divided into two main types: continuous control and pulse control. We, firstly, consider how a continuous control measure affects the dynamical behavior of model (1) in this section.

Through discussion of Theorem 2 in Section 2, we know that (i) language X is permanent and language Y and bilingual speakers B are extinct for $(1-s) < s(1-k)$ and that (ii) language Y is permanent and language X and bilingual speakers B are extinct for $s < (1-s)(1-k)$. Considering the similarities of the two cases, we only need to consider case (i); that is, $(1-s) < s(1-k)$.

3.1. Protecting of Language Y with Continuous Control for $(1-s) < s(1-k)$. We consider model (1) with parameters (2)

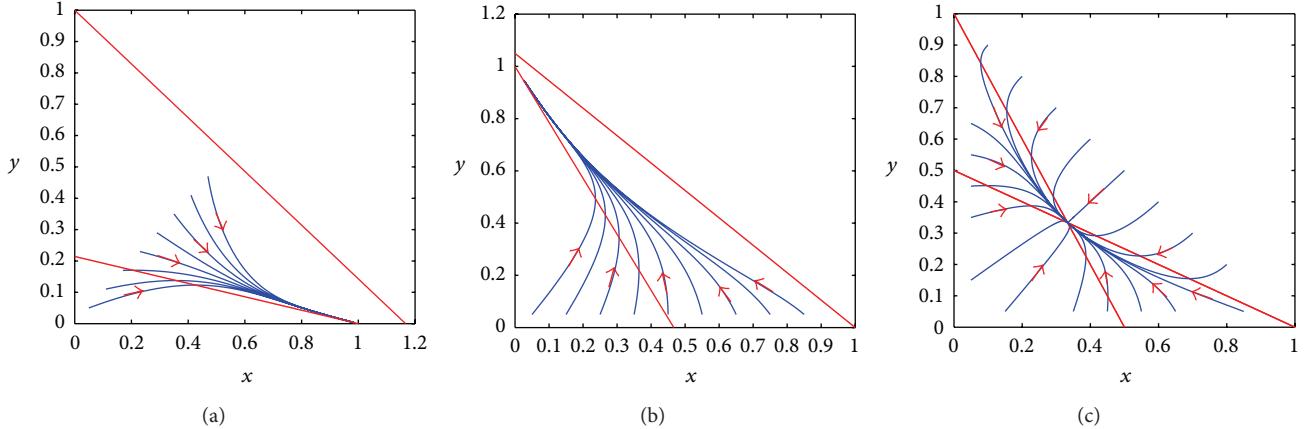


FIGURE 1: The asymptotical stability of equilibria of model (1) with $c = 0.5$ and $\alpha = 1$: (a) the stability of trivial equilibrium $(1, 0, 0)$ with $k = 0.5$ and $s = 0.7$; (b) the stability of trivial equilibrium $(0, 1, 0)$ with $k = 0.3$ and $s = 0.7$; and (c) the stability of positive equilibrium $(x^*, y^*, 1 - x^* - y^*)$ with $k = 0.5$ and $s = 0.5$.

and continuous control measure. The control system is modeled by the following differential equations:

$$\begin{aligned}\frac{dx}{dt} &= cs(1-k)(y+b)(1-y) - c(1-s)x(1-x) - c\mu x, \\ \frac{dy}{dt} &= c(1-s)(1-k)(x+b)(1-x) - csy(1-y), \\ \frac{db}{dt} &= ck(1-s)x(1-x) + cksy(1-y) \\ &\quad - c(1-k)b[(1-s)(1-x) + s(1-y)] + c\mu x,\end{aligned}\tag{10}$$

where μ is a controlled variable, which means that the fraction of language X becomes the bilingual B per unit of time. Note that $x + y + b = 1$; model (10) can be written as follows:

$$\begin{aligned}\frac{dx}{dt} &= cs(1-k)(1-x)(1-y) - c(1-s)x(1-x) - c\mu x, \\ \frac{dy}{dt} &= c(1-s)(1-k)(1-y)(1-x) - csy(1-y).\end{aligned}\tag{11}$$

We denote the right-side of model (11) by $f(x, y)$ and $g(x, y)$, respectively. We, here, discuss the existence and asymptotical stability of positive equilibrium of model (11).

From $g(x, y) = 0$, it is easily shown that

$$y_\mu = y(x) = \frac{(1-k)(1-s)(1-x)}{s}. \quad (12)$$

Substituting it in $f(x, y)$, it follows that $f(1, y(1)) = -\mu < 0$ and

$$f(0, y(0)) = cs(1-k) \left(1 - \frac{(1-k)(1-s)}{s}\right) > 0 \quad (13)$$

due to conditions $(1-s) < s(1-k)$ and $k, s \in (0, 1)$. Hence, there is at least one $x_\mu^* \in (0, 1)$ such that $f(x_\mu^*, y(x_\mu^*)) = 0$. Furthermore, it is also easy to calculate that

$$\begin{aligned} \frac{df(x, y(x))}{dx} &= c(1-k)[-s + 2(1-k)(1-s)(1-x)] \\ &\quad + c(1-s)(-1+2x) - c\mu, \end{aligned} \tag{14}$$

$$\frac{d^2 f(x, y(x))}{dx^2} = 2(1-s) \left[1 - (1-k)^2 \right] > 0 \quad (15)$$

for all $x \in [0, 1]$. This together with (14) gives that

$$\frac{df(x, y(x))}{dx} < \frac{df(1, y(1))}{dx} = (1-s) - s(1-k) - \mu < 0 \quad (16)$$

for all $x \in (0, 1)$, where inequality $(1-s) < s(1-k)$ is used. Hence, x_μ^* is unique. That is, model (II) has a unique positive equilibrium $E_*(x_\mu^*, y_\mu^*)$ in the interior of Ω .

Similar to the discussion of Theorem 2, we can get the locally asymptotical stability of equilibrium $E_\mu(x_\mu^*, y_\mu^*)$. Next, we discuss the global behavior of equilibrium $E_\mu(x_\mu^*, y_\mu^*)$. Let $B(x, y) = (1-x)^{-1}(1-y)^{-1}$; we have

$$\frac{\partial(Bf)}{\partial x} + \frac{\partial(Bg)}{\partial y} = -\frac{c(1-s)}{1-y} - \frac{cs}{1-x} - \frac{c\mu}{(1-x)^2(1-y)} < 0. \quad (17)$$

So model (11) has no closed orbit lying entirely in the interior of Ω . We, therefore, can show from the above discussion that equilibrium $E_\mu(x_\mu^*, y_\mu^*)$ is globally asymptotically stable in the interior of Ω . The result can be written by the following theorem.

Theorem 4. For any $\mu \in (0, 1)$, if $(1 - s) < s(1 - k)$, then model (11) has a globally asymptotically stable positive equilibrium $E_\mu(x^*, y^*)$ in the interior of Ω .

Remark 5. From Theorem 2, it follows that language X is permanent and that language Y and bilingual speakers B will eventually disappear for $(1 - s) < s(1 - k)$ in model (1). However, if we introduce a control variable μ (no matter how small it is) in model (1), languages X and Y are coexistent and tend to positive equilibrium $(x_\mu^*, y_\mu^*, 1 - x_\mu^* - y_\mu^*)$. The coexistent state, of course, depends upon the controlled strength μ . This implies that μ is a sensitive controlled parameter for the protection of endangered language.

3.2. Analysis of Optimal Control. Optimal control techniques are of great use in developing the optimal strategies to protect endangered civilization. To solve the challenges of obtaining an optimal control measure, we use optimal control theory; for more details, see Lenhart and Workman [27]. In model (10), for the optimal control problem, we consider a control variable $\mu(t) \in U_{ad}$; here, $U_{ad} = \{\mu : \mu(t) \text{ is measurable, } 0 \leq \mu(t) \leq 0.9 \text{ for all } t \in [0, t_{final}]\}$ indicates an admissible control. In this optimal problem, we assume a restriction on the control variable $\mu(t)$ such that $0 \leq \mu(t) \leq 0.9$, because conversion of all of language X at one time is impossible. In case of no control, the fraction of language X increases while the fractions of language Y and bilingual B die out. Therefore, the biological meaning of an optimal control in this problem is that the adequate levels for the fractions of language X and bilingual B are built.

Now, we consider an optimal control problem to maximize the objective functional

$$J(\mu) = \int_0^{t_{final}} \left(C_1 b - \frac{1}{2} C_2 \mu^2 \right) dt \quad (18)$$

subject to model (10). The first term represents the benefit of bilingual B , and the other term is systemic cost of control measure. The positive constants C_1 and C_2 balance the size of the terms B and μ . Our goal is maximizing the fraction of bilingual B and minimizing the systemic cost to the control measure. This seeks an optimal control μ^* such that

$$J(\mu^*) = \max \{J(\mu) : \mu \in U_{ad}\} \quad (19)$$

subject to model (10). It is obvious that the integrand of objective functional J is a convex function of control variable μ and that state model satisfies the Lipschitz property with respect to the state since state solutions are bounded. The existence of an optimal control follows [28].

The necessary conditions that an optimal must satisfy come from the Pontryagin's Maximum Principle in [28]. This principle converts optimal control problem (10) and (18) into a problem of maximizing pointwise a Hamiltonian H with respect to μ as follows:

$$\begin{aligned} H = C_1 b - \frac{1}{2} C_2 \mu^2 + \lambda_1 [& cs(1 - k)(y + b)(1 - y) \\ & - c(1 - s)x(1 - x) - c\mu x] \\ & + \lambda_2 [c(1 - s)(1 - k)(x + b)(1 - x) - csy(1 - y)] \end{aligned}$$

$$\begin{aligned} & + \lambda_3 \{ ck(1 - s)x(1 - x) + cksy(1 - y) - c(1 - k)b \\ & \times [(1 - s)(1 - x) + s(1 - y)] + c\mu x \}, \end{aligned} \quad (20)$$

where λ_1 , λ_2 , and λ_3 are adjoint variables.

In the following theorem, we derive the necessary conditions for the optimal control problem.

Theorem 6. Let (x^*, y^*, b^*) be an optimal state solution with associated optimal control variable μ^* for the maximized object functional $J(\mu)$ subject to control model (10). Then, for model (10), there exist adjoint variables λ_1 , λ_2 , and λ_3 such that

$$\begin{aligned} \frac{d\lambda_1}{dt} = \lambda_1 [& c(1 - s)(1 - x^*) - c(1 - s)x^* + c\mu^*] \\ & + \lambda_2 [c(1 - s)(1 - k)(x^* + b^*) - c(1 - s) \\ & \times (1 - k)(1 - x^*)] \\ & + \lambda_3 [ck(1 - s)x^* - ck(1 - s)(1 - x^*) \\ & - c(1 - k)(1 - s)b^* - c\mu^*], \end{aligned} \quad (21)$$

$$\begin{aligned} \frac{d\lambda_2}{dt} = \lambda_1 [& cs(1 - k)(y^* + b^*) - cs(1 - k)(1 - y^*)] \\ & + \lambda_2 [cs(1 - y^*) - csy^*] \\ & + \lambda_3 [cksy^* - cks(1 - y^*) - cs(1 - k)b^*], \end{aligned} \quad (22)$$

$$\begin{aligned} \frac{d\lambda_3}{dt} = -C_1 - \lambda_1 cs(1 - k)(1 - y^*) - \lambda_2 c(1 - s) \\ \times (1 - k)(1 - x^*) \\ & + \lambda_3 c(1 - k)[(1 - s)(1 - x^*) + s(1 - y^*)] \end{aligned} \quad (23)$$

with transversality conditions $\lambda_i(t_{final}) = 0$, $i = 1, 2, 3$. Furthermore, the optimal control is given by

$$\mu^*(t) = \min \left\{ 0.9, \max \left\{ 0, \frac{(\lambda_3 - \lambda_1)cx^*}{A} \right\} \right\}. \quad (24)$$

Proof. From Pontryagin's Maximum Principle, adjoint variables λ_i ($i = 1, 2, 3$) can be written as

$$\frac{d\lambda_1}{dt} = -\frac{\partial H}{\partial x}, \quad \frac{d\lambda_2}{dt} = -\frac{\partial H}{\partial y}, \quad \frac{d\lambda_3}{dt} = -\frac{\partial H}{\partial b}. \quad (25)$$

These are just differential equations (21)–(23) with transversality conditions $\lambda_i(t_{final}) = 0$ ($i = 1, 2, 3$). Furthermore, by optimality condition, we have that

$$\left. \frac{\partial H}{\partial \mu} \right|_{\mu=\mu^*} = -C_2 \mu^* - c\lambda_1 x^* + c\lambda_3 x^* = 0. \quad (26)$$

This shows that

$$\mu^* = \frac{(\lambda_3 - \lambda_1)cx^*}{C_2}. \quad (27)$$

Using the property of control space, we get that

$$\mu^* = \begin{cases} 0, & \frac{(\lambda_3 - \lambda_1) cx^*}{C_2} \leq 0, \\ \frac{(\lambda_3 - \lambda_1) cx^*}{C_2}, & 0 < \frac{(\lambda_3 - \lambda_1) cx^*}{C_2} < 0.9, \\ 0.9, & \frac{(\lambda_3 - \lambda_1) cx^*}{C_2} \geq 0.9. \end{cases} \quad (28)$$

This can be rewritten in compact notation, which is just (24). This completes the proof. \square

3.3. Numerical Simulation and Discussion. Here, we discuss how the continuous control measure affects the protection of endangered language and the existence and stability of positive equilibrium for model (10) by numerical simulations. Firstly, we choose the same parameters as in Figure 1(a); that is, $c = k = 0.5$ and $s = 0.7$. Besides that, we choose control variable $\mu = 0.25$. Figure 2(a), however, is completely different from Figure 1(a), which shows that model (10) has a globally asymptotically stable positive equilibrium $(x_\mu^*, y_\mu^*, b_\mu^*) = (0.4194, 0.1244, 0.4562)$. Furthermore, from the discussion of Theorem 4, it follows that values x_μ^* and y_μ^* will decrease and increase with increase in control strength μ , respectively, which is shown in Figures 2(b) and 2(c). The strong consistency between theoretical result and real situation is obviously observed.

In addition, in Figure 2(c), it is interesting to note that control strength μ is close to 1 (where $\mu = 0.95$), but language X runs around 0.2. This means that, no matter how strong the control strength μ is, language X will not fade. Namely, control measure can only protect endangered language Y , but not result in extinction of language X . Actually, in the real world, the own characteristics are very important factors in determining the development of languages.

The plots in Figure 3(a) show three adjoint variables λ_1 , λ_2 , and λ_3 in the optimality system. We solve these adjoint equations by a backward Runge-Kutta fourth-order procedure because of the transversality conditions for more details, see Lenhart and Workman [27]. In Figure 3(b), dotted line and solid line represent languages X , Y and bilingual B in model (10) without and with continuous control, respectively. We see that the fractions of language Y and bilingual B in population decrease more when there is no control. In this case, most of this population goes to language X . If we apply the continuous control measure, however, the fraction of language Y slowly falls, the fraction of language X decreases quite a lot, and the fraction of bilingual B is quite greater than the fraction in the case without control, since our main object is maximizing the fraction of bilingual B . In Figure 3(c), the control μ is plotted as a function of time for three different values of weight factor C_2 : 0.045, 0.1, and 0.5. The control variable μ for the associated weight factor $\mu = 0.045$ is much larger than the other two values. Note that, in general, as C_2 decreases, the amount of μ increases. The same results can also be obtained from the expression of μ in (27). The associated weight factor C_2 also plays a significant role

in keeping the balance of the size of fraction an optimal problem.

4. Protection of Language Y with Impulsive Control for $(1-s) < s(1-k)$

For reasons of protecting culture diversity, in this section, we will consider how the state-dependent pulse control measure affects the prevention of endangered language. Since learning cycle is very brief in contrast to the life cycle of a person, naturally we suppose that the procedure of learning is pulse effect.

As for the protection of endangered languages for the state-dependent control measure, we construct the following controlled model which is modeled by differential equations with state-dependent pulse effect:

$$\begin{aligned} \frac{dx}{dt} &= yP_{YX} + bP_{BX} - x(P_{XY} + P_{XB}), \\ \frac{dy}{dt} &= xP_{XY} + bP_{BY} - y(P_{YX} + P_{YB}), \\ \frac{db}{dt} &= xP_{XB} + yP_{YB} - b(P_{BY} + P_{BX}), \\ x &< E_X, \\ x(t^+) &= (1 - \theta)x(t), \\ y(t^+) &= y(t), \\ b(t^+) &= b(t) + \theta x(t), \\ x &= E_X. \end{aligned} \quad (29)$$

The meaning of model (29) is as follows: when the fraction of language X reaches the critical threshold value E_X at time t_E , controlling measure (for example, encouraging some speaker of language X to study language Y) is taken and the fractions of language X and bilingual speakers B immediately become $(1 - \theta)x(t_E)$ and $b(t_E) + \theta x(t_E)$, respectively.

Remark 7. It is obvious that the fractions of language Y and bilingual B are rather small and in danger of becoming extinct when the fraction of language X reaches the critical threshold value E_X . In this case, the effective measure is taken to prevent the loss of language Y . The times of control measures are obviously related to the state of language X .

Remark 8. The critical threshold value E_X represents the fraction of monolingual speakers X in population, and θ represents the strength of control measure. Numerical simulations in Section 4.3 show that these are crucial parameters in model (29).

For model (29), which is equivalent to the following reduced model since the population has a constant size $x + y + b = 1$,

$$\begin{aligned} \frac{dx}{dt} &= c(1-x)[(1-k)s(x+b) - x(1-s)], \\ \frac{db}{dt} &= c(1-x)[kx - b(1-s-k)] - bcs(x+b), \\ x &\neq E_X, \end{aligned}$$

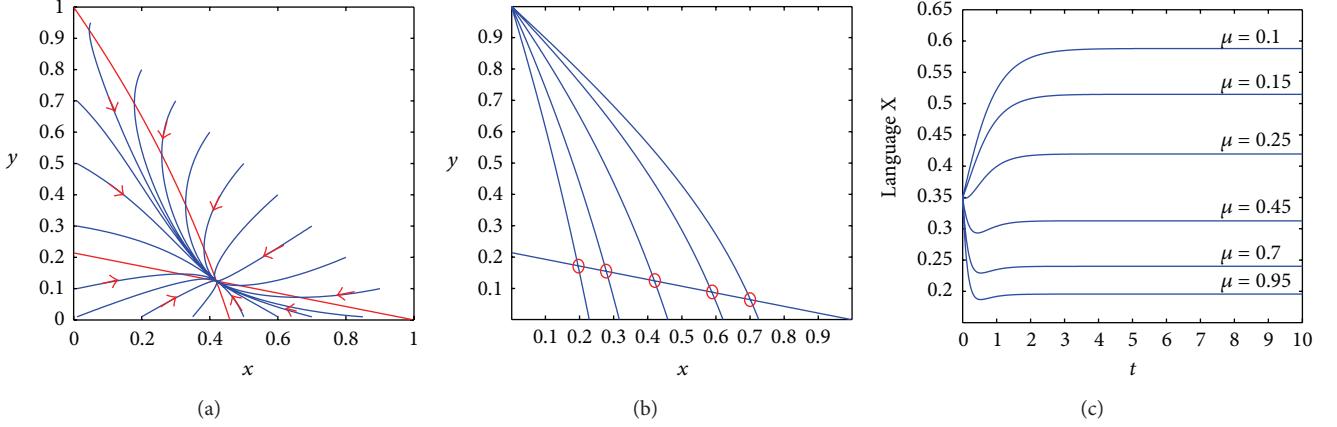


FIGURE 2: The effect of control measure on the existence and stability of positive equilibrium for model (10) with $c = k = 0.5$, $s = 0.7$, and $\alpha = 1$.

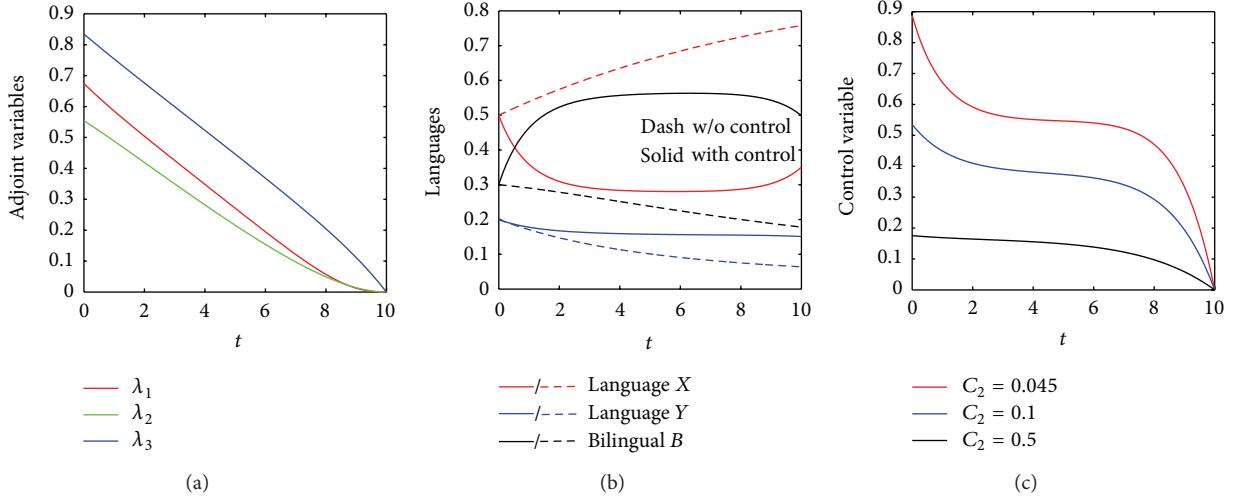


FIGURE 3: The optimal adjoint variables, states, and control variable for the optimal control problem with $c = k = 0.5$, $s = 0.7$, $\alpha = 1$, and $C_1 = 0.1$: (a) the optimal adjoint variables λ_1 , λ_2 , and λ_3 ; (b) the optimal states of languages X , Y and bilingual B ; and (c) the control variable μ with different weight factor C_2 .

$$\begin{aligned} x(t^+) &= (1 - \theta)x(t), \\ b(t^+) &= b(t) + \theta x(t), \\ x &= E_X. \end{aligned} \tag{30}$$

By the biological background of model (30), we only consider model in the biological meaning region $\Omega_0 = \{(x, b) : x \geq 0, b \geq 0, x + b \leq 1\}$. Obviously, the global existence and uniqueness of solution of model (30) are guaranteed by the smoothness properties of the right-side of model (30); for more details, see Lakshmikantham et al. [29].

4.1. Preliminaries. To discuss the dynamic behavior of model (30), we define two cross-sections to the phase space of model (30) by sections

$$\begin{aligned} \Gamma_\theta &= \{(x, b) : x = (1 - \theta)E_X, b \geq 0\}, \\ \Gamma_E &= \{(x, b) : x = E_X, b \geq 0\}. \end{aligned} \tag{31}$$

Suppose that point $A_n(E_X, b_n)$ is on section Γ_E ; then, trajectory

$$\begin{aligned} O^+(A_n, t_n) &= \{(x(t), b(t)) : (x(t), b(t)) \in \Omega_0, \\ &t \geq t_n, x(t_n) = E_X, b(t_n) = b_n\} \end{aligned} \tag{32}$$

jumps to point $A_n^+((1 - \theta)E_X, \hat{b}_n)$ on section Γ_θ due to pulse effects $x(t_n^+) = (1 - \theta)x(t_n)$ and $b(t_n^+) = b(t_n) + \theta x(t_n)$, then intersects section Γ_E at point $A_{n+1}(E_X, b_{n+1})$, and finally jumps to point $A_{n+1}^+((1 - \theta)E_X, \hat{b}_{n+1})$ on section Γ_θ again, where b_{n+1} is decided by b_n and parameters θ , E_X . Therefore, we defined a Poincaré map of section Γ_E as follows:

$$b_{n+1} = F(b_n, \theta, E_X). \tag{33}$$

Let $z(t) = (x(t), b(t))$ be a solution of model (30) with initial condition $z_0 = z(t_0) = ((1 - \theta)E_X, \hat{b}_0)$. Trajectory $O^+(z_0, t_0)$ starts from point $A_0((1 - \theta)E_X, \hat{b}_0)$; first, it intersects section Γ_E at point $B_1(E_X, b_1)$, then jumps to point

$A_1((1 - \theta)E_X, \hat{b}_1)$ on section Γ_θ due to pulse effects, and finally reaches point $B_2(E_x, b_2)$ on section Γ_E again. Repeating the above process, we have two points' sequences $\{A_{n-1}((1 - \theta)E_X, \hat{b}_{n-1})\}$ and $\{B_n(E_X, b_n)\}$ ($n = 1, 2, \dots$). We notice that the coordinates satisfy the relation $\hat{b}_n = b_n + \theta E_X$ ($n = 1, 2, \dots$).

Let \mathbb{S} be an arbitrary set in \mathbb{R}^2 , and let P be an arbitrary point in \mathbb{R}^2 . The distance between point P and set \mathbb{S} is defined by $d(P, \mathbb{S}) = \inf_{P_0 \in \mathbb{S}} |P - P_0|$. For the convenience of statement in the rest of this paper, we introduce some definitions.

Definition 9. Trajectory $O^+(z_0, t_0)$ of model (30) is said to be order- k periodic if there exists a positive integer $k \geq 1$ such that k is the smallest integer for $\hat{b}_0 = \hat{b}_k$.

Definition 10 (orbital stability [30]). Trajectory $O^+(X_0, t_0)$ is said to be orbitally stable if, given that $\varepsilon > 0$, there exists a constant $\delta = \delta(\varepsilon) > 0$ such that, for other solution $z^*(t)$ of model (30), when $d(z^*(t_0), O^+(z_0, t_0)) < \delta$, one has that $d(z^*(t), O^+(z_0, t_0)) < \varepsilon$ for all $t > t_0$.

Definition 11 (orbitally asymptotical stability [30]). Trajectory $O^+(z_0, t_0)$ is said to be orbitally asymptotically stable if it is orbitally stable, and there exists a constant $\eta > 0$ such that, for any other solution $z^*(t)$ of model (30), when $d(z^*(t_0), O^+(z_0, t_0)) < \eta$, $\lim_{t \rightarrow \infty} d(z^*(t), O^+(z_0, t_0)) = 0$.

Next, we consider the following autonomous model with pulse effects:

$$\begin{aligned} \frac{dx}{dt} &= f(x, y), & \frac{dy}{dt} &= g(x, y), & \varphi(x, y) &\neq 0, \\ \Delta x &= \xi(x, y), & \Delta y &= \eta(x, y), & \varphi(x, y) &= 0, \end{aligned} \quad (34)$$

where $f(x, y)$ and $g(x, y)$ are continuous differential functions defined on \mathbb{R}^2 and $\varphi(x, y)$ is a sufficiently smooth function with $\nabla \varphi(x, y) \neq 0$. Let $(\mu(t), \nu(t))$ be a positive T -periodic solution of model (34). The following result comes from Corollary 2 of Theorem 1 in [31].

Lemma 12 (analogue of Poincaré criterion). *If the Floquet multiplier μ satisfies condition $|\mu| < 1$, where*

$$\mu = \prod_{j=1}^n \kappa_j \exp \left\{ \int_0^T \left[\frac{\partial f(\mu(t), \nu(t))}{\partial x} + \frac{\partial g(\mu(t), \nu(t))}{\partial y} \right] dt \right\} \quad (35)$$

with

$$\begin{aligned} \kappa_j &= \left(\left(\frac{\partial \eta}{\partial y} \frac{\partial \varphi}{\partial x} - \frac{\partial \eta}{\partial x} \frac{\partial \varphi}{\partial y} + \frac{\partial \varphi}{\partial x} \right) f_+ \right. \\ &\quad \left. + \left(\frac{\partial \xi}{\partial x} \frac{\partial \varphi}{\partial y} - \frac{\partial \xi}{\partial y} \frac{\partial \varphi}{\partial x} + \frac{\partial \varphi}{\partial y} \right) g_+ \right) \\ &\quad \times \left(\frac{\partial \varphi}{\partial x} f + \frac{\partial \varphi}{\partial y} g \right)^{-1} \end{aligned} \quad (36)$$

and $f_+, g_+, \partial \xi / \partial x, \partial \xi / \partial y, \partial \eta / \partial x, \partial \eta / \partial y, \partial \varphi / \partial x$, and $\partial \varphi / \partial y$ have been calculated at the point $(\mu(\tau_j), \nu(\tau_j))$, $f_+ = f(\mu(\tau_j^+), \nu(\tau_j^+))$, $g_+ = g(\mu(\tau_j^+), \nu(\tau_j^+))$, and τ_j ($j \in N$) is the time of the j th jump, then, $(\mu(t), \nu(t))$ is orbitally asymptotically stable.

4.2. Main Results. On the existence of positive order-1 periodic solution for model (30), we have the following theorem.

Theorem 13. *For any $\theta \in (0, 1)$ and $E_X \in (0, 1)$, model (30) admits a positive order-1 periodic solution.*

Proof. Let point $A_0((1 - \theta)E_X, \hat{s}_0) \in \Gamma_\theta$ for sufficiently small \hat{s}_0 with $\hat{s}_0 \leq \theta E_X$. In view of the geometrical structure of the phase space of model (30), trajectory $O^+(A_0, t_0)$ of model (30) starts from initial point A_0 and intersects section Γ_E at point $B_1(E_X, s_1)$. And then, trajectory $O^+(A_0, t_0)$ jumps to point $A_1(\hat{s}_1, (1 - \theta)E_X)$ on section Γ_θ due to pulse control effects $x(t^+) = (1 - \theta)x(t)$ and $b(t^+) = b(t) + \theta x(t)$. Since $\hat{s}_0 \leq \theta E_X$, it follows that point A_1 is above point A_0 . Furthermore, point B_2 is above point B_1 . Otherwise, this is a contradiction with the uniqueness of solution for model (30). By (33), we have that $s_2 = F(s_1, \theta, E_X)$ and

$$F(s_1, \theta, E_X) - s_1 = s_2 - s_1 > 0. \quad (37)$$

On the other hand, let $\lambda_0 = 1 - (1 - \theta)E_X$. For critical point $C_0((1 - \theta)E_X, \lambda_0)$, trajectory $O^+(C_0, t_0)$ starts from initial point C_0 ; it intersects section Γ_E at point $D_1(E_X, \lambda_1)$, then jumps to point $C_1((1 - m)E_X, \lambda_1^+)$ on section Γ_θ , and finally reaches point $D_2(E_X, \lambda_2)$ on section Γ_E again. Since $x + b \leq 1$, one has that $(1 - m)E_X + \lambda_1^+ \leq 1 = (1 - \theta)E_X + \lambda_0$. That is, $\lambda_1^+ \leq \lambda_0$. If $\lambda_1^+ = \lambda_0$, namely, points C_0 and C_1 coincide, then model (30) has a positive order-1 periodic solution. Otherwise, $\lambda_1^+ < \lambda_0$; that is, point C_0 is above point C_1 , and point D_1 is above point D_2 . By (33), we can get that $\lambda_2 = F(\lambda_1, \theta, E_X)$ and

$$F(\lambda_1, \theta, E_X) - \lambda_1 = \lambda_2 - \lambda_1 < 0. \quad (38)$$

This together with (37) yields that Poincaré map (33) has a fixed point; that is, model (30) has a positive order-1 periodic solution. This completes the proof. \square

Now, on the orbital stability of positive order-1 periodic solution of model (30), we have the following result.

Theorem 14. *Let $(\varphi(t), \psi(t))$ be a positive order-1 periodic solution of model (30) with period T . For any $\theta \in (0, 1)$ and $E_X \in (0, 1)$, if the Floquet multiplier μ satisfies condition $|\mu| < 1$, where*

$$\begin{aligned} |\mu| &= |\kappa| \exp \left\{ \int_0^T \left[c(1 - \varphi(t))(3s + k - ks - 2 \right. \right. \\ &\quad \left. \left. - cs(\varphi(t) + \psi(t)) - cs\psi(t) \right] dt \right\} \end{aligned} \quad (39)$$

with

$$\kappa = \frac{(1-k)s(E_X + \psi(T)) - (1-\theta)E_X(1-s)}{(1-k)s(E_X + \psi(T)) - E_X(1-s)}, \quad (40)$$

then $(\varphi(t), \psi(t))$ is orbitally asymptotically stable.

Proof. Suppose that $(\varphi(t), \psi(t))$ intersects sections Γ_θ and Γ_E at points $C^*((1-\theta)E_X, \theta E_X + \psi(T))$ and $C(E_X, \psi(T))$, respectively. Comparing with model (34), we have that

$$\begin{aligned} f(x, b) &= c(1-x)[(1-k)s(x+b) - x(1-s)], \\ g(x, b) &= c(1-x)[kx - b(1-s-k)] - bcs(x+b), \end{aligned} \quad (41)$$

and $\xi(x, b) = -\theta x$, $\eta(x, b) = \theta x$, $\phi(S, I) = x - E_X$, $(\varphi(T), \psi(T)) = (E_X, \psi(T))$, and $(\varphi(T^+), \psi(T^+)) = ((1-\theta)E_X, \psi(T) + \theta E_X)$. Thus,

$$\begin{aligned} \frac{\partial f}{\partial x} &= -c[(1-k)s(x+b) - x(1-s)] \\ &\quad + c(1-x)(2s - ks - 1), \\ \frac{\partial g}{\partial b} &= -c(1-x)(1-s-k) - cs(x+b) - csb, \\ \frac{\partial \xi}{\partial x} &= -\theta, \quad \frac{\partial \eta}{\partial x} = \theta, \\ \frac{\partial \phi}{\partial x} &= 1, \quad \frac{\partial \xi}{\partial b} = \frac{\partial \eta}{\partial b} = \frac{\partial \phi}{\partial b} = 0. \end{aligned} \quad (42)$$

Furthermore, it follows from (42) that

$$\begin{aligned} \kappa &= \left(\left(\frac{\partial \eta}{\partial b} \frac{\partial \phi}{\partial x} - \frac{\partial \eta}{\partial x} \frac{\partial \phi}{\partial b} + \frac{\partial \phi}{\partial x} \right) f_+ \right. \\ &\quad \left. + \left(\frac{\partial \xi}{\partial x} \frac{\partial \phi}{\partial b} - \frac{\partial \xi}{\partial b} \frac{\partial \phi}{\partial x} + \frac{\partial \phi}{\partial b} \right) g_+ \right) \times \left(\frac{\partial \phi}{\partial x} f + \frac{\partial \phi}{\partial b} g \right)^{-1} \\ &= \frac{f_+(\varphi(T^+), \psi(T^+))}{f(\varphi(T), \psi(T))} \\ &= ((1-(1-\theta)E_X)[(1-k)s(E_X + \psi(T)) \\ &\quad - (1-\theta)E_X(1-s)]) \\ &\quad \times ((1-E_X)[(1-k)s(E_X \\ &\quad + \psi(T)) - E_X(1-s)])^{-1}, \end{aligned} \quad (43)$$

$$\begin{aligned} \mu &= \kappa \exp \left\{ \int_0^T [-c((1-k)s(\varphi(t) + \psi(t)) \right. \\ &\quad \left. - \varphi(t)(1-s)) + c(1-\varphi(t)) \right. \\ &\quad \times (2s - ks - 1) - c(1-\varphi(t))(1-s-k) \\ &\quad \left. - cs(\varphi(t) + \psi(t)) - cs\psi(t)] dt \right\}. \end{aligned} \quad (44)$$

On the other hand, integrating both sides of the first equation of model (30) along the orbit $\bar{C}^+ \bar{C}$, we have that

$$\begin{aligned} \ln \frac{1 - E_X}{1 - (1-\theta)E_X} &= \int_{(1-\theta)E_X}^{E_X} \frac{dx}{1-x} \\ &= - \int_0^T c [(1-k)s(\varphi(t) + \psi(t)) \\ &\quad - \varphi(t)(1-s)] dt. \end{aligned} \quad (45)$$

From (43)–(45), we can obtain that

$$\begin{aligned} |\mu| &= \left| \frac{(1-k)s(E_X + \psi(T)) - (1-\theta)E_X(1-s)}{(1-k)s(E_X + \psi(T)) - E_X(1-s)} \right| \\ &\quad \times \exp \left\{ \int_0^T [c(1-\varphi(t))(3s+k-ks-2) \right. \\ &\quad \left. - cs(\varphi(t) + \psi(t)) - cs\psi(t)] dt \right\}. \end{aligned} \quad (46)$$

By condition (39), we see that model (30) satisfies all conditions of Lemma 12. Therefore, order-1 periodic solution $(\varphi(t), \psi(t))$ of model (30) is orbitally asymptotically stable and has asymptotic phase property. This completes the proof. \square

Remark 15. From Theorem 14, though condition (39) of Theorem 14 is hard to test, yet it is weak since the second and third items of the exponent term of the right-side of (39) are negative.

Next, we give a more general result on the existence and stability of positive order-1 periodic solutions of model (30).

Theorem 16. For any $\theta \in (0, 1)$ and $E_X \in (0, 1)$, model (30) has a positive order-1 periodic solution which is orbitally asymptotically stable.

Proof. Let $\lambda^* = 1 - (1-\theta)E_X$, and suppose that the trajectory $O^+(C_0, t_0)$ of model (30) starts from critical point $C_0((1-\theta)E_X, \lambda^*)$ and intersects section Γ_E at point $C^*(E_X, \beta^*)$. In view of the geometrical construction of phase space of model (30) and that $(1, 0)$ is a globally asymptotically stable node, we obtain that trajectory of model (30) which starts from point $((1-\theta)E_X, b)$ with $b \in (0, \lambda^*)$ will intersect section Γ_E at point (E_X, \hat{b}) ; then, $\hat{b} \in (0, \beta^*)$. So we only need to consider trajectories of model (30) which start from the point (E_X, b) on section Γ_E , where $b \in (0, \beta^*)$.

Suppose that trajectory $O^+(D_1, t_0)$ of model (30) which starts from initial point $D_1(E_X, b_1)$ ($0 < b_1 < \beta^*$) jumps to point $D_1^+((1-\theta)E_X, \hat{b}_1)$ on section Γ_θ due to pulse effects and then reaches section Γ_E at point $D_2(E_X, b_2)$, where $\hat{b}_1 \in (0, \lambda^*)$ and $b_2 \in (0, \beta^*)$. Repeating the above process, we have pulse points sequences $\{D_n(E_X, b_n)\}$ and $\{D_n^+((1-\theta)E_X, \hat{b}_n)\}$, where $b_n \in (0, \beta^*)$ and $\hat{b}_n \in (0, \lambda^*)$. Furthermore, this follows from Poincaré map (33) that $b_{n+1} = F(b_n, \theta, E_X)$ ($n = 1, 2, \dots$). On the other hand, for any two points $D_i(E_X, b_i)$ and $D_j(E_X, b_j)$ on section Γ_E , where $b_i, b_j \in (0, \beta^*)$ and $b_i < b_j$, in view of pulse effects,

point $D_i^+((1-\theta)E_X, \hat{b}_i)$ is below point $D_j^+((1-\theta)E_X, \hat{b}_j)$. And then, trajectories $O^+(D_i, t_0)$ and $O^+(D_j, t_0)$ intersect section Γ_E at points $D_{i+1}(E_X, b_{i+1})$ and $D_{j+1}(E_X, b_{j+1})$, respectively. Therefore, from the geometrical construction of the phase space of model (30), we claim that

$$0 < b_{i+1} < b_{j+1} < \beta^*. \quad (47)$$

In fact, if inequality (47) does not hold, that is, $b_{i+1} \geq b_{j+1}$, then, it follows that point D_{i+1} is above point D_{j+1} or that two points coincide. Furthermore, we can obtain that trajectories $O^+(D_i, t_0)$ and $O^+(D_j, t_0)$ intersect at a point $D^*(\hat{x}, \hat{b})$. It is indicated that there are two different solutions which start from point D^* . This is a contradiction with the uniqueness of solution of model (30). So inequality (47) is thus valid.

Now, for any $b_0 \in (0, \beta^*)$, from Poincaré map (33) of section Γ_E , we have $b_1 = F(b_0, \theta, E_X)$, $b_2 = F(b_1, \theta, E_X)$, and $b_{n+1} = F(b_n, \theta, E_X)$ ($n = 2, 3, 4, \dots$). In particular, if $b_0 = b_1$, then model (30) has a positive order-1 periodic solution, if $b_0 \neq b_1$, then it follows from the fact (47) that $b_0 \neq b_2$.

Next, we discuss the general circumstance; that is, $b_0 \neq b_1 \neq b_2 \neq \dots \neq b_k$ ($k > 2$).

(a) If $b_0 < b_1$, from (47), we obtain that $b_1 < b_2$. Then, $b_0 < b_1 < b_2$. In this case, repeating the above process, we have that

$$0 < b_0 < b_1 < b_2 < b_3 < \dots < \beta^*. \quad (48)$$

(b) If $b_0 > b_1$, similar to (a), we have that

$$0 < \dots < b_3 < b_2 < b_1 < b_0 < \beta^*. \quad (49)$$

Therefore, in case (a), we can get that $\lim_{n \rightarrow \infty} b_n = \eta$. This implies that model (30) has an orbitally asymptotically stable positive order-1 periodic solution. Similarly, we can obtain the same result in case (b). This completes the proof. \square

Remark 17. According to the equivalence of models (29) and (30), from Theorem 16, we also obtained that model (29) has a positive order-1 periodic solution which is orbitally asymptotically stable. At the same time, it also implies that languages X and Y are coexistent and have a stable equilibrium state under state-dependent impulsive control strategy.

Remark 18. Similar results can also be obtained from the case that $s < (1-s)(1-k)$; we, hence, omit them here.

4.3. Numerical Simulation and Discussion. In this subsection, some numerical simulations are carried out to illustrate main results and the feasibility of state-dependent pulse feedback control measure. Firstly, we choose model parameters $c = k = 0.5$ and $s = 0.7$ and control parameters $E_X = 0.8$ and $\theta = 0.15$. From Figure 4, we see that state-dependent control measure plays an important role in preventing the disappearance of endangered language. Under state-dependent impulsive control measure, the downward trend for language Y was controlled effectively, and the fractions

of language X and bilingual B are kept within reasonable levels. Furthermore, numerical simulations also show that the fractions of languages X , Y and bilingual B , though from different initial states, are stabilized in the same state. The corresponding numerical results are presented in Figures 4 and 5. Namely, model (29) has a positive order-1 periodic solution, which is orbitally asymptotically stable. This is certainly the case as shown in Theorems 13–16. Again from Figure 5(a)–5(c), the periodic solution is orbitally asymptotically stable instead of being Lyapunov asymptotically stable. In fact, it also shows exactly how different the two stabilities are.

Next, we investigate what effect has the choice of controlling parameters on the dynamical behavior of model (29) using numerical modeling method. We first choose $E_X = 0.8$ and parameter θ to be 0.05, 0.15, 0.25, and 0.35, respectively. From Figure 6, we note that the length of times intervals between two control strategies are closely geared to the strength of control measure θ and that the time interval increases with the increasing of θ . Again from Figures 6(b) and 6(c), it is obvious that the fractions of language Y and bilingual B in population could maintain higher level for a long time due to larger θ . Of course, the cost of control measure is related to its strength. Furthermore, similar results can also be obtained from $\theta = 0.15$ and letting E_X be 0.9, 0.85, 0.8, and 0.7 in Figure 7. It is not hard to imagine, however, that the cost of control measure is very high if the fraction of monolingual speakers is incredibly low in the population. This is because if the fraction of monolingual speakers is high in the population, then it is extremely difficult to encourage monolingual speaker of the mainstream language to study endangered language. Of course, it is also not a good measure for the protection of endangered language. How do we choose appropriate parameters such that the fractions of language Y and bilingual B are kept at reasonable levels with the minimal cost of control measure? It is an interesting problem; at the same time, it is extremely difficult.

5. Concluding Remarks

The dynamic behavior of two-language competitive model (1) with parameters (2) and $\alpha = 1$ is analyzed systematically in this paper. By the linearization and Bendixson-Dulac theorem on dynamical system, some sufficient conditions on the globally asymptotical stability of the trivial equilibria, the existence, the local stability, and the global stability of positive equilibrium of model (1) are presented. The theoretical results show that languages X and Y are coexistent by adjusting the values of model parameters k and s .

And when considering the protection of endangered language, model (1) with (2) is extended to model (10) assessing the impact of continuous control measure. The theoretical results and numerical simulations indicate that the existence and stability of model (10) are sensitive to control parameter. Furthermore, using the optimal control theory, we derived and analyzed the conditions for optimality of the endangered language. Our results say that the optimal control has a very desirable effect for maintaining the fraction

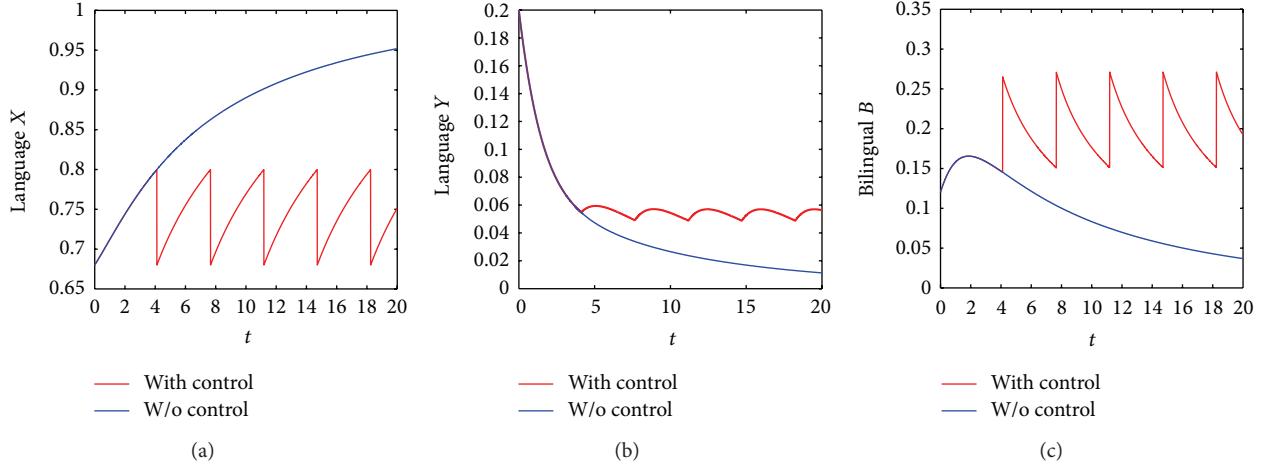


FIGURE 4: The contrast of the dynamical behaviors of model (29) with state-dependent impulsive control measure and without control, where $c = k = 0.5$, $s = 0.7$, $E_X = 0.8$, and $\theta = 0.15$.

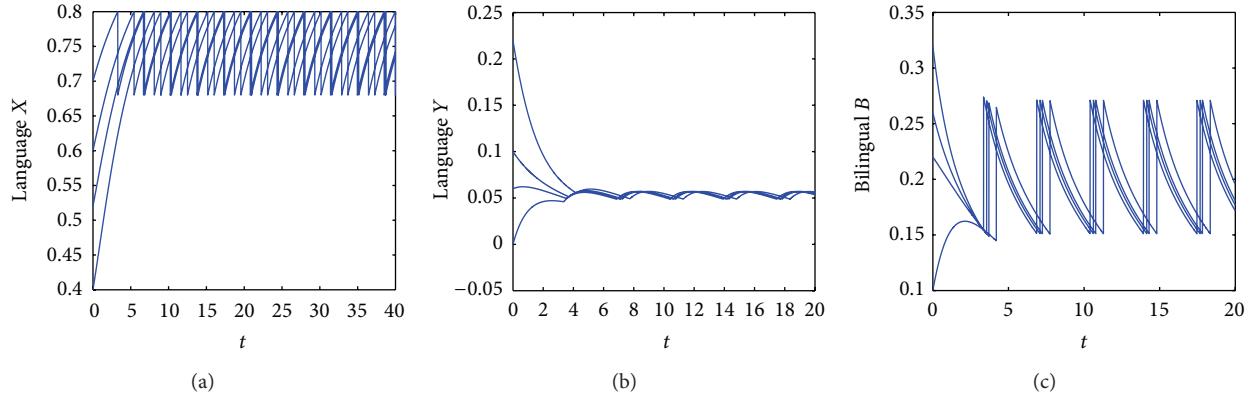


FIGURE 5: The orbitally asymptotical stability of positive order-1 periodic solution for model (29).

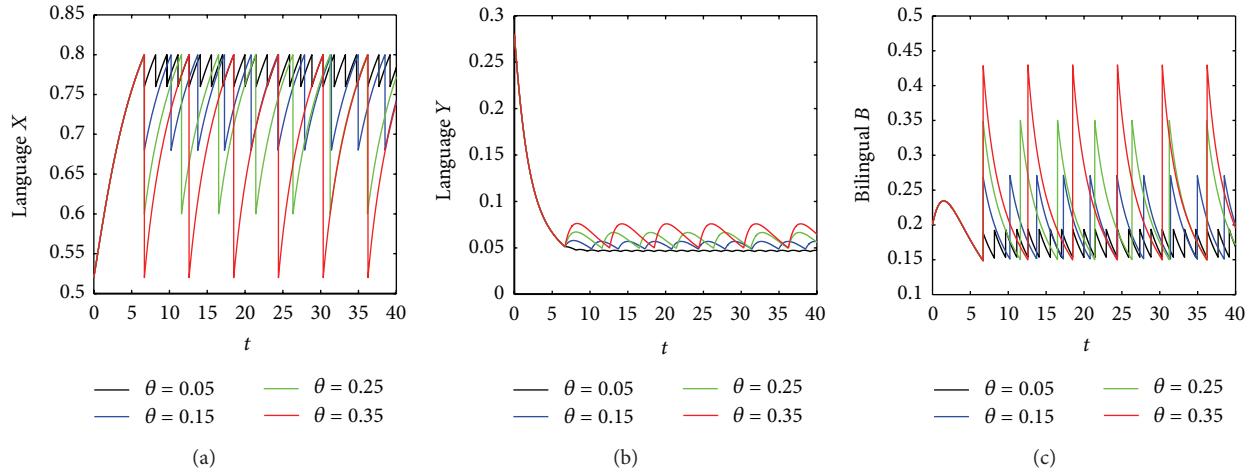


FIGURE 6: The contrast of the dynamical behaviors of model (29) with different state-dependent impulsive control parameters, where $E_X = 0.8$ and $\theta = 0.05, 0.15, 0.25$, and 0.35 , respectively.

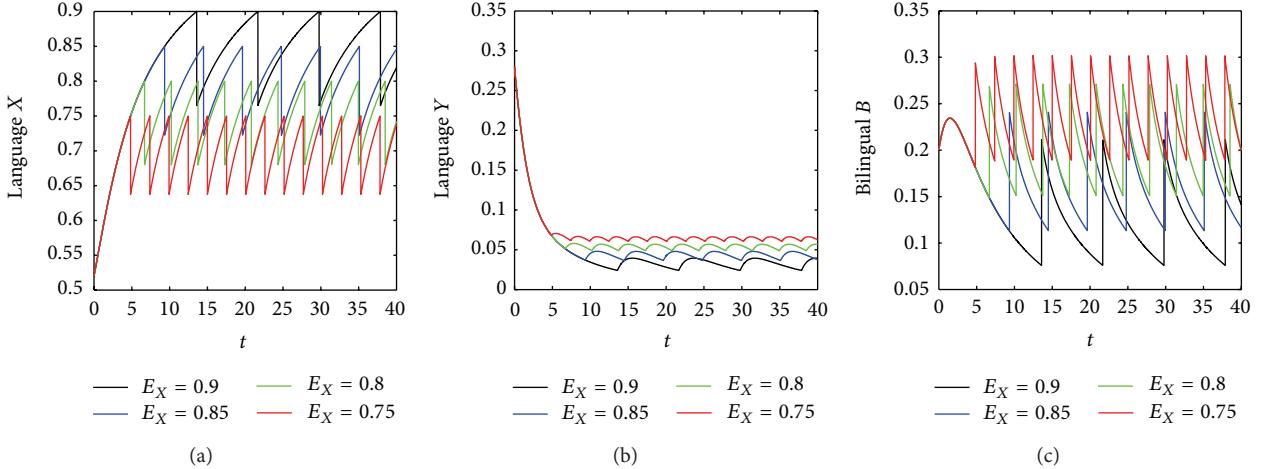


FIGURE 7: The contrast of the dynamical behaviors of model (29) with different state-dependent impulsive control parameters, where $\theta = 0.15$ and $E_X = 0.9, 0.85, 0.8$, and 0.75 , respectively.

of bilingual B , and some comparisons between with and without control are made in the figures.

Finally, the dynamic behavior of model (1) with state-dependent pulse control measure, that is, model (29), is studied in Section 4. The state-dependent pulse control measure causes the complexity for the dynamic behavior of model (29) such as frequent switching between states, irregular motion, and some uncertainties. This is the distinguished feature compared with continuous control measure. By the Poincaré map, analogue of Poincaré criterion, and qualitative analysis method, some sufficient conditions on the existence and orbitally asymptotical stability of positive order-1 periodic solution are presented. This amounts to the fact that we can control the fractions of languages X , Y and bilingual B at reasonable levels by adjusting control parameters. Theoretical basis for finding a new measure to protect the endangered language is provided.

Acknowledgments

The authors wish to thank the reviewers and the handling editor, Prof. Bashir Ahmad, for their comments and suggestions, which led to a great improvement to the presentation of this work. This paper has been partially supported by the Scientific Research Programmes of Colleges in Xinjiang (Grant no. XJEDU2011S08), the Natural Science Foundation of Xinjiang (Grant no. 2011211B08), the National Natural Science Foundation of China (Grants nos. 11001235, 11271312, and 11261056), the China Postdoctoral Science Foundation (Grants nos. 20110491750 and 2012T50836), and the Ministerio de Economía y Competitividad (Spain), Project no. MTM2010-15314 and has been co-financed by the European Community Fund FEDER.

References

- [1] W. J. Sutherland, “Parallel extinction risk and global distribution of languages and species,” *Nature*, vol. 423, no. 6937, pp. 276–279, 2003.
- [2] D. M. Abrams and S. H. Strogatz, “Modelling the dynamics of language death,” *Nature*, vol. 424, no. 6951, p. 900, 2003.
- [3] X. Castelló, V. M. Eguíluz, and M. San Miguel, “Ordering dynamics with two non-excluding options: Bilingualism in language competition,” *New Journal of Physics*, vol. 8, article 308, 2006.
- [4] I. Dyen, J. B. Kruskal, and P. Black, “An indo-european classification: a lexicostatistical experiment,” *American Philosophical Society*, vol. 82, p. 5, 1992.
- [5] T. Gong, L. Shuai, M. Tamariz, and G. Jäger, “Studying language change using Price equation and Pólya-urn dynamics,” *PLoS ONE*, vol. 7, no. 3, Article ID e33171, 2012.
- [6] J. A. Hawkins, “Innateness and function in language universals,” in *The Evolution of Human Languages*, J. A. Hawkins and G. M. Murray, Eds., Addison-Wesley, Reading, Mass, USA, 1992.
- [7] D. Helbing, A. Szolnoki, M. Perc, and G. Szabó, “Evolutionary establishment of moral and double moral standards through spatial interactions,” *PLoS Computational Biology*, vol. 6, no. 4, Article ID e1000758, 2010.
- [8] J. Mira and A. Plaiedes, “Interlinguistic similarity and language death dynamics,” *Europhysics Letters*, vol. 69, no. 6, pp. 1031–1034, 2005.
- [9] M. A. Nowak, N. L. Komarova, and P. Niyogi, “Computational and evolutionary aspects of language,” *Nature*, vol. 417, no. 6889, pp. 611–617, 2002.
- [10] F. Petroni and M. Serva, “Language distance and tree reconstruction,” *Journal of Statistical Mechanics*, vol. 2008, no. 8, Article ID P08012, 2008.
- [11] C. Schulze, D. Stauffer, and S. Wichmann, “Birth, survival and death of languages by Monte Carlo simulation,” *Communications in Computational Physics*, vol. 3, no. 2, pp. 271–294, 2008.
- [12] M. Patriarca and E. Heinsalu, “Influence of geography on language competition,” *Physica A*, vol. 388, no. 2-3, pp. 174–186, 2009.
- [13] M. Patriarca and T. Leppänen, “Modeling language competition,” *Physica A*, vol. 338, no. 1-2, pp. 296–299, 2004.
- [14] D. Stauffer, X. Castelló, V. M. Eguíluz, and M. San Miguel, “Microscopic Abrams-Strogatz model of language competition,” *Physica A*, vol. 374, no. 2, pp. 835–842, 2007.

- [15] J. Mira, L. F. Seoane, and J. J. Nieto, "The importance of interlinguistic similarity and stable bilingualism when two languages compete," *New Journal of Physics*, vol. 13, Article ID 033007, 2011.
- [16] V. Otero-Espinar, L. F. Seoane, J. J. Nieto, and J. Mira, "An analytic solution of a model of language competition with bilingualism and interlinguistic similarity," *Physica D*, vol. 264, pp. 17–26, 2013.
- [17] K. Blayneh, Y. Cao, and H.-D. Kwon, "Optimal control of vector-borne diseases: treatment and prevention," *Discrete and Continuous Dynamical Systems B*, vol. 11, no. 3, pp. 587–611, 2009.
- [18] H. R. Joshi, "Optimal control of an HIV immunology model," *Optimal Control Applications and Methods*, vol. 23, no. 4, pp. 199–213, 2002.
- [19] T. K. Kar and A. Batabyal, "Stability analysis and optimal control of an SIR epidemic model with vaccination," *BioSystems*, vol. 104, no. 2-3, pp. 127–135, 2011.
- [20] D. Kirschner, S. Lenhart, and S. Serbin, "Optimal control of the chemotherapy of HIV," *Journal of Mathematical Biology*, vol. 35, no. 7, pp. 775–792, 1997.
- [21] A. A. Lashari and G. Zaman, "Optimal control of a vector borne disease with horizontal transmission," *Nonlinear Analysis: Real World Applications*, vol. 13, no. 1, pp. 203–212, 2012.
- [22] M. Ozari, A. A. Lashari, I. H. Jung, and K. O. Okosun, "Stability analysis and optimalcontrol of a Vector-Borne disease with nonlinear incidence," *Discrete Dynamics in Nature and Society*, vol. 2012, Article ID 595487, 21 pages, 2012.
- [23] G. Zaman, Y. Han Kang, and I. H. Jung, "Stability analysis and optimal vaccination of an SIR epidemic model," *BioSystems*, vol. 93, no. 3, pp. 240–249, 2008.
- [24] E. Jung, S. Lenhart, and Z. Feng, "Optimal control of treatments in a two-strain tuberculosis model," *Discrete and Continuous Dynamical Systems B*, vol. 2, no. 4, pp. 473–482, 2002.
- [25] L. F. Nie, Z. D. Teng, and B. Z. Guo, "A state dependent pulse control strategy for a SIRS epidemic system," *Bulletin of Mathematical Biology*, vol. 75, no. 10, pp. 1697–1715, 2013.
- [26] L. Nie, Z. Teng, and A. Torres, "Dynamic analysis of an SIR epidemic model with state dependent pulse vaccination," *Nonlinear Analysis: Real World Applications*, vol. 13, no. 4, pp. 1621–1629, 2012.
- [27] S. Lenhart and J. T. Workman, *Optimal Control Applied to Biological Models*, Chapman & Hall/CRC, London, UK, 2007.
- [28] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The Mathematical Theory of Optimal Process*, vol. 14, Gordon & Breach Science, New York, NY, USA, 1986.
- [29] V. Lakshmikantham, D. D. Bainov, and P. S. Simeonov, *Theory of Impulsive Differential Equations*, World Scientific, Singapore, 1989.
- [30] J. Hale and H. Kocak, *Dynamics and Bifurcations*, Springer, New York, NY, USA, 1991.
- [31] P. S. Simeonov and D. D. Bainov, "Orbital stability of periodic solutions of autonomous systems with impulse effect," *International Journal of Systems Science*, vol. 19, no. 7, pp. 2561–2585, 1988.

Research Article

Multiscale Characterization of Sea Clutter by Scale-Dependent Lyapunov Exponent

Jing Hu^{1,2} and Jianbo Gao^{1,2}

¹ School of Mechanical Engineering, Guangxi University, 100 Daxue Road, Nanning, Guangxi 530005, China

² PMB Intelligence LLC, Sunnyvale, CA 94087, USA

Correspondence should be addressed to Jianbo Gao; jbao.pmb@gmail.com

Received 13 August 2013; Accepted 10 October 2013

Academic Editor: J. A. Tenreiro Machado

Copyright © 2013 J. Hu and J. Gao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Determining whether sea clutter radar returns are stochastic or deterministic is crucial to the successful modelling of sea clutter as well as to facilitate target detection within sea clutter. Despite extensive studies of sea clutter using distributional analysis, chaos analysis, and fractal analysis, the nature of sea clutter is still not well understood. Realizing that the difficulty in sea clutter modeling is due to the multiscale nature of sea clutter, we employ a new multiscale complexity measure, the scale-dependent Lyapunov exponent (SDLE), to better characterize the nonstationary and multiscale nature of sea clutter. SDLE has been shown to readily characterize major models of complex time series, including deterministic chaos, noisy chaos, stochastic oscillations, random $1/f$ processes, random Levy processes, and complex time series with multiple scaling behaviors. With SDLE, we are able to directly show that sea clutter is not chaotic. More importantly, we find a new scaling law suggesting noisy dynamics for sea clutter. The new scaling law has an interesting interpretation in terms of intrinsic predictability of sea clutter, and provides an excellent new means of detecting targets within sea clutter.

1. Introduction

Sea clutter, the backscatter of the transmitted radar signal by elements of the sea surface such as ocean sprays, capillary waves, wind waves, and swells, is highly complicated. The complexity of the signals comes from two sources, the rough sea surface, sometimes oscillatory, sometimes turbulent, and the multipath propagation of radar backscatter. From Figure 1, it is clear that the signal is not purely random, since the waveform can be fairly smooth on short time scales (Figure 1(a)). However, the signal is highly nonstationary, since the frequency of the signal (Figures 1(a) and 1(b)) and the randomness of the signal (Figures 1(c) and 1(d)) change over time drastically. Therefore, Fourier analysis or deterministic chaotic analysis of sea clutter cannot be effective. From Figure 1(e), where $X_t^{(m)} = (X_{tm-m+1} + \dots + X_{tm})/m$, it can be concluded that neither autoregressive (AR) models nor textbook fractal models can describe the data. This is because AR modeling requires exponentially decaying autocorrelation (which amounts to $\text{Var}(X_t^{(m)}) \sim m^{-1}$), while fractal modeling requires the variation between $\text{Var}(X_t^{(m)})$

and m to follow a power-law [1]. (To see better the variation of $\text{Var}(X_t^{(m)})$ with m , $\log_2[m^2 \text{Var}(X_t^{(m)})]$ is plotted against $\log_2 m$ in the figure.) However, neither AR nor true fractal behavior is observed in Figure 1(e).

Understanding of sea clutter will depend not only on finding suitable models to describe the surface scattering but also on knowledge of the complex behavior of the sea [2]. As radars become more complex, to fully utilize their power, it has become increasingly important to accurately model their interaction with environment with ever greater fidelity. Therefore, the study of sea clutter continues to be an important aspect of the development of radar systems.

Over the past several decades, a lot of efforts have been made to study sea clutter. Fundamental studies of sea clutter can be roughly classified into three categories: (i) distributional analysis of sea clutter [3–15], (ii) analysis of sea clutter using fractal theory [1, 16–25], and (iii) analysis of sea clutter using chaos theory [26–37]. Despite these extensive studies, however, the nature of sea clutter is still not well understood. As a result, the important problem of target detection within sea clutter remains to be a tremendous

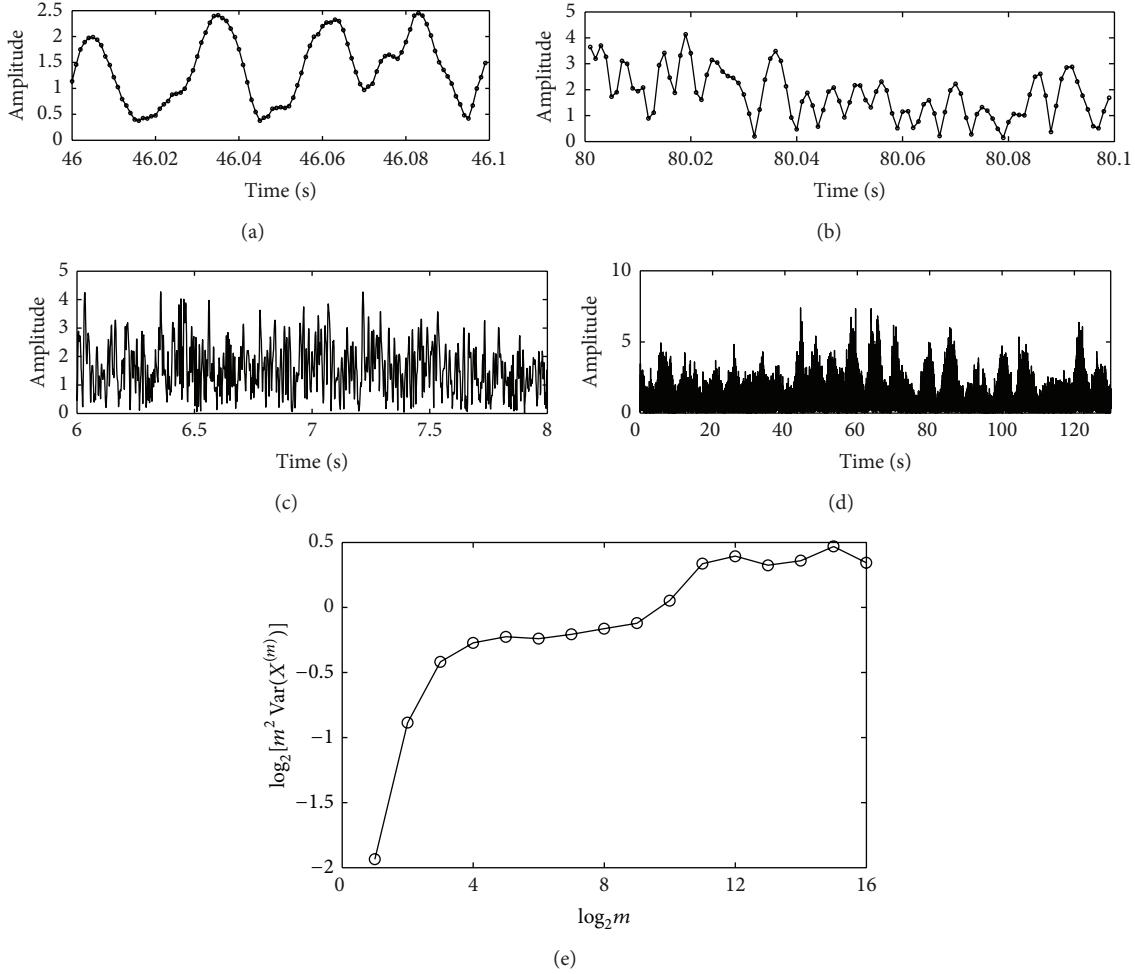


FIGURE 1: ((a) and (b)) Two 0.1 s duration sea clutter signal; (c) a 2 s duration sea clutter signal; (d) the entire sea clutter signal (of about 130 s); and (e) $\log_2[m^2 \text{Var}(X^{(m)})]$ versus $\log_2 m$.

challenge [16, 22, 38–44]. To shed new light on the problem, in this study, we carry out a multiscale analysis of sea clutter, using a recently developed multiscale complexity measure, the scale-dependent Lyapunov exponent (SDLE) [1, 45].

The most salient feature of multiscale signals is that the signals behave differently, depending upon the scale at which the data are examined [1]. To understand why multiscale analysis of sea clutter is essential, it is important to highlight a few empirical observations that point to the multiscale nature of sea clutter: (1) real sea clutter data are highly nonstationary, on both short and long time scales. This has motivated us to fit the differentiated data of sea clutter by Tsallis distribution [13]. (2) In the time scale range of 0.01 to 4 s, structure-function based multifractal characterization detects targets within sea clutter with very high accuracy, although the power-law fractal scaling is not very well defined [1, 23–25]; (3) when analyzed by multiplicative cascade multifractal [35], sea clutter again shows a fractal scaling break on time scales shorter than about 0.01 s. When the envelope of the sea clutter data is extracted by finding all the local

maxima and subsequently analyzed, the cascade multifractal characterization becomes excellent; furthermore, log-normal distribution fits the envelope data of sea clutter excellently, but not the original sea clutter data [35]. These observations motivate us to carry out a multiscale analysis of sea clutter by using scale-dependent Lyapunov exponent (SDLE) [1, 45], to shed new light on the nature of sea clutter.

The remainder of the paper is organized as follows. In Section 2, we briefly describe the SDLE, focusing on its properties that are most relevant to sea clutter analysis. In Section 3, we use the SDLE to analyze sea clutter, consider whether such characterization can help detect targets within sea clutter, and compare SDLE with the conventional largest positive Lyapunov exponent. In Section 4, we make a few concluding remarks.

2. SDLE as a Multiscale Complexity Measure

2.1. SDLE: Definition and Computation. SDLE is a recently developed multiscale complexity measure [1, 45]. It has been

used to characterize brainwave EEG data [46, 47], heart rate variability (HRV) [48–51], financial time series [52], Earth's geodynamo [53], among others. SDLE is defined in a phase space through consideration of an ensemble of trajectories [1, 45]. In the case of a scalar time series $x(1), x(2), \dots, x(n)$, a suitable phase space may be obtained by using time delay embedding [54] to construct vectors of the following form:

$$V_i = [x(i), x(i+L), \dots, x(i+(m-1)L)], \quad (1)$$

where m and L are called the embedding dimension and the delay time, respectively. For chaotic systems, m and L have to be chosen according to certain optimization criterion [1]. For a stochastic process, which is infinite-dimensional, the embedding procedure transforms a self-affine stochastic process to a self-similar process in a phase space, and often $m = 2$ is sufficient (this issue will be further explained later) [1, 45].

We will now be more concrete. Denote the initial distance between two nearby trajectories by ε_0 , and their *average distances* at time t and $t + \Delta t$, respectively, by ε_t and $\varepsilon_{t+\Delta t}$, where Δt is small. The SDLE $\lambda(\varepsilon_t)$ is defined by [1, 45]

$$\varepsilon_{t+\Delta t} = \varepsilon_t e^{\lambda(\varepsilon_t)\Delta t}, \quad \text{or} \quad \lambda(\varepsilon_t) = \frac{\ln \varepsilon_{t+\Delta t} - \ln \varepsilon_t}{\Delta t}. \quad (2)$$

Or equivalently by

$$\frac{d\varepsilon_t}{dt} = \lambda(\varepsilon_t) \varepsilon_t, \quad \text{or} \quad \frac{d \ln \varepsilon_t}{dt} = \lambda(\varepsilon_t). \quad (3)$$

To compute SDLE, we can start from an arbitrary number of shells as follows:

$$\varepsilon_k \leq \|V_i - V_j\| \leq \varepsilon_k + \Delta\varepsilon_k, \quad k = 1, 2, 3, \dots, \quad (4)$$

where V_i and V_j are reconstructed vectors, ε_k (the radius of the shell) and $\Delta\varepsilon_k$ (the width of the shell) are arbitrarily chosen small distances ($\Delta\varepsilon_k$ is not necessarily a constant). Then, we monitor the evolution of all pairs of points (V_i, V_j) within a shell and take average. Equation (2) can now be written as

$$\lambda(\varepsilon_t) = \frac{\langle \ln \|V_{i+t+\Delta t} - V_{j+t+\Delta t}\| - \ln \|V_{i+t} - V_{j+t}\| \rangle}{\Delta t}, \quad (5)$$

where t and Δt are integers in unit of the sampling time, and the angle brackets denote average within a shell. Note that this computational procedure is similar to that for computing the so-called time-dependent exponent curves [55].

Note that the initial set of shells serve as initial values of the scales; through evolution of the dynamics, they will automatically converge to the range of inherent scales. This is emphasized by the subscript t in ε_t , when the scales become inherent, t can then be dropped. Also, note that when analyzing chaotic time series, the following condition:

$$|j - i| \geq (m - 1)L \quad (6)$$

needs to be imposed when finding pairs of vectors within a shell, to eliminate the effects of tangential motions [1]. This

condition is often also sufficient for an initial scale to converge to the inherent scales [1].

To better understand SDLE, it is instructive to point out a relation between SDLE and the largest positive Lyapunov exponent λ_1 for true chaotic signals. It is given by [1]

$$\lambda_1 = \int_0^{\varepsilon^*} \lambda(\varepsilon) p(\varepsilon) d\varepsilon, \quad (7)$$

where ε^* is a scale parameter (e.g., used for renormalization when using Wolf et al.'s algorithm [56]), $p(\varepsilon)$ is the probability density function for the scale ε given by

$$p(\varepsilon) = Z \frac{dC(\varepsilon)}{d\varepsilon}, \quad (8)$$

where Z is a normalization constant satisfying $\int_0^{\varepsilon^*} p(\varepsilon) d\varepsilon = 1$ and $C(\varepsilon)$ is the well-known Grassberger-Procaccia's correlation integral [57].

SDLE has distinctive scaling laws for different types of time series. Those most relevant to sea clutter analysis are listed here.

- (1) For clean chaos on small scales, and noisy chaos with weak noise on intermediate scales,

$$\lambda(\varepsilon) = \lambda_1. \quad (9)$$

As an operational definition for chaos, we define chaos to be observing scaling of (9) on a scale range of $(\varepsilon, r\varepsilon)$, where $r > 1$ is a coefficient [1, 45]. When low-dimensional chaos is concerned, one may require $r \geq 2$.

- (2) For clean chaos on large scales where memory has been lost and for noisy chaos (including noise-induced chaos [58–60]) on small scales,

$$\lambda(\varepsilon) \sim -\gamma \ln \varepsilon, \quad (10)$$

where $\gamma > 0$ is a parameter.

- (3) For random $1/f^{2H+1}$ processes, where $0 < H < 1$ is called the Hurst parameter which characterizes the correlation structure of the process: depending on whether H is smaller than, equal to, or larger than $1/2$, the process is said to have antipersistent, short-range, or persistent long-range correlations [1, 23] as follows:

$$\lambda(\varepsilon) \sim \varepsilon^{-1/H}. \quad (11)$$

Note that $H = 1/2$ for the standard Brownian motion and $H = 1/3$ for turbulence. Note also that this scaling is quite independent of m , except that the scale range defining (11) shrinks when m increases. Because of this, $m = 2$ again is preferred.

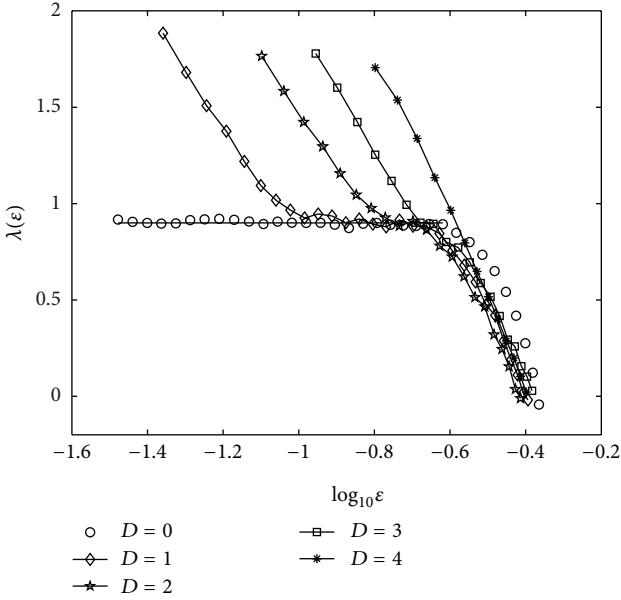


FIGURE 2: $\lambda(\epsilon)$ curves for the clean and the noisy Lorenz system.

To better appreciate the above properties, it is helpful to examine the chaotic Lorenz system with stochastic forcing as an example as follows:

$$\begin{aligned} \frac{dx}{dt} &= -10(x - y) + D\eta_1(t), \\ \frac{dy}{dt} &= -xz + 28x - y + D\eta_2(t), \\ \frac{dz}{dt} &= xy - \frac{8}{3}z + D\eta_3(t), \end{aligned} \quad (12)$$

where $D\eta_i(t)$, $i = 1, 2, 3$ are independent Gaussian noise forcing terms with mean 0 and variance D^2 . Note that the system with a different parameter set was studied in [1, 45]. The system is solved using the scheme of Exact propagator [61], where the exact solution of the Lorenz system is solved using a 4th order Runge-Kutta method with a time-step of $h = 0.002$, and then a term $D\sqrt{h}W$, where W is a Gaussian noise of mean 0 and variance 1, is added to the corresponding equations to take into account the noise. Figure 2 shows five curves, for the cases of $D = 0, 1, 2, 3, 4$. The computations are done using $m = 4$, $L = 2$, and 10000 points of the x component of the Lorenz system sampled at a time interval of 0.06. For the clean system, we observe two scaling laws. One is (9), $\lambda(\epsilon) \approx 0.9$, for small ϵ , the other is (10), for large ϵ where memory has been lost. For the noisy system, the scale region where the scaling law of (9) shrinks when the stochastic forcing is increased. Interestingly, although the part of the curve with $\lambda(\epsilon) \sim -\gamma \ln \epsilon$ shifts to the right when noise is increased, the parameter γ appears to not depend on the noise strength.

2.2. Coping with Nonstationarity. Since sea clutter is nonstationary [1, 13], it is important to understand how SDLE

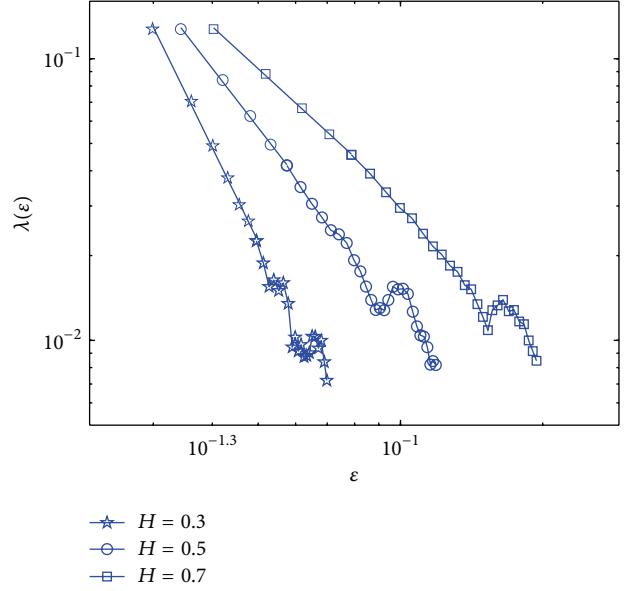


FIGURE 3: $\lambda(\epsilon)$ versus ϵ curves for perturbed $1/f$ processes. Curves for three different H are plotted. To put all the curves on one plot, the curves for different H (except the smallest one considered here) are arbitrarily shifted rightward.

deals with nonstationarity before it is used to analyze sea clutter. Noting that sea clutter has fractal $1/f$ scaling behavior in the time scale range from 0.01 to 4 s [1, 23–25], we first study if the $1/f$ feature remains robust under the following perturbations.

- (1) Shift a $1/f^\beta$, $\beta = 2H+1$ process downward or upward at randomly chosen points in time by an arbitrary amount. For convenience, we call this procedure type 1 nonstationarity and the processes obtained broken- $1/f^\beta$ processes.
- (2) At randomly chosen time intervals, concatenate randomly broken- $1/f^\beta$ processes and oscillatory components or superimpose oscillatory components on broken- $1/f^\beta$ processes. This procedure causes a different type of nonstationarity, which for convenience we will call type 2 nonstationarity.

We call the resulting random processes perturbed $1/f^\beta$ processes. Three examples of the $\lambda(\epsilon)$ curves for such processes, where the frequency of the perturbations are, on average, 1% of the simulated data, are shown in Figure 3. We observe that (11) still holds very well when $\lambda(\epsilon) > 0.02$. Therefore, SDLE can readily characterize $1/f$ processes perturbed by either of the nonstationarities considered.

To understand why the SDLE can deal with type 1 nonstationarity, it suffices to note that type 1 nonstationarity causes shifts of the trajectory in phase space; the greater the nonstationarity, the larger the shifts. The SDLE, however, cannot be affected much by shifts, especially large ones, since it is based on the coevolution of pairs of vectors within chosen small shells. In fact, the effect of shifts is to exclude a few pairs of vectors that were originally counted in the ensemble

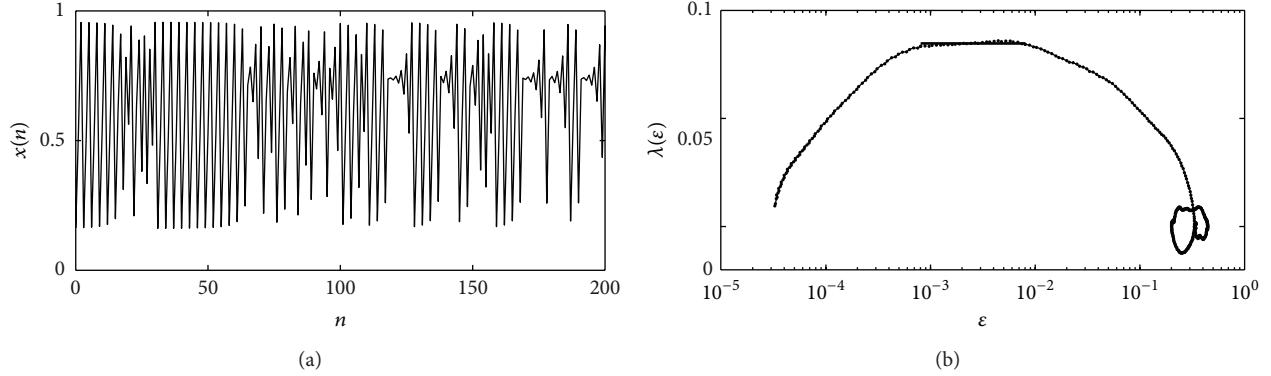


FIGURE 4: (a) An intermittent time series generated by the logistic map. (b) The SDLE curve for the time series.

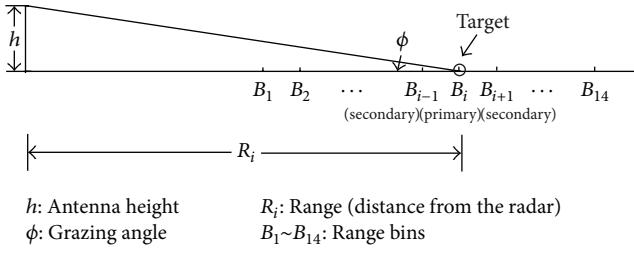


FIGURE 5: A schematic showing how the sea clutter data were collected.

average. Therefore, so long as the shifts are not too frequent, the effect of shifts can be neglected, since ensemble average within a shell involves a large number of pairs of vectors.

Let us now turn to type 2 nonstationarity, which involves oscillatory components. Being regular, oscillatory components can only affect $\lambda(\varepsilon)$ where it is close to 0. Therefore, type 2 nonstationarity cannot affect the positive portion of $\lambda(\varepsilon)$ either.

Note that similar types of perturbations have been carried out to clean and noisy chaotic data, and the major features of (9) and (10) are also robust, as expected.

Finally, we examine whether SDLE can detect chaos from an intermittent time series with a long laminar phase during which neighboring trajectories do not diverge, and a rapid divergence over a small part of the state space. Such a case is most relevant to the transitions from a quiet sea surface to a turbulent sea disturbed by a sudden gust of wind, and vice versa. For this purpose, we study the logistic map as follows:

$$x_{n+1} = ax_n(1 - x_n), \quad a = 3.8284. \quad (13)$$

Figure 4(a) shows an intermittent time series. We note that the chaotic part is very short compared with the periodic part. Figure 4(b) shows the SDLE curve for the time series. We clearly observe a plateau defining a chaotic motion. Also, note that the $\ln \varepsilon$ scaling on very small scales indicates the transitions from periodic to chaotic motions, and vice versa (it is unstable, but not chaotic in the rigorous sense).

3. Multiscale Analysis of Sea Clutter

3.1. Sea Clutter Data. Fourteen sea clutter datasets were obtained from a website maintained by Professor Simon Haykin: <http://soma.ece.mcmaster.ca/iphix/dartmouth/datasets.html>.

The measurement was made using the McMaster IPIX radar at Dartmouth, Nova Scotia, Canada. The radar was mounted in a fixed position on land 25–30 m above sea level, with an operating (carrier) frequency of 9.39 GHz (and hence a wavelength of about 3 cm). It was operated at low grazing angles, with the antenna dwelling in a fixed direction, illuminating a patch of ocean surface. The measurements were performed with the wave height in the ocean varying from 0.8 to 3.8 m (with peak heights up to 5.5 m) and the wind conditions varying from still to 60 km/hr (with gusts up to 90 km/hr). For each measurement, 14 areas, called antenna footprints or range bins, were scanned. Their centers are depicted as B_1, B_2, \dots, B_{14} in Figure 5. The distance between two adjacent range bins was 15 m. One or a few range bins (say, B_{i-1}, B_i and B_{i+1}) hit a target, which was a spherical block of styrofoam of diameter 1 m wrapped with wire mesh. The locations of the three targets were specified by their azimuthal angle and distance to the radar. They were $(128^\circ, 2660 \text{ m})$, $(130^\circ, 5525 \text{ m})$, and $(170^\circ, 2655 \text{ m})$, respectively. The range bin where the target is strongest is labeled as the primary target bin. Due to the drift of the target, bins adjacent to the primary target bin may also have hit the target. They are called secondary target bins. For each range bin, there were 2^{17} complex numbers, sampled with a frequency of 1000 Hz.

3.2. Sea Clutter Analysis. The SDLE curves can be readily computed from sea clutter data. Figure 6(a) shows a typical example. We do not observe the chaotic scaling of (9) on any significant scale ranges. Therefore, sea clutter is not chaotic. While this result is consistent with the results of [29, 30, 33–37], it merits noting that the SDLE method is a direct method and is more rigorous.

More interestingly, all the sea clutter data, whether it is with target or without target, yields the scaling of (10). Therefore, (10) describes a prevailing feature of sea clutter. To understand which feature that is as well as to understand

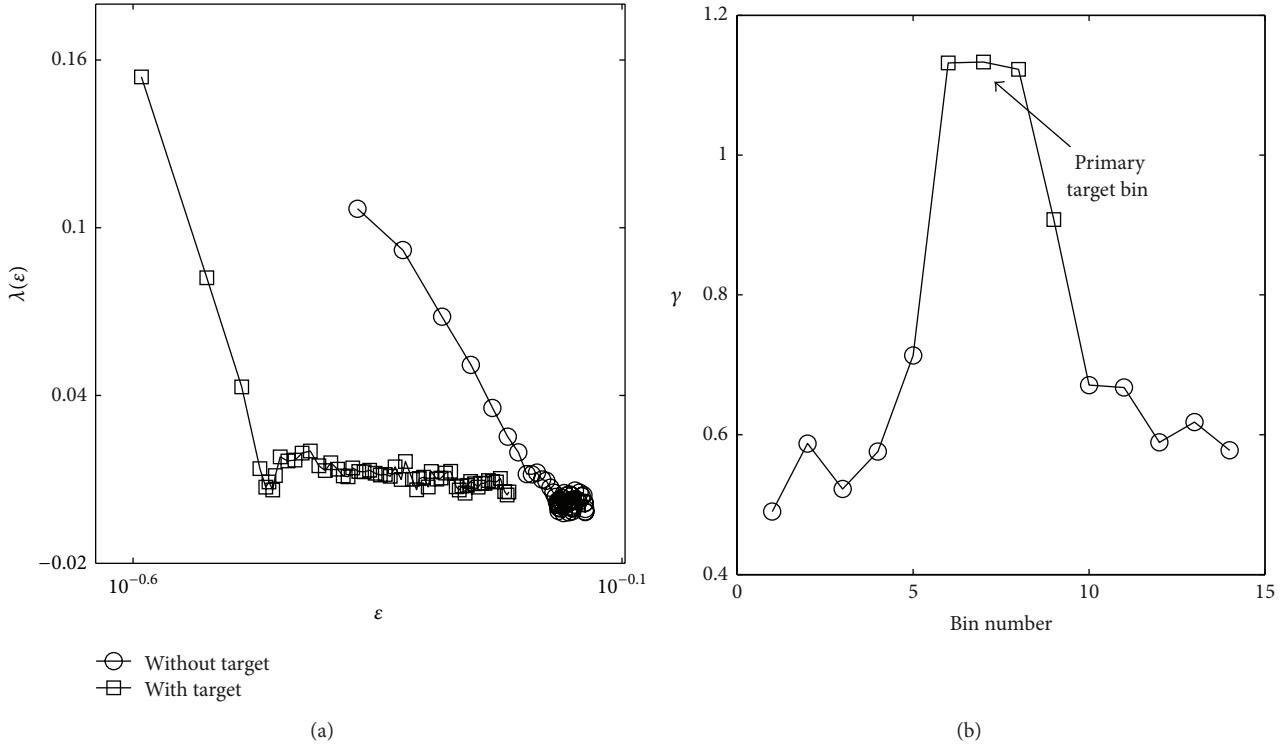


FIGURE 6: (a) $\lambda(\varepsilon)$ curves for sea clutter. The embedding parameters are $m = 4$, $L = 1$. (b) Target detection using γ parameter.

the physical significance of γ , let us consider the problem of predictability with sea clutter. Specifically, we consider the concept of the error *doubling time*, defined as the average time for an initial error ε_0 to double: an error grows faster when the doubling time is shorter. In other words, a faster error growth means a more rapid loss of initial information.

What is the error doubling time for sea clutter data? To find this, we use (3) to get the following:

$$\ln \varepsilon_t = \ln \varepsilon_0 + \int_0^t \lambda(\varepsilon_t) dt. \quad (14)$$

Letting $\varepsilon_{T_{db}} = 2\varepsilon_0$, we find the error doubling time T_{db} given by

$$\ln 2 = \int_0^{T_{db}} \lambda(\varepsilon_t) dt. \quad (15)$$

As the first approximation, we may consider $1/\lambda(\varepsilon)$ to be proportional to the doubling time. When $\lambda(\varepsilon)$ is given by (10), then the error doubling time is inversely proportional to γ . Therefore, γ characterizes the speed of information loss. More rigorously, we can prove that $-\ln \varepsilon_0 \gg 1$ decays to a saturation or limiting value of 0 exponentially as follows:

$$-\ln \varepsilon_t = -\ln \varepsilon_0 e^{-\gamma t}. \quad (16)$$

A zero magnitude of $\ln \varepsilon_t$ amounts to a unit error. Therefore, it is proper to call γ the relaxation parameter of information loss. The fact that γ is larger for sea clutter with target implies that information loss in sea clutter with target is

faster. Physically, this may be attributed to the generation of turbulence near the target.

Can the γ parameter be used to help detect targets within sea clutter? The answer is positive, as shown in Figure 6(b). Clearly, γ is very effective in distinguishing sea clutter with and without targets. For automatic detection purposes, we have analyzed almost 400 datasets available to us and found that a threshold value of 0.9 for γ yields an almost perfect classification as follows.

- (i) Hypothesis H_0 : sea clutter without target, $\gamma < 0.9$.
- (ii) Hypothesis H_1 : sea clutter with target, $\gamma \geq 0.9$.

When a signal is truly chaotic, SDLE is related to the conventional positive Lyapunov exponent by (7). Now that we have found sea clutter not to be chaotic, it is interesting to find out how different SDLE is from λ_1 . We have computed λ_1 for all the sea clutter data using Wolf et al.'s algorithm [56]. Wolf et al.'s algorithm works as follows: one selects a reference trajectory and follows the divergence of its neighboring trajectory from it. Denote the reference and the neighboring trajectories by $X_i = [x(i), x(i+L), \dots, x(i+(m-1)L)]$, $X_j = [x(j), x(j+L), \dots, x(j+(m-1)L)]$, $i = 1, 2, \dots$, and $j = K, K+1, \dots$, respectively. At the start of the time (which corresponds to $i = 1$), X_K is usually taken as the nearest neighbor of X_1 . That is, $j = K$ minimizes the distance between X_j and X_1 . When time evolves, the distance between X_i and X_j also changes. Let the spacing between the two trajectories at time t_i and t_{i+1} be d'_i and d_{i+1} , respectively.

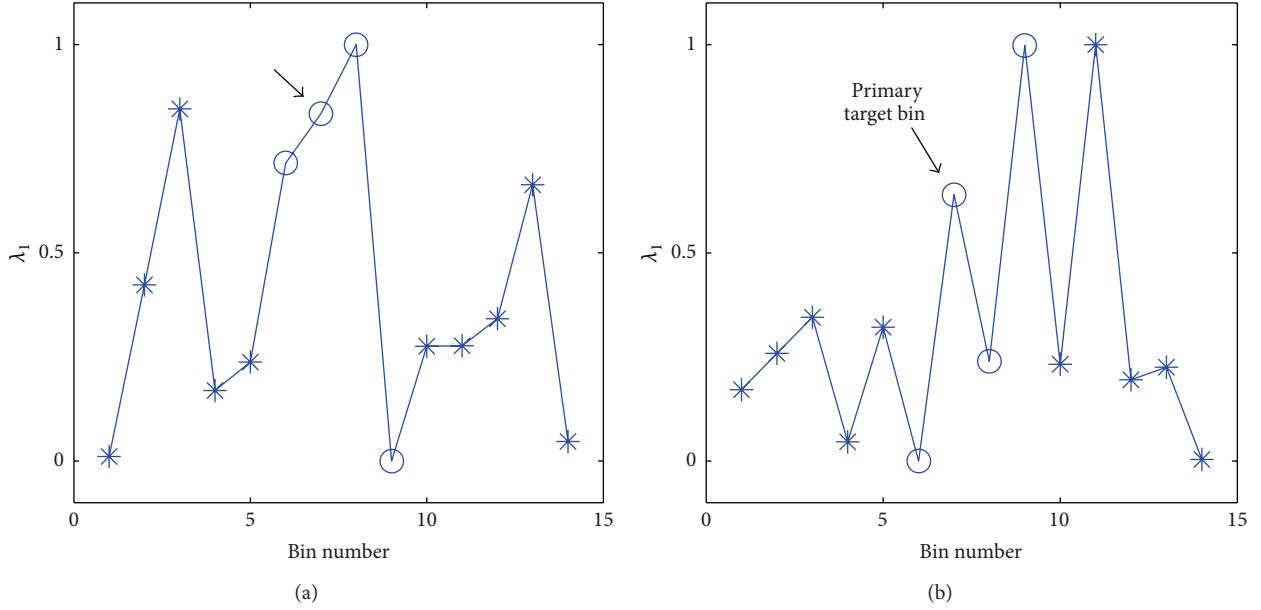


FIGURE 7: ((a) and (b)) The variations of the Lyapunov exponent λ estimated by conventional methods versus the 14 range bins for two HH measurements. Open circles denote the range bins with target, while asterisks denote the bins without target. The primary target bin is explicitly indicated by an arrow. For convenience, the largest λ_c is taken as 1 unit.

Assuming that $d_{i+1} \sim d'_i e^{\lambda_1(t_{i+1}-t_i)}$, the rate of divergence of the trajectory, λ_1 , over a time interval of $t_{i+1} - t_i$ is then

$$\frac{\ln(d_{i+1}/d'_i)}{t_{i+1} - t_i}. \quad (17)$$

To ensure that the separation between the two trajectories is always small, when d_{i+1} exceeds certain threshold value, it has to be renormalized; a new point in the direction of the vector of d_{i+1} is picked up so that d'_{i+1} is very small compared to the size of the attractor. After n repetitions of stretching and renormalizing the spacing, one obtains the following formula:

$$\begin{aligned} \lambda_1 &= \sum_{i=1}^{n-1} \left[\frac{t_{i+1} - t_i}{\sum_{i=1}^{n-1} (t_{i+1} - t_i)} \right] \left[\frac{\ln(d_{i+1}/d'_i)}{t_{i+1} - t_i} \right] \\ &= \frac{\sum_{i=1}^{n-1} \ln(d_{i+1}/d'_i)}{t_n - t_1}. \end{aligned} \quad (18)$$

Figures 7(a) and 7(b) show representative λ_1 for two measurements, where the largest λ_1 among the 14 bins of a measurement is taken as 1 unit. Clearly, λ_1 does not have any power of detecting targets within sea clutter. This suggests that λ_1 only characterizes noise in sea clutter, whether they hit targets or not. Therefore, the effectiveness of using (10) to detect targets within sea clutter suggests that SDLE is fundamentally different from the conventional Lyapunov exponent.

4. Concluding Remarks

To better characterize the nonstationary and multiscale nature of sea clutter, in this study, we have employed SDLE to analyze low-grazing angle sea clutter. We are able to directly show that sea clutter is not chaotic and reveal a new scaling law suggesting noisy dynamics for sea clutter. The new scaling law has an interesting interpretation in terms of intrinsic predictability of sea clutter. The fact that it can be used to accurately detect targets within sea clutter implies that it is a fundamental feature of sea clutter.

References

- [1] J. B. Gao, Y. H. Cao, W. W. Tung, and J. Hu, *Multiscale Analysis of Complex Time Series: Integration of Chaos and Random Fractal Theory, and Beyond*, John Wiley & Sons, Hoboken, NJ, USA, 2007.
- [2] L. B. Wetzel, "Sea clutter," in *Radar Handbook*, M. I. Skolnik, Ed., chapter 13, McGraw-Hill, New York, NY, USA, 2nd edition, 1990.
- [3] F. A. Fay, J. Clarke, and R. S. Peters, "Weibull distribution applied to sea-clutter," in *Proceedings of the IEE International Radar'77 Conference*, pp. 101–104, London, UK, October 1977.
- [4] F. E. Nathanson, *Radar Design Principles*, McGraw-Hill, New York, NY, USA, 1969.
- [5] G. V. Trunk and S. F. George, "Detection of targets in non-Gaussian sea clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 6, no. 5, pp. 620–628, 1970.
- [6] H. C. Chan, "Radar sea-clutter at low grazing angles," *IEE Proceedings F*, vol. 137, no. 2, pp. 102–112, 1990.
- [7] E. Jakeman and P. N. Pusey, "A model for non-Rayleigh sea echo," *IEEE Transactions on Antennas and Propagation*, vol. 24, no. 6, pp. 806–814, 1976.

- [8] K. D. Ward, C. J. Baker, and S. Watts, "Maritime surveillance radar—part 1: radar scattering from the ocean surface," *IEE Proceedings F*, vol. 137, no. 2, pp. 51–62, 1990.
- [9] T. J. Nohara and S. Haykin, "Canadian east coast radar trials and the K -distribution," *IEE Proceedings F*, vol. 138, no. 2, pp. 80–88, 1991.
- [10] S. Sayama and M. Sekine, "Log-normal, log-Weibull and K -distributed sea clutter," *IEICE Transactions on Communications*, vol. 85, no. 7, pp. 1375–1381, 2002.
- [11] F. Gini, "Performance analysis of two structured covariance matrix estimators in compound-Gaussian clutter," *Signal Processing*, vol. 80, no. 2, pp. 365–371, 2000.
- [12] R. J. A. Tough and K. D. Ward, "Diffusive and related models of some properties of sea clutter," *Journal of Physics Condensed Matter*, vol. 14, no. 33, pp. 7737–7748, 2002.
- [13] J. Hu, W. W. Tung, and J. B. Gao, "A new way to model nonstationary sea clutter," *IEEE Signal Processing Letters*, vol. 16, no. 2, pp. 129–132, 2009.
- [14] F. Gini, M. Montanari, and L. Verrazzani, "Maximum likelihood, ESPRIT, and periodogram frequency estimation of radar signals in K -distributed clutter," *Signal Processing*, vol. 80, no. 6, pp. 1115–1126, 2000.
- [15] F. Geni, A. Farina, and M. Montanari, "Vector subspace detection in compound-Gaussian clutter part II: performance analysis," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 4, pp. 1312–1323, 2002.
- [16] M. Martorella, F. Berizzi, and E. D. Mese, "On the fractal dimension of sea surface backscattered signal at low grazing angle," *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 5, pp. 1193–1204, 2004.
- [17] J. Chen, T. K. Y. Lo, H. Leung, and J. Litva, "The use of fractals for modeling EM waves scattering from rough sea surface," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, no. 4, pp. 966–972, 1996.
- [18] F. Berizzi, E. Dalle Mese, and M. Martorella, "A sea surface fractal model for ocean remote sensing," *International Journal of Remote Sensing*, vol. 25, no. 7–8, pp. 1265–1270, 2004.
- [19] F. Berizzi and E. Dalle-Mese, "Scattering from a 2-D sea fractal surface: fractal analysis of the scattered signal," *IEEE Transactions on Antennas and Propagation*, vol. 50, no. 7, pp. 912–925, 2002.
- [20] G. Franceschetti, A. Iodice, M. Migliaccio, and D. Riccio, "Scattering from natural rough surfaces modeled by fractional Brownian motion two-dimensional processes," *IEEE Transactions on Antennas and Propagation*, vol. 47, no. 9, pp. 1405–1415, 1999.
- [21] T. Lo, H. Leung, J. Litva, and S. Haykin, "Fractal characterization of sea-scattered signals and detection of sea-surface targets," *IEE Proceedings F*, vol. 140, no. 4, pp. 243–250, 1993.
- [22] C. P. Lin, M. Sano, S. Sayama, and M. Sekine, "Detection of radar targets embedded in sea ice and sea clutter using fractals, wavelets, and neural networks," *IEICE Transactions on Communications*, vol. 83, no. 9, pp. 1916–1929, 2000.
- [23] J. B. Gao, J. Hu, W. W. Tung, Y. H. Cao, N. Sarshar, and V. P. Roychowdhury, "Assessment of long-range correlation in time series: how to avoid pitfalls," *Physical Review E*, vol. 73, no. 1, Article ID 016117, 2006.
- [24] J. Hu, J. B. Gao, F. L. Posner, Y. I. Zheng, and W. W. Tung, "Target detection within sea clutter: a comparative study by fractal scaling analyses," *Fractals*, vol. 14, no. 3, pp. 187–204, 2006.
- [25] J. Hu, W. W. Tung, and J. B. Gao, "Detection of low observable targets within sea clutter by structure function based multifractal analysis," *IEEE Transactions on Antennas and Propagation*, vol. 54, no. 1, pp. 136–143, 2006.
- [26] S. Haykin and S. Puthusserpady, "Chaotic dynamics of sea clutter," *Chaos*, vol. 7, no. 4, pp. 777–802, 1997.
- [27] S. Haykin, *Chaotic Dynamics of Sea Clutter*, vol. 17 of *Adaptive and Learning Systems for Signal Processing, Communications, and Control*, John Wiley & Sons, Hoboken, NJ, USA, 1999.
- [28] S. Haykin, R. Barker, and B. W. Currie, "Uncovering nonlinear dynamics—the case study of sea clutter," *Proceedings of the IEEE*, vol. 90, no. 5, pp. 860–881, 2002.
- [29] M. Davies, "Looking for non-linearities in sea clutter," in *Proceedings of the IEE Radar and Sonar Signal Processing*, Peebles, UK, July 1998.
- [30] M. R. Cowper, B. Mulgrew, and C. P. Unsworth, "Investigation into the use of nonlinear predictor networks to improve the performance of maritime surveillance radar target detectors," *IEE Proceedings—Radar, Sonar and Navigation*, vol. 148, no. 3, pp. 103–111, 2001.
- [31] M. R. Cowper and B. Mulgrew, "Nonlinear processing of high resolution radar sea clutter," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN '99)*, pp. 2633–2638, July 1999.
- [32] J. L. Noga, *Bayesian state-space modelling of spatio-temporal non-Gaussian radar returns [Ph.D. thesis]*, Cambridge University, 1998.
- [33] C. P. Unsworth, M. R. Cowper, S. McLaughlin, and B. Mulgrew, "False detection of chaotic behavior in the stochastic compound k -distribution model of radar sea clutter," in *Proceedings of the 10th IEEE Workshop on Statistical Signal and Array Processing (SSAP '00)*, pp. 296–300, August 2000.
- [34] C. P. Unsworth, M. R. Cowper, S. McLaughlin, and B. Mulgrew, "Re-examining the nature of radar sea clutter," *IEE Proceedings—Radar, Sonar and Navigation*, vol. 149, no. 3, pp. 105–114, 2002.
- [35] J. B. Gao and K. Yao, "Multifractal features of sea clutter," in *Proceedings of the IEEE Radar Conference*, pp. 500–505, Long Beach, Calif, USA, April 2002.
- [36] J. B. Gao, S. K. Hwang, H. F. Chen, Z. Kang, K. Yao, and J. M. Liu, "Can sea clutter and indoor radio propagation be modeled as strange attractors?" in *Proceedings of the 7th Experimental Chaos Conference*, vol. 676, pp. 154–161, San Diego, Calif, USA, August 2002.
- [37] M. McDonald and A. Damini, "Limitations of nonlinear chaotic dynamics in predicting sea clutter returns," *IEE Proceedings—Radar, Sonar and Navigation*, vol. 151, no. 2, pp. 105–113, 2004.
- [38] T. Thayaparan and S. Kennedy, "Detection of a manoeuvring air target in sea-clutter using joint time-frequency analysis techniques," *IEE Proceedings—Radar, Sonar and Navigation*, vol. 151, no. 1, pp. 19–30, 2004.
- [39] G. Davidson and H. D. Griffiths, "Wavelet detection scheme for small targets in sea clutter," *Electronics Letters*, vol. 38, no. 19, pp. 1128–1130, 2002.
- [40] T. K. Bhattacharya and S. Haykin, "Neural network-based adaptive radar detection scheme for small ice targets in sea clutter," *Electronics Letters*, vol. 28, no. 16, pp. 1528–1529, 1992.
- [41] T. K. Bhattacharya and S. Haykin, "Neural network-based radar detection for an ocean environment," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 2, pp. 408–420, 1997.

- [42] H. Leung, N. Dubash, and N. Xie, "Detection of small objects in clutter using a GA-RBF neural network," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 1, pp. 98–118, 2002.
- [43] N. Xie, H. Leung, and H. Chan, "A multiple-model prediction approach for sea clutter modeling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 6, pp. 1491–1502, 2003.
- [44] C. P. Lin, M. Sano, S. Obi, S. Sayama, and M. Sekine, "Detection of oil leakage in SAR images using wavelet feature extractors and unsupervised neural classifiers," *IEICE Transactions on Communications*, vol. 83, no. 9, pp. 1955–1962, 2000.
- [45] J. B. Gao, J. Hu, W. W. Tung, and Y. H. Cao, "Distinguishing chaos from noise by scale-dependent Lyapunov exponent," *Physical Review E*, vol. 74, no. 6, Article ID 066204, 9 pages, 2006.
- [46] J. B. Gao, J. Hu, and W. W. Tung, "Complexity measures of brain wave dynamics," *Cognitive Neurodynamics*, vol. 5, no. 2, pp. 171–182, 2011.
- [47] J. B. Gao, J. Hu, and W. W. Tung, "Entropy measures for biological signal analyses," *Nonlinear Dynamics*, vol. 68, no. 3, pp. 431–444, 2012.
- [48] J. Hu, J. B. Gao, W. W. Tung, and Y. H. Cao, "Multiscale analysis of heart rate variability: a comparison of different complexity measures," *Annals of Biomedical Engineering*, vol. 38, no. 3, pp. 854–864, 2010.
- [49] J. Hu, J. B. Gao, and W. W. Tung, "Characterizing heart rate variability by scale-dependent Lyapunov exponent," *Chaos*, vol. 19, no. 2, Article ID 028506, 6 pages, 2009.
- [50] J. B. Gao, J. Hu, W. W. Tung, and E. Blasch, "Multiscale analysis of physiological data by scale-dependent Lyapunov exponent," *Frontiers in Fractal Physiology*, 2012.
- [51] J. B. Gao, B. M. Gurbaxani, J. Hu et al., "Multiscale analysis of heart rate variability in nonstationary environments," *Frontiers in Computational Physiology and Medicine*, 2013.
- [52] J. B. Gao, J. Hu, W. W. Tung, and Y. Zheng, "Multiscale analysis of economic time series by scale-dependent Lyapunov exponent," *Quantitative Finance*, vol. 13, no. 2, pp. 265–274, 2013.
- [53] D. A. Ryan and G. R. Sarson, "The geodynamo as a low-dimensional deterministic system at the edge of chaos," *Euro-physics Letters*, vol. 83, no. 4, Article ID 49001, 6 pages, 2008.
- [54] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence*, D. A. Rand and L. S. Young, Eds., vol. 898 of *Lecture Notes in Mathematics*, pp. 366–381, Springer, Berlin, Germany, 1981.
- [55] J. B. Gao and Z. M. Zheng, "Direct dynamical test for deterministic chaos and optimal embedding of a chaotic time series," *Physical Review E*, vol. 49, no. 5, pp. 3807–3814, 1994.
- [56] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, "Determining Lyapunov exponents from a time series," *Physica D*, vol. 16, no. 3, pp. 285–317, 1985.
- [57] P. Grassberger and I. Procaccia, "Characterization of strange attractors," *Physical Review Letters*, vol. 50, no. 5, pp. 346–349, 1983.
- [58] J. B. Gao, S. K. Hwang, and J. M. Liu, "When can noise induce chaos?" *Physical Review Letters*, vol. 82, no. 6, pp. 1132–1135, 1999.
- [59] J. B. Gao, G. C. Chen, S. K. Hwang, and J. M. Liu, "Noise-induced chaos," *International Journal of Modern Physics B*, vol. 13, no. 28, pp. 3283–3305, 1999.
- [60] S. K. Hwang, J. B. Gao, and J. M. Liu, "Noise-induced chaos in an optically injected semiconductor laser model," *Physical Review E*, vol. 61, no. 5, pp. 5162–5170, 2000.
- [61] R. Mannella, "Integration of stochastic differential equations on a computer," *International Journal of Modern Physics C*, vol. 13, no. 9, pp. 1177–1194, 2002.

Research Article

Restoration of Natural Frequency of Cracked Cantilever Beam Using CNT Composite Patch: A Finite Element Study

Mahmoud Nadim Nahas and Mahmoud Ali Alzahrani

Mechanical Engineering Department, King Abdulaziz University, P.O. Box 80204, Jeddah 21589, Saudi Arabia

Correspondence should be addressed to Mahmoud Nadim Nahas; mnahas@kau.edu.sa

Received 16 September 2013; Revised 21 October 2013; Accepted 1 November 2013

Academic Editor: Dumitru Baleanu

Copyright © 2013 M. N. Nahas and M. A. Alzahrani. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Cyclic loadings cause fatigue to the elements of machines leading to crack initiation and propagation. This phenomenon decreases the age of the elements. In particular, cracks decrease the stiffness of the parts and lower the parts natural frequency, leading to failure under normal working conditions. This paper introduces a new application to carbon nanotube (CNT) composites in the repairing process of a cracked specimen to restore the natural frequency of the specimen. Commonly, patches are made of high strength and high stiffness materials. This paper shows that even low stiffness materials, such as epoxy reinforced with CNT, can contribute to the repair of a cracked specimen. A 2D finite element (FE) simulation is used to study the effects of bonding CNT composite patches over the crack location to repair cracked metal specimens. The effects of the patch thickness, length, and CNTs weight concentration ratio are investigated. Results showed an increase in the natural frequency of 31% compared to the cracked specimen at a crack depth of 70% of the beam depth and at a distance of 20% of the total beam length from the support.

1. Introduction

Vibrational faults in machines are of major concern in the engineering field. In machines, cracked parts due to fatigue loading can cause a drop in the part natural frequency. This drop can lead to resonant vibrations under normal loading conditions, which may cause failure to the part. A lot of work had been conducted on the change in the natural frequency of a cracked beam compared to a healthy beam. Khajji et al. [1] constructed a closed form analytical solution to the change in the natural frequency of a cracked uniform beam. The crack was modeled using a massless, linear, elastic, and rotational spring under six different boundary conditions and different parameters (such as crack location and crack depth). Kim and Stubbs [2] and Lee [3] conducted an experimental work on the detection of crack locations and its size by relating the fractional changes in the natural frequency of the cracked beam to the healthy beam due to the presence of cracks. Zheng and Kessissoglou [4] studied the free vibration of a cracked beam using the finite elements analysis (FEA) to obtain the natural frequencies and mode shapes of the

beam. Loya et al. [5] studied the natural frequency of simply supported cracked Timoshenko beam by modeling the cracked region with a massless torsional and normal spring (for the vertical displacement and rotational displacement). Zhong and Oyadiji [6] constructed an analytical model to help in the prediction of the natural frequencies of a cracked simply supported beam with a stationary roving mass. The roving mass helps in detecting the presence of cracks by roving the mass from one end of the beam to another. Since cracks cause a reduction in the stiffness of the beam, this causes a reduction in the beam natural frequency when the roving mass is located near the crack. Matbuly et al. [7] studied the natural frequency of a functionally graded cracked beam resting on a Winkler-Pasternak foundation using the line spring model for formulating and applying the method of differential quadrature to solve it. For the strength of bonded patches, much research had been conducted, with different materials, to evaluate its effectiveness under various conditions. Osnes et al. [8] had studied the strength of bonded joints using the traditional strength of materials approach, which did not agree with experimental results, and inelastic

TABLE 1: Mechanical properties of epoxy and aluminum specimens.

Mechanical properties	Steel	Epoxy			
		0 wt.%	2 wt.%	4 wt.%	6 wt.%
E (GPa)	210	3.11	4.2	4.7	4.98
ν	0.3	0.35	0.3492	0.3476	0.3468
ρ (Kg/m ³)	7860	1170	1181	1191	1202

fracture mechanics, which showed good agreement with the experimental results. Hosseini-Toudeshky et al. [9] studied how the curing temperature affects the residual stresses in fiber metal laminate (FML) patches, which showed no effect on the repairing process for the cracked panels. Nahas [10] performed an experimental investigation on the repairing effects of fiber composite patches on the fatigue life of a cracked aluminum specimen. Ariaei et al. [11] used piezoelectric patches to repair cracked beams by applying an external voltage to actuate the piezoelectric patches bonded on the beam. Then the patches reduce the singularity produced by the crack tip. This helped in decreasing the maximum deflection of the beam, due to the reduction of the beam stiffness, to approach the maximum deflection of a healthy beam. Brown et al. [12, 13] used self-healing epoxy matrix composite that incorporates healing agent to hinder fatigue cracks. Self-healing was also used by Blaiszik et al. [14] and Williams et al. [15].

In this paper, a FEA is conducted to examine the repair of the natural frequency of a cracked steel beam (at different locations and depth) by adhesively bonding an epoxy patch (with and without multiwalled carbon nanotubes (MWCNTs) reinforcement). The properties of the steel specimen, epoxy, and multiwalled carbon nanotubes (MWCNTs) are discussed in Section 2. The modeling of the parts and the boundary conditions are discussed in Section 3. The FEA results of the analysis are summarized in Section 4. Discussions of the results are given in Section 4 and conclusions of the results are given in Section 5.

2. Mechanical Properties

CNTs are known for their high mechanical properties [16]. They have shown great potential in the reinforcement of composite materials. Many research works were carried out to predict the mechanical properties of CNT reinforced composite [17–22]. Omidi et al. [23] experimented on the effects of adding MWCNTs into an epoxy patch on Young's modulus and strength of the CNT/Epoxy composite. These results were used with the work done by Rokni et al. [24, 25] to construct a finite elements model to determine the optimum distribution of CNTs to obtain the highest natural frequency in a microbeam. Different weight percentage (wt.%) of CNTs was used (from 0.5% to 10%). The mechanical properties used in this paper (Young's modulus E , density ρ , and Poisson's ratio) of the epoxy patch (with different wt.% of CNTs) and the steel specimen are shown in Table 1.

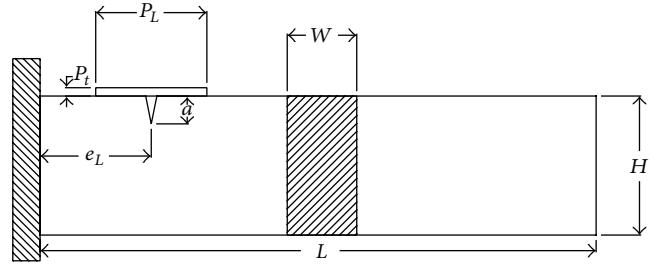


FIGURE 1: Beam dimensions.

3. Modeling and FE Idealization

A 2D plane stress FE model was constructed in ABAQUS/Standard with dimensions similar to those in [1] and was verified with the results in the literature. Figure 1 shows the dimensions of the beam used in the FE analysis where L is the beam length which is 100 mm, H is the beam depth which is 25 mm, and W is the beam width which is 12.5 mm. The ratio a/H (the ratio of the crack depth to the beam depth) was taken as 0.7 and the crack position e_L was taken at a location of 20% of the beam total length away from the support location. These values, according to [1], show large decrease in the beam original natural frequency compared to the other values used in the work. The epoxy patch has a length of P_L and thickness of P_t .

Figure 2 shows the FE models and the boundary conditions used in the analysis. The nodes to the left end were constrained from movement in the x and y directions, while the nodes to the right end were left unconstrained. An 8-node quadrilateral element was used in the model. The repair of the natural frequency was investigated under different patch thickness and length (P_t and P_L) values with different CNT wt.% (2 wt.%, 4 wt.% and 6 wt.%). P_t values were taken from 0.5 to 3 mm, while P_L was taken from 10 to 30 mm. Then the effects of the crack depth and location were investigated when P_t was 1.5 mm and P_L was 20 mm.

4. Results and Discussions

The first natural frequency was obtained for the model of the original beam (the uncracked beam), and then a crack with a value of $a/H = 0.7$ and $e = 0.2$ was introduced. The natural frequency was then obtained from ABAQUS for the cracked beam, firstly without the epoxy patch and then with the epoxy patch. Figure 3 shows the FE contour plot of the first mode shape of the beam. The epoxy patch was added to the model at different P_t values (0.5, 1, 1.5, 2, 2.5, and 3 mm). Figure 4 shows the graph of the increase of natural frequency of the cracked beam with different P_t and CNT wt.%. The graph shows that, with the increase of the patch thickness, more repairing is observed in the natural frequency of the patched cracked beam in comparison to the unrepaired beam. A patch thickness of 3 mm increased the natural frequency of the cracked beam by 41% at 6 wt.% CNT, but the increase in P_t can lead to stress concentrations, so P_t was taken as 1.5 mm for the rest of the analysis.

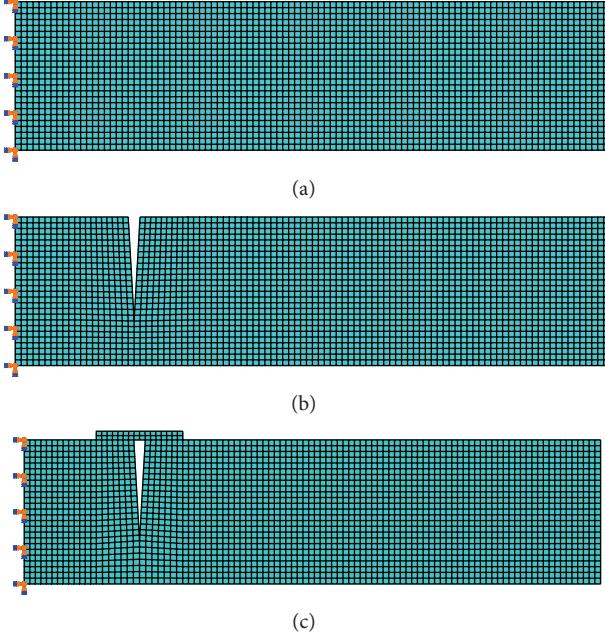


FIGURE 2: FE model boundary conditions, (a) original beam, (b) cracked beam, and (c) epoxy patched beam.

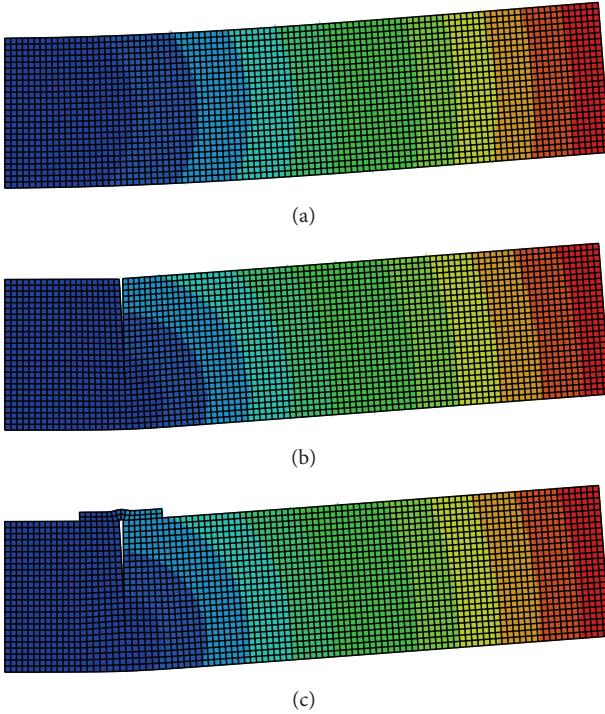


FIGURE 3: FE first mode shape contour plot, (a) original beam, (b) cracked beam, and (c) epoxy patched beam.

Different P_L values (10, 15, 20, 25, and 30 mm) were used in order to investigate the effects of the patch length on the increase in the natural frequency of the beam. Figure 5 shows that P_L has no effect on the increase in the natural frequency of the cracked beam. Only CNT's concentration increased the natural frequency, but short length patches are weaker than

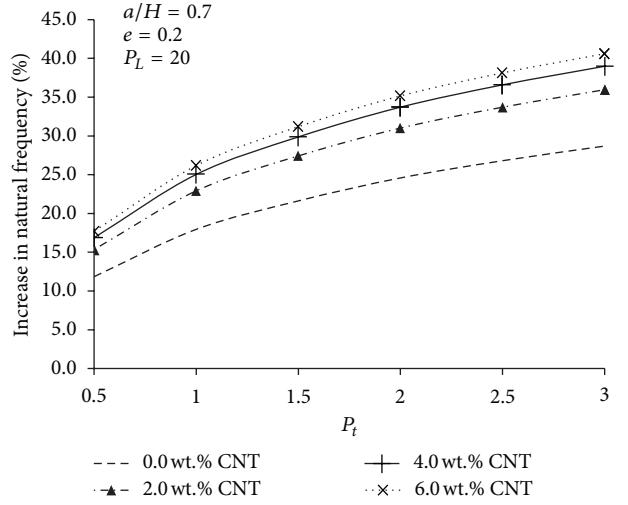


FIGURE 4: Increase in cracked beam natural frequency with different P_t and wt.% of CNT.

long ones, due to the small interfacial area between the patch and the beam, so P_L was taken as 20 mm in further analysis.

To investigate the changes in the natural frequency of the cracked beam with the repairing patch at different crack depths (a/H) and crack locations (e), a patch with $P_t = 1.5$ mm and $P_L = 20$ mm was used. a/H was changed from 20% to 70% (0.2, 0.35, 0.5, 0.6, and 0.7) and e from 20% to 80% (0.2, 0.4, 0.5, 0.6, and 0.8). Figure 6 shows the relative natural frequency of the cracked beam with and without CNT composite patches at different crack depths. It is clear from the graph that at small crack depths, the effects of the epoxy patches are negligible since the decrease in the natural frequency is also low. Figure 7 shows the increase in the natural frequency of the repaired beam in comparison to the unrepairs one. An increase of 31% in the natural frequency is observed at 6 wt.% CNTs and $e = 0.2$. Also, a noticeable increase in the repairing takes place when a/H goes above 0.5.

Figure 8 shows the relative natural frequency but with different crack locations. The same is also observed here; the repair has no significant change in the crack locations that produce low decrease in the natural frequency of the beam, while at $e = 0.2$, the relative natural frequency of the repaired beam increased with increasing CNTs wt%. Figure 9 shows the increase in natural frequency with different e values; the higher the CNTs wt%, the higher the increase in the natural frequency of the beam.

It is noticed from the graphs that increasing the contents of CNT in the epoxy matrix does not always cause a noticeable increase in the repair of the cracked beam natural frequency. Figure 10 shows the increase in the natural frequency of the cracked beam with different concentration ratios at $a/H = 0.7$, $e = 0.2$, $P_L = 20$ mm, and $P_t = 1.5$ mm. The graph shows that, after 4.0 wt.% CNTs, the increase in the natural frequency converges to 31%; further increase will cause no significant change in the natural frequency of the repaired beam.

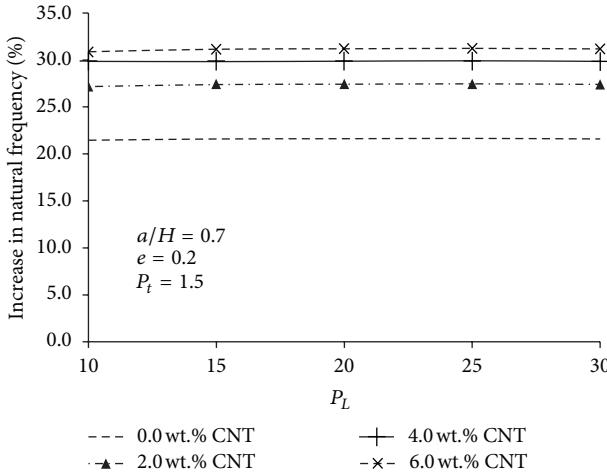


FIGURE 5: Increase in cracked beam natural frequency with different P_L and wt.% of CNT.

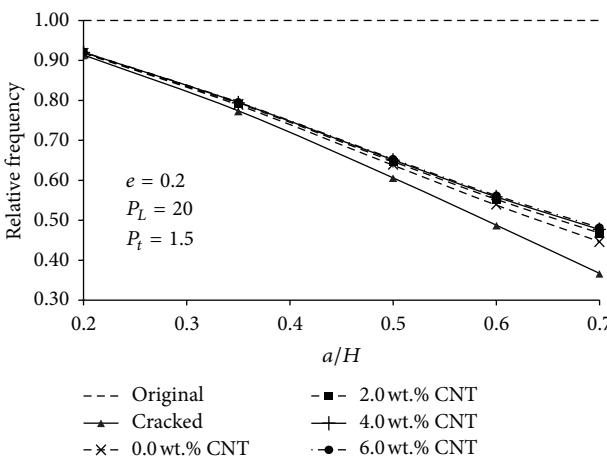


FIGURE 6: Relative natural frequency at different a/H and wt.% of CNT.

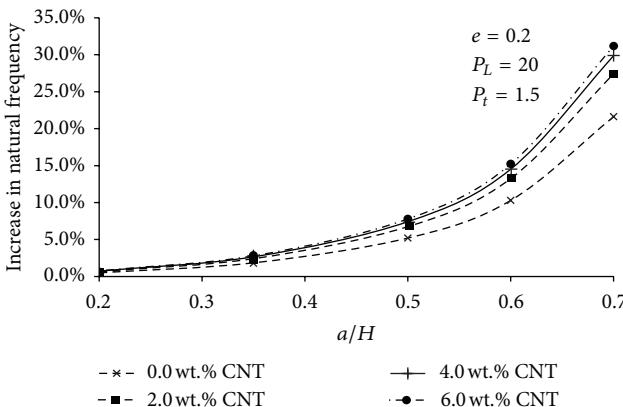


FIGURE 7: Increase in cracked beam natural frequency at different a/H and wt.% of CNT.

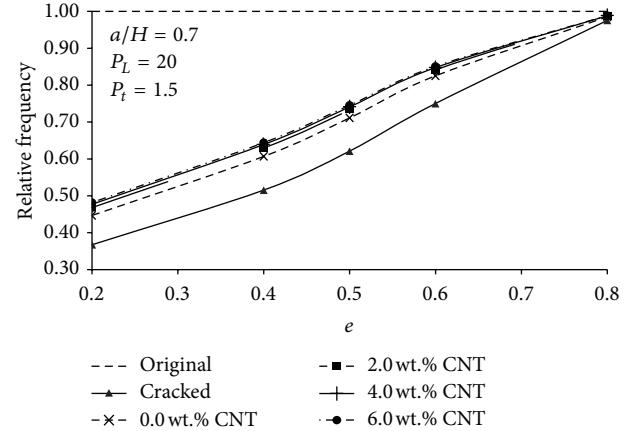


FIGURE 8: Relative natural frequency at different e and wt.% of CNT.

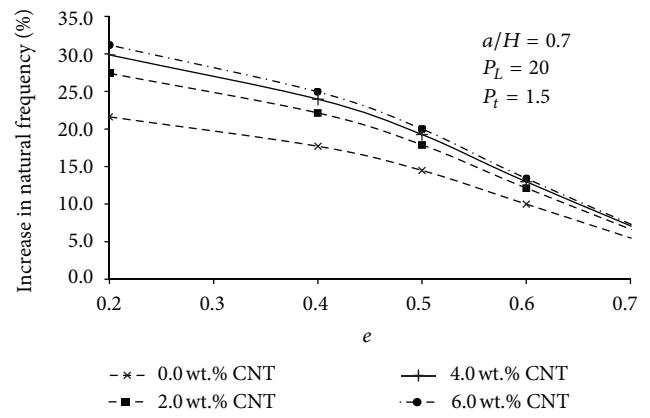


FIGURE 9: Increase in cracked beam natural frequency at different e and wt.% of CNT.

The present work is a start for more work on the repair of materials using CNT composite patches. In particular, the authors intend to conduct the following two studies.

- Experimental investigation of cracked specimens repaired by CNT composite patches to restore the natural frequency of the specimens. The CNT patch will be bonded over the crack location to repair it. The specimens will then be subjected to cyclic loadings to study the influence of the repair on the natural frequency. Different patch thickness and patch length values will be considered. In addition, different CNTs weight concentration ratios will be used. The facilities and equipment available at the Nanocenter of King Abdulaziz University will be used in this study.
- On the theoretical side, the authors will conduct further analysis on the stress intensity factor at the crack tip as was done by Bachir et al. [26] and will use the J -integral concept of Eshelby-Rice to add results for fractal cracks. The J -integral for a fractal is path dependent [27]. This is why a J -integral fractal should be the rate of release of potential energy per unit of measurement of the fractal crack growth [28]. The

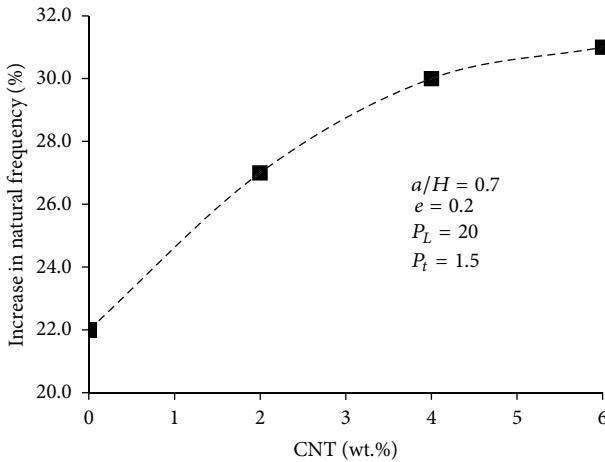


FIGURE 10: Increase in natural frequency with different CNT ratios.

authors also intend to compare how well will the CNT reinforced epoxy patches perform when compared with highly stiff composites such as carbon epoxy [29], boron epoxy [30, 31], and graphite epoxy [32].

5. Conclusion

A numerical FEA using ABAQUS/Standard was carried out to observe the repair of the natural frequency of a cracked steel cantilever beam with an adhesively bonded epoxy patch (with and without CNT reinforcement at different wt.%). The first natural frequency was obtained for the original beam and then the natural frequency of the cracked beam ($a/H = 0.7$ and $e = 0.2$) with different P_t and P_L was obtained to determine the suitable values of P_t and P_L . It is shown that P_L has no effect on the repair of the natural frequency, while as P_t increases, the natural frequency increases. The effect of a/H and e was investigated at P_t and P_L values of 1.5 and 20 mm, respectively. The results show that the repair had insignificant effect on locations and depths that produced low decrease in the beam natural frequency no matter which CNT wt.% was used. But at locations where there was a significant drop in the natural frequency, the effects of CNT's wt.% became insignificant at concentrations more than 4.0 wt.%. Further investigation can be carried out to see the effect of layered epoxy patches with different P_L to reduce stress concentration when the patch thickness increases as in [8]. Experimental validation of the findings will be carried out in more details as the facilities and equipment are available for the authors. Further investigation will be done by inspecting the effects of the patches on the strain energy release rate at the crack tip which can be used as a supporting result on the claim of the effects of the epoxy patches.

Acknowledgment

This paper was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under Grant no.

(21-135-D1432). The authors, therefore, acknowledge with thanks DSR technical and financial support.

References

- [1] N. Khajri, M. Shafiei, and M. Jalalpour, "Closed-form solutions for crack detection problem of Timoshenko beams with various boundary conditions," *International Journal of Mechanical Sciences*, vol. 51, no. 9-10, pp. 667–681, 2009.
- [2] J. T. Kim and N. Stubbs, "Crack detection in beam-type structures using frequency data," *Journal of Sound and Vibration*, vol. 259, no. 1, pp. 145–160, 2003.
- [3] J. Lee, "Identification of multiple cracks in a beam using natural frequencies," *Journal of Sound and Vibration*, vol. 320, no. 3, pp. 482–490, 2009.
- [4] D. Y. Zheng and N. J. Kessissoglou, "Free vibration analysis of a cracked beam by finite element method," *Journal of Sound and Vibration*, vol. 273, no. 3, pp. 457–475, 2004.
- [5] J. A. Loya, L. Rubio, and J. Fernández-Sáez, "Natural frequencies for bending vibrations of Timoshenko cracked beams," *Journal of Sound and Vibration*, vol. 290, no. 3–5, pp. 640–653, 2006.
- [6] S. Zhong and S. O. Oyadiji, "Analytical predictions of natural frequencies of cracked simply supported beams with a stationary roving mass," *Journal of Sound and Vibration*, vol. 311, no. 1–2, pp. 328–352, 2008.
- [7] M. S. Matbuly, O. Ragh, and M. Nassar, "Natural frequencies of a functionally graded cracked beam using the differential quadrature method," *Applied Mathematics and Computation*, vol. 215, no. 6, pp. 2307–2316, 2009.
- [8] H. Osnes, D. McGeorge, J. R. Weitzenböck, and G. O. Guthu, "Predicting failure of bonded patches using a fracture mechanics approach," *International Journal of Adhesion and Adhesives*, vol. 37, pp. 102–111, 2012.
- [9] H. Hosseini-Toudeshky, M. Sadighi, and A. ojdani, "Effects of curing thermal residual stresses on fatigue crack propagation of aluminum plates repaired by FML patches," *Composite Structures*, vol. 100, pp. 154–162, 2013.
- [10] M. N. Nahas, "Experimental investigation of fatigue of cracked aluminum specimens repaired with fibre composite patches," *Journal of Reinforced Plastics and Composites*, vol. 11, no. 8, pp. 932–938, 1992.
- [11] A. Ariaei, S. Ziae Rad, and M. Ghayour, "Repair of a cracked Timoshenko beam subjected to a moving mass using piezoelectric patches," *International Journal of Mechanical Sciences*, vol. 52, no. 8, pp. 1074–1091, 2010.
- [12] E. N. Brown, S. R. White, and N. R. Sottos, "Retardation and repair of fatigue cracks in a microcapsule toughened epoxy composite. Part I: manual infiltration," *Composites Science and Technology*, vol. 65, no. 15–16, pp. 2466–2473, 2005.
- [13] E. N. Brown, S. R. White, and N. R. Sottos, "Retardation and repair of fatigue cracks in a microcapsule toughened epoxy composite. Part II: in situ self-healing," *Composites Science and Technology*, vol. 65, no. 15–16, pp. 2474–2480, 2005.
- [14] B. J. Blaiszik, N. R. Sottos, and S. R. White, "Nanocapsules for self-healing materials," *Composites Science and Technology*, vol. 68, no. 3–4, pp. 978–986, 2008.
- [15] H. R. Williams, R. S. Trask, and I. P. Bond, "Self-healing sandwich panels: restoration of compressive strength after impact," *Composites Science and Technology*, vol. 68, no. 15–16, pp. 3171–3177, 2008.

- [16] S. Iijima, "Helical microtubules of graphitic carbon," *Nature*, vol. 354, no. 6348, pp. 56–58, 1991.
- [17] K. I. Tserpes, P. Papanikos, G. Labeas, and S. G. Pantelakis, "Multi-scale modeling of tensile behavior of carbon nanotube-reinforced composites," *Theoretical and Applied Fracture Mechanics*, vol. 49, no. 1, pp. 51–60, 2008.
- [18] M. M. Shokrieh and R. Rafiee, "Investigation of nanotube length effect on the reinforcement efficiency in carbon nanotube based composites," *Composite Structures*, vol. 92, no. 10, pp. 2415–2420, 2010.
- [19] M. M. Shokrieh and R. Rafiee, "On the tensile behavior of an embedded carbon nanotube in polymer matrix with non-bonded interphase region," *Composite Structures*, vol. 92, no. 3, pp. 647–652, 2010.
- [20] M. Kulkarni, D. Carnahan, K. Kulkarni, D. Qian, and J. L. Abot, "Elastic response of a carbon nanotube fiber reinforced polymeric composite: a numerical and experimental study," *Composites Part B*, vol. 41, no. 5, pp. 414–421, 2010.
- [21] G. I. Giannopoulos, S. K. Georgantzinos, and N. K. Anifantis, "A semi-continuum finite element approach to evaluate the Young's modulus of single-walled carbon nanotube reinforced composites," *Composites Part B*, vol. 41, no. 8, pp. 594–601, 2010.
- [22] S. K. Georgantzinos, G. I. Giannopoulos, and N. K. Anifantis, "Investigation of stress-strain behavior of single walled carbon nanotube/rubber composites by a multi-scale finite element method," *Theoretical and Applied Fracture Mechanics*, vol. 52, no. 3, pp. 158–164, 2009.
- [23] M. Omidi, D. T. Hossein Rokni, A. S. Milani, R. J. Seethaler, and R. Arasteh, "Prediction of the mechanical characteristics of multi-walled carbon nanotube/epoxy composites using a new form of the rule of mixtures," *Carbon*, vol. 48, no. 11, pp. 3218–3228, 2010.
- [24] H. Rokni, A. S. Milani, and R. J. Seethaler, "Maximum natural frequencies of polymer composite micro-beams by optimum distribution of carbon nanotubes," *Materials and Design*, vol. 32, no. 6, pp. 3389–3398, 2011.
- [25] H. Rokni, A. S. Milani, and R. J. Seethaler, "2D optimum distribution of carbon nanotubes to maximize fundamental natural frequency of polymer composite micro-beams," *Composites Part B*, vol. 43, no. 2, pp. 779–785, 2012.
- [26] B. Bachir Bouadjra, W. Oudad, A. Albedah, F. Benyahia, and M. Belhouari, "Effects of the adhesive disband on the performances of bonded composite repairs in aircraft structures," *Materials and Design*, vol. 37, pp. 89–95, 2012.
- [27] A. Yavari, S. Sarkani, and E. T. Moyer Jr., "The mechanics of self-similar and self-affine fractal cracks," *International Journal of Fracture*, vol. 114, no. 1, pp. 1–27, 2002.
- [28] X. J. Yang, *Advanced Local Fractional Calculus and Its Applications*, World Science Publisher, New York, NY, USA, 2012.
- [29] R. Chandra and K. Guruprasad, "Numerical estimation of stress intensity factors in patched cracked plates," *Engineering Fracture Mechanics*, vol. 27, no. 5, pp. 559–569, 1987.
- [30] S. M. Fekih, A. Albedah, F. Benyahia, M. Belhouari, B. Bachir Bouadjra, and A. Miloudi, "Optimisation of the sizes of bonded composite repair in aircraft structures," *Materials and Design*, vol. 41, pp. 171–176, 2012.
- [31] H. Hosseini-Toudeshky, B. Mohammadi, and H. R. Daghyani, "Mixed-mode fracture analysis of aluminium repaired panels using composite patches," *Composites Science and Technology*, vol. 66, no. 2, pp. 188–198, 2006.
- [32] M. R. Ayatollahi and R. Hashemi, "Computation of stress intensity factors (KI, KII) and T-stress for cracks reinforced by composite patching," *Composite Structures*, vol. 78, no. 4, pp. 602–609, 2007.

Research Article

Modeling and Flocking Consensus Analysis for Large-Scale UAV Swarms

Li Bing, Li Jie, and Huang KeWei

School of Mechatronical Engineering, Beijing Institute of Technology, Beijing 100081, China

Correspondence should be addressed to Huang KeWei; binlin_lee@163.com

Received 31 July 2013; Accepted 7 October 2013

Academic Editor: J. A. Tenreiro Machado

Copyright © 2013 Li Bing et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Recently, distributed coordination control of the unmanned aerial vehicle (UAV) swarms has been a particularly active topic in intelligent system field. In this paper, through understanding the emergent mechanism of the complex system, further research on the flocking and the dynamic characteristic of UAV swarms will be given. Firstly, this paper analyzes the current researches and existent problems of UAV swarms. Afterwards, by the theory of stochastic process and supplemented variables, a differential-integral model is established, converting the system model into Volterra integral equation. The existence and uniqueness of the solution of the system are discussed. Then the flocking control law is given based on artificial potential with system consensus. At last, we analyze the stability of the proposed flocking control algorithm based on the Lyapunov approach and prove that the system in a limited time can converge to the consensus direction of the velocity. Simulation results are provided to verify the conclusion.

1. Introduction

UAV is an advanced system with high autonomy for intelligent combat [1]. In the future, UAVs will be used for complex tasks, such as surveillance, reconnaissance, and precision strike missions. Many organizations have foreseen that in the near future, swarms of UAVs will replace single ones for more complicated missions in more uncertain and possibly hostile environments [2]. Therefore, many researchers are studying groups of cooperative UAVs.

(A) *Related Work on the UAV Swarms Problem.* The new challenges imposed by UAV swarms have attracted many researchers. New control mechanisms, application domains, simulation models, and simulation tools have been developed to tackle issues in different aspects of the swarm. Currently, a completely new topic is opening up in the area of UAV swarms performing different missions cooperatively. [3], the method of evolutionary pinning control is applied to UAV swarms successfully. Path planning and routing are investigated in [4, 5], using multiobjective evolutionary algorithm. The path planning problem in three-dimensional environment without any obstacles is addressed in [6, 7] and with only static obstacles in [8]. [9], cooperative searching problem

is discussed for the purpose of detecting moving and evading targets in a hazardous environment. A similar cooperative searching problem is also discussed in [10, 11]. Reference [12] investigates the automatic target recognition (ATR) problem in UAV control and proposes a distributed strategy for UAV swarms. Task allocation problem is discussed in [13–18] using different methodologies. Some applications of using a UAV swarm to search and destroy targets could be found in [19–22]. [23], a swarm simulator for target searching is implemented with Java. Garcia introduces a multi-UAVs simulator implemented with X-Plane—a commercial flight simulator [24]. Russell et al. present a parallel swarm simulation environment which utilizes an existing parallel emulation and simulation tool called SPEEDS [25]. MASON [26] is a general purpose multiagents simulation library utilized in our previous work, along with MATLAB based UAV simulator [27].

(B) *Related Work on the Consensus Flocking Problem.* Most research on formation of agent swarms uses distributed techniques by Reynolds' seminal work on the mobility of flocks [28], which prescribes three fundamental operations for each robot to realize distributed flocking—separation, alignment, and cohesion [29, 30]. One of the earliest attempts to realize

flocking through a set of basic behaviors including safe wandering, aggregation, dispersion, and homing to implement flocking is by Mataric [31]. Kelley and Keating realize flocking with robots based on leader-following behavior [32]. Hayes and Dormiani-Tabatabaei [33] propose a flocking algorithm based on two behaviors: collision avoidance and velocity-matching flock centering. Holland et al. [34] propose a flocking algorithm for UAV similar to Reynolds'. A host is used as an intermediate station for receiving each UAV's range, bearing, and velocity and sending them to other UAVs to simulate the sensing process of one UAV for perceiving range, bearing, and heading of its neighbors. Ferrante et al. [35] introduce a new communication strategy called the information aware communication for alignment behavior. Recently, Stranieri et al. [36] perform flocking with a swarm of behaviorally heterogeneous mobile robots.

In this paper, we consider models for flocking swarms. Firstly, a mathematical model of cooperative system is established by using Markov stochastic process and calculus analysis. Then, the control law for UAV swarm is established based on artificial potential field. At last, we analyze the stability of the proposed flocking control algorithm based on the Lyapunov approach and prove the conclusion that the system in a limited time can converge to the consensus direction of the velocity. Simulation results are provided to verify the conclusion.

2. The Model of the UAVs Swarms

2.1. Differential Integral Model. Let $C(t)$ denote the state of the UAV swarms at time t ; $C(t) = 0$ identifies the state that UAV swarms are stable at time t . The state of UAV i at time t is denoted by $c_i(t) = (p_i(t), o_i(t))$, in which the first element $p_i(t) = (x_i(t), y_i(t), z_i(t))$ is the UAV's position in the environment at time t , and the second element $o_i(t)$ is the UAV's orientation. The UAV's dynamics is subject to its physical curvature radius constraints, leading to the fact that it can only change its orientation by at most one step, which is described as go straight, go up, go down, turn left, turn upper left, turn lower left, turn right, turn upper right, and turn lower right.

In order to obtain Markov random process, the new state of process is derived by supplement of variable [37, 38], which is described as follows:

$$\begin{aligned} P_i(t) &= P(C(t) = i), \quad i = 0, 1, 2, \dots, N, \\ P_{jk}(x, t) &= P(C(t) = j, x < y_{j,k}(t) < x + dx), \quad (1) \\ j &= N + 1, \dots, M, \quad k = 0, 1, 2, \dots, N, \end{aligned}$$

where $y_i(t)$ is the dwell time after state i . So it is easy to verify that $\{U(t), 0 \leq t \leq T_0\} = (C(t), y(t))$ is a broad Markov random process.

The probability of state transition after Δt can be obtained using total probability theorem:

$$\begin{aligned} P_0(t + \Delta t) &= \left(\sum_{i=0}^M P(C(t) = i, C(t + \Delta t) = i) \right) \\ &= P_0(t) \left(1 - \sum_{i=1}^N \lambda_{0i} \Delta t + o(\Delta t) \right) + \sum_{i=1}^N \lambda_{i0} P_k(t) \Delta t \\ &\quad + \sum_{j=N+1}^M \int_0^T P_j(x, t) r_j(x) \Delta t dx + o(\Delta t). \end{aligned} \quad (2)$$

According to (2) we can get the all probability:

$$\begin{aligned} P_k(t + \Delta t) &= P_k(t) \left(1 - \sum_{i=1, i \neq k}^N \lambda_{ki} \Delta t + o(\Delta t) \right) \\ &\quad + \sum_{i=1, i \neq k}^N \lambda_{ik} P_k(t) \Delta t + o(\Delta t), \\ k &= 1, 2, \dots, N, \\ P_j(x, t) &= P(C(t) = j; \\ &\quad \text{the time of self-organised } C_{\text{system}} \text{ is } x, \\ &\quad C(t + \Delta t) = j) \\ &= P_j(x, t) \left(1 - r_j(x) \Delta t + o(\Delta t) \right), \\ j &= N + 1, \dots, M, \end{aligned} \quad (3)$$

where λ_{ij} is the average sustained rate of each state and $r_j(x)$ is the average repair rate at state j . Similarly, the expression of state transition rate for $P_j(x + \Delta x, t + \Delta t)$ can be derivated.

Differentiate the expression for state transition probability to derive its limit. Then the mathematical model can be described using integral-differential equations as follows:

$$\begin{aligned} \frac{dP_0(t)}{dt} &+ \sum_{i=0, i \neq k}^N \lambda_{0k} P_0(t) \\ &= \sum_{i=0, i \neq k}^N \lambda_{i0} P_i(t) \\ &\quad + \sum_{j=N+1}^M \int_0^T r_j(x) p_j(x, t) dx \\ \frac{dP_k(t)}{dt} &+ \sum_{i=0, i \neq k}^N \lambda_{ki} P_k(t) \\ &= \sum_{i=0, i \neq k}^N \lambda_{ik} P_i(t), \\ k &= 1, 2, \dots, N, \end{aligned}$$

$$\frac{\partial p_j(x, t)}{\partial x} + \frac{\partial p_j(x, t)}{\partial t} + r_j(x) p_j(x, t) = 0,$$

$$j = N + 1, \dots, M.$$

(4)

$$\begin{aligned} & \times \exp(-a_0(t-s)) ds \\ &= \exp(-a_0 t) + \sum_{i=1}^N \lambda_{i0} \int_0^{T_0} P_i(s) \exp(-a_0(t-s)) \times ds \\ &+ \sum_{i=N+1}^M \int_0^{T_0} P_i(0, \tau) d\tau \\ &\times \int_0^{T_0 - \tau} \exp\left(-a_0(T_0 - \tau) + a_0 v - \int_0^v r_i(\mu) d\mu\right) r_i(v) dv \\ &= \exp(-a_0 t) + \sum_{i=1}^N \lambda_{i0} \int_0^{T_0} P_i(s) k_0 ds \\ &+ \sum_{i=N+1}^M \int_0^{T_0} P_i(0, \tau) K_i(t - \tau) d\tau, \end{aligned} \quad (6)$$

The boundary and initial conditions are

$$P_0(0) = 1, \quad P_1(0) = 0, \dots, P_i(0) = 0, \dots, P_N(0) = 0,$$

$$P_j(0, t) = \sum_{i=0}^N \lambda_{ij} P_i(t), \quad j = N + 1, \dots, M. \quad (5)$$

Theorem 1. *The reliability of coordination system has uniqueness and nonnegative solution on $C[0, T]$.*

Proof. According to the initial conditions we can get the analytic solution of the partial differential equation [39, 40].

Set

$$\begin{aligned} a_0 &= \sum_{i=1}^N \lambda_{0i}, \dots, a_j = \sum_{i=1}^N \lambda_{ji}, \dots, a_N = \sum_{i=1}^N \lambda_{Ni}, \\ P_j(x, t) &= P_i(0, t-x) \exp\left(-\int_0^x r_i(\mu) d\mu\right), \\ i &= N + 1, \dots, M \\ \frac{dP_0(t)}{dt} &= -a_0 P_0(t) + \sum_{i=0, i \neq k}^N \lambda_{i0} P_i(t) \\ &+ \sum_{j=N+1}^M \int_0^T P_j(0, t-x) \exp\left(-\int_0^x r_i(\mu) d\mu\right) \times r_i(x) dx, \end{aligned}$$

$$\begin{aligned} P_0(t) &= \exp(-a_0 t) + \sum_{i=1}^N \lambda_{i0} \int_0^{T_0} P_i(s) \exp(-a_0(t-s)) ds \\ &+ \int_0^{T_0} \left[\sum_{i=N+1}^M \int_0^s P_i(0, \tau) \exp\left(-\int_0^{s-\tau} r_i(\mu) d\mu\right) \times r_i(s-\tau) d\tau \right] \end{aligned}$$

where

$$\begin{aligned} k_0 &= \exp(-a_0(T-s)) \\ k_i(T-\tau) &= \int_0^{T-\tau} k_0 * \exp\left(a_0 v - \int_0^v r_i(\mu) d\mu\right) r_i(v) dv, \\ i &= N + 1, \dots, M, \\ P_j(t) &= \sum_{i=0, i \neq j}^N \int_0^T \exp(-a_j(T-s)) \lambda_{ij} P_i(s) ds \\ &= \sum_{i=0, i \neq j}^N k_j \lambda_{ij} P_i(s) ds, \quad j = 1, 2, \dots, N, \end{aligned} \quad (7)$$

where $k_j = \exp(-a_j(T-s))$, $j = 1, 2, \dots, N$.

So we can get the following equation:

$$\begin{aligned} P_j(0, t) &= \sum_{i=0}^N \lambda_{ij} \sum_{l=0, l \neq j}^N \int_0^T k_j \lambda_{lj} P_l(s) ds \\ &= \int_0^T k_j \left(\sum_{i=0}^N \lambda_{ij} \right) \left(\sum_{l=0, l \neq j}^N \lambda_{lj} P_l(s) \right) ds, \\ j &= N + 1, \dots, M. \end{aligned} \quad (8)$$

Assuming

$$\begin{aligned} P(t) &= (P_0(t), P_1(t), \dots, P_N(t), P_{N+1}(0, t), \dots, P_M(0, t)), \\ f(t) &= (f_0(t), f_1(t), \dots, f_N(t), f_{N+1}(0, t), \dots, f_M(0, t)) \\ &= (\exp(-a_0 t), 0, \dots, 0, \dots, 0), \end{aligned} \quad (9)$$

then, the solution of the system can be converted into vectors format as follows:

$$P(t) = f(t) + \int_0^T k(T-s) P(s) ds,$$

$$k(T-\tau) = \begin{pmatrix} 0 & \lambda_{10}k_0 & \cdots & \lambda_{N0}k_0 & k_{N+1}(T-s) & \cdots & k_M(T-s) \\ \lambda_{01}k_1 & 0 & \cdots & \lambda_{N1}k_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \lambda_{0N}k_N & \lambda_{2N}k_N & \cdots & 0 & 0 & \cdots & 0 \\ \left(\sum_{l=0, l \neq j}^N \lambda_{lj} \right) \lambda_{0N+1} & \left(\sum_{l=0, l \neq j}^N \lambda_{lj} \right) \lambda_{1N+1} & \cdots & \left(\sum_{l=0, l \neq j}^N \lambda_{lj} \right) \lambda_{NN+1} & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \left(\sum_{l=0, l \neq j}^N \lambda_{lj} \right) \lambda_{0M} & \left(\sum_{l=0, l \neq j}^N \lambda_{lj} \right) \lambda_{1M} & \cdots & \left(\sum_{l=0, l \neq j}^N \lambda_{lj} \right) \lambda_{NM} & 0 & \cdots & 0 \end{pmatrix}. \quad (10)$$

□

Any component of $F(t)$ and $G(t-h)$ vector is nonnegative. The functions $F_j(t)$ and $G_j(t-h)$, $j = 1, 2, \dots, M$ are limitary on the domain $0 < T < +\infty$. The solution of integral equation is unique and nonnegative on $C[0, T]$. So the reliability of coordination system has unique and nonnegative solution on $C[0, T]$.

2.2. Probabilistic Analysis Based on State Transformation. The behavior evolution of the UAV swarm system is a limited Markov decision process. Suppose that the probability distribution of the system state is $P(N, t)$ at time t . Then at time $t+\tau$, the probability distribution is $P(N, t+\tau)$. According to the relationship of the probability density at different time, the marginal probability density $P(N, t+\tau)$ is $P(N, t+\tau) = \int P(\eta, t)P(N, t+\tau | \eta, t)d\eta$.

And the time derivative of the $P(N, t+\tau)$ is

$$\frac{dP(N, t)}{dt} = \lim_{\tau \rightarrow 0} \frac{P(N, t+\tau) - P(N, t)}{\tau}. \quad (11)$$

Define $W_t(\eta, N)$ as the transition probability density from state η to state N in unit time during time interval $[t, t+\tau]$. So the transition probability from state η to state N during time interval $[t, t+\tau]$ is $\tau W_t(\eta, N)$. Then the probability by which the transition does not happen is

$$1 - \tau \int W_t(\eta, N) \delta(N - \eta) d\eta, \quad (12)$$

where $\delta(N - \eta) = P(N, t+\tau | \eta, t)$, when $\tau = 0$. Thus,

$$\frac{dP(N, t)}{dt} = \int \{W(\eta, n)P(\eta, t) - W(n, \eta)P(n, t)\} d\eta. \quad (13)$$

Equation (13) describes the evolution of the system states over time, which is the primary equation model of the UAV swarms behavior.

3. Flocking Control of UAV Swarms

3.1. Flocking Control Law. In this section, first we design a distributed flocking control law. Assuming that each UAV senses its own position and velocity and is able to obtain its neighbors' position and velocity, the UAV swarms form flocking behaviour model structure control law as follows:

$$U_i = \text{Uniformity}(\nu) \sum_{j=1, j \neq i}^N f(p_{i,j})(p_i - p_j) + \beta(p_i - p_{\text{goal}}) + \sum_{j=1, j \neq i}^N a_{i,j}(p_{i,j})(v_j - v_i), \quad (14)$$

where $\text{Uniformity}(\nu) = (\alpha \sum_{i \neq j} \|v_i - v_j\|_2)$. Note that alignment at a common velocity is equivalent to $\text{Uniformity}(\nu) = 0$. $p_{i,j}$ is the distance between the individual i and j . U_{ij} is potential function and satisfies the condition [29, 30]:

- (i) $U_{iR}(\|p_i - p_j\|) \rightarrow \infty$, $\|p_i - p_j\| < R_{\min}$,
- (ii) $R_{\min} \leq \|p_i - p_j\| \leq R_{\max}$, $\exists! \min U_{iR}$.

3.2. Stability Analysis. Consider the following positive semidefinite function:

$$E = \frac{1}{2} \left(\sum_{i=1}^N \left(\sum_{j=1, j \neq i}^N U_{i,j} + K_i U_{i,\text{goal}} \right) + v_i^T v_i \right). \quad (15)$$

In order to facilitate writing, we simplify the certification process variable substitution as follows:

$$\begin{aligned}
\bar{U}_i &= \left(\sum_{j=1, j \neq i}^N U_{i,j} + K_i U_{i,\text{goal}} \right), \\
\dot{E} &= \frac{1}{2} \sum_{i=1}^N \dot{\bar{U}}_i - \sum_{i=1}^N v_i^T U_i \\
&= \frac{1}{2} \sum_{i=1}^N \dot{\bar{U}}_i + \sum_{i=1}^N v_i^T \left(- \sum_{j=1, j \neq i}^N \nabla_{p_i} U_{ij} (p_{i,j}) - K_i \nabla_{p_i} U_{iA} (p_{i,j}) \right. \\
&\quad \left. - \sum_{j=1, j \neq i}^N \alpha_{i,j} (v_i - v_j) \right) \\
&= \sum_{i=1}^N v_i^T \nabla_{p_i} \bar{U}_i + \sum_{i=1}^N v_i^T \left(- \sum_{j=1, j \neq i}^N \nabla_{p_i} U_{ij} (p_{i,j}) \right. \\
&\quad \left. - K_i \nabla_{p_i} U_{i,\text{goal}} (p_{i,\text{goal}}) \right. \\
&\quad \left. - \sum_{j=1, j \neq i}^N \alpha_{i,j} (v_i - v_j) \right) \\
&= \sum_{i=1}^N v_i^T \nabla_{p_i} \left(\sum_{j=1, j \neq i}^N U_{i,j} + K_i U_{i,\text{goal}} \right) \\
&\quad + \sum_{i=1}^N v_i^T \left(- \sum_{j=1, j \neq i}^N \nabla_{p_i} U_{ij} - K_i \nabla_{p_i} U_{i,\text{goal}} \right. \\
&\quad \left. - \sum_{j=1, j \neq i}^N \alpha_{i,j} (v_i - v_j) \right) \\
&= \sum_{i=1}^N v_i^T \left(- \sum_{j=1, j \neq i}^N \alpha_{i,j} (v_i - v_j) \right) \\
&= -v^T (L_C \otimes I_3) v,
\end{aligned} \tag{16}$$

where L_c is UAV swarms system satisfying the Laplacian matrix of the communication conditions. Therefore, the quadratic form is explicitly described as follows:

$$\dot{E} = -v_x^T L_c v_x - v_y^T L_c v_y - v_z^T L_c v_z \leq 0. \tag{17}$$

Consider the following collections: $\{v_i, P_{i,j} \mid E \leq C\}$ is a closed set. The following is to verify that it is a compact set, and there is a clear conclusion that $P_{i,j} \leq C$. Similarly $v_i^T v_i \leq C$, $\|v_i\| \leq C^{1/2}$, and according to the definition of the potential field we obtain $\|P_{i,j}\| \leq U_{i,j}^{-1}(C(N-1))$. According to the LaSalle invariance principle, the system will converge to the largest invariant set in the area and meet $\dot{E} = 0$. According to $\dot{E} = 0$, when the system enters the steady state, the speed of each individual is equal, and all individuals move to the target position P_{goal} , making the overall potential energy minimum.

Theorem 2. Consider the UAV swarms consisting of N UAVs. The position of individual i is p_i . All individuals in the swarms will eventually build up to the spherical region:

$$\|p_i - p_c\| \leq \frac{2\lambda_{\max}(G)(N-1)lC}{\lambda_{\min}(A^T G + G A)}. \tag{18}$$

Proof. Consider

$$\begin{aligned}
\dot{P}_c &= V_c \\
\dot{V}_c &= \frac{1}{N} \sum_{i=1}^N \dot{v}_i = \frac{1}{N} \sum_{i=1}^N \left(- \sum_{j=1, j \neq i}^N \nabla_{p_i} U_{ij} (\|p_i - p_j\|) \right. \\
&\quad \left. - \beta_i \nabla_{p_i} U_{iA} (\|p_i - p_c\|) - \sum_{j=1, j \neq i}^N \alpha_{i,j} (v_i - v_j) \right) \tag{19}
\end{aligned}$$

where $P_c(t) = (1)/N \sum_{i=1}^N p_i(t)$, $V_c = (1)/N \sum_{i=1}^N \alpha_{i,j} v_i$.

By making the variable replacement $\varepsilon_{ip} = P_i - P_c$, $\varepsilon_{iv} = v_i - v_c$, we get

$$\begin{aligned}
\dot{\varepsilon}_{ip} &= \left(- \sum_{j=1, j \neq i}^N \nabla_{p_i} U_{ij} (\|p_i - p_j\|) \right. \\
&\quad \left. - \beta_i \nabla_{p_i} U_{iA} (\|p_i - p_c\|) \sum_{j=1, j \neq i}^N \alpha_{i,j} (v_i - v_j) \right) \tag{20} \\
&= -\beta_i \varepsilon_{ip} - N \left(\sum_{j=1, j \neq i}^N \alpha_{i,j} \right) \varepsilon_{iv} \\
&\quad - \sum_{j=1, j \neq i}^N \nabla_{p_i} U_{ij} (\|p_i - p_j\|).
\end{aligned}$$

Then

$$\begin{aligned}
\dot{\zeta}_i &= \begin{pmatrix} \dot{\varepsilon}_{ip} \\ \dot{\varepsilon}_{iv} \end{pmatrix} = \begin{pmatrix} 0 \\ -\beta_i - N \left(\sum_{j=1, j \neq i}^N \alpha_{i,j} \right) \end{pmatrix} \zeta_i \\
&\quad - \begin{pmatrix} 0 \\ \sum_{j=1, j \neq i}^N \nabla_{p_i} U_{ij} (\|p_i - p_j\|) \end{pmatrix}, \tag{21} \\
A &= \begin{pmatrix} 0 & 1 \\ -\beta_i & -N \left(\sum_{j=1, j \neq i}^N \alpha_{i,j} \right) \end{pmatrix}, \\
B &= \begin{pmatrix} 0 \\ \sum_{j=1, j \neq i}^N \nabla_{p_i} U_{ij} (\|p_i - p_j\|) \end{pmatrix}.
\end{aligned}$$

Since $K_i, N > 0$,

$$\begin{aligned} \lambda(A) &= \frac{-N(\sum_{j=1, j \neq i}^N \alpha_{i,j})}{2} \\ &\pm \frac{\sqrt{(N(\sum_{j=1, j \neq i}^N \alpha_{i,j}))^2 - 4k_i}}{2}, \end{aligned} \quad (22)$$

$\operatorname{Re}(\lambda(A)) < 0$.

The largest and the smallest eigenvalues of symmetric positive definite matrix S are $\lambda_{\max}(S)$ and $\lambda_{\min}(S)$, respectively. The symmetric positive definite matrix S with appropriate dimensions satisfies the following conclusion [41, 42].

$$\lambda_{\min}(S) \|\zeta_i\|^2 \leq \zeta_i^T S \zeta_i \leq \lambda_{\max}(S) \|\zeta_i\|^2. \quad (23)$$

Finally, select Lyapunov function

$$E_i = \zeta_i^T G \zeta_i, \quad G^T = G, \quad \lambda(G) > 0. \quad (24)$$

Time derivative can be obtained:

$$\begin{aligned} \dot{E}_i &= \dot{\zeta}_i^T G \zeta_i + \zeta_i^T G \dot{\zeta}_i = (A \zeta_i - B)^T G \zeta_i + \zeta_i^T G (A \zeta_i - B) \\ &= -\zeta_i^T (A^T G + G A) \zeta_i - 2\zeta_i^T G B, \\ U_{ij} (\|p_i - p_j\|) \|p_i - p_j\| &\leq l C. \end{aligned} \quad (25)$$

Therefore, according to the above formula we obtain

$$\dot{E}_i \leq -\lambda_{\min}(A^T G + G A) \|\zeta_i\|^2 + 2\lambda_{\max}(G) \|\zeta_i\| (N-1) l C. \quad (26)$$

When

$$\|\zeta_i\| > \frac{2\lambda_{\max}(G)(N-1)lC}{\lambda_{\min}(A^T G + G A)}, \quad (27)$$

$\dot{E}_i < 0$. The system continues to move closer to the population centre. Therefore, eventually the system stabilizes at a known system of

$$\|\zeta_i\| \leq \frac{2\lambda_{\max}(G)(N-1)lC}{\lambda_{\min}(A^T G + G A)}. \quad (28)$$

□

4. Simulation of System Flocking Formation Behavior

According to the UAV's physical characteristics, this paper will discretize the time with high frequency. Thus, a UAV i makes its path decision $P_i(t+1)$ at time-step t and will execute an action as the following equation:

$$P_i(t+1) = P_i(t) + v \Delta t. \quad (29)$$

The movement of the individual is not only controlled by itself but also affected by the state of other individuals.

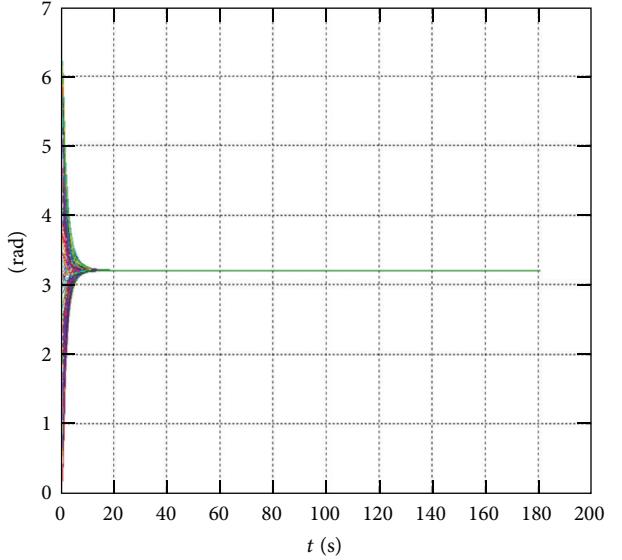


FIGURE 1: Velocities with respect to time.

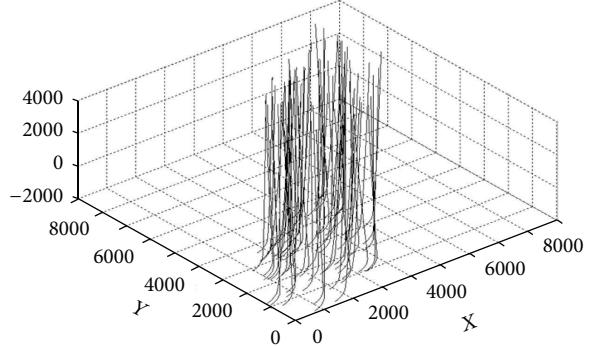


FIGURE 2: Trajectories with respect to time.

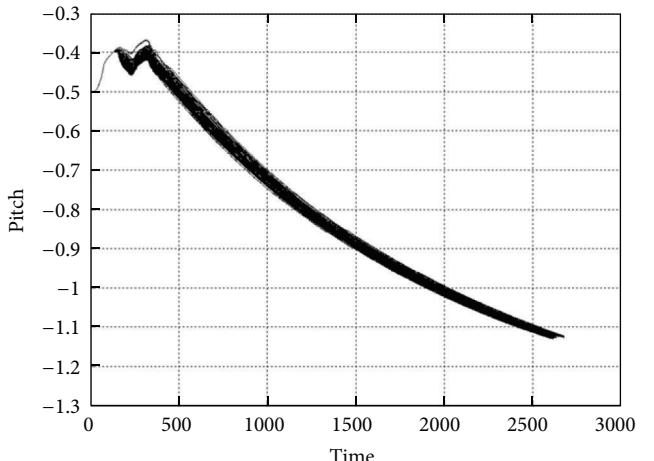


FIGURE 3: Angle of the Pitch with respect to time.

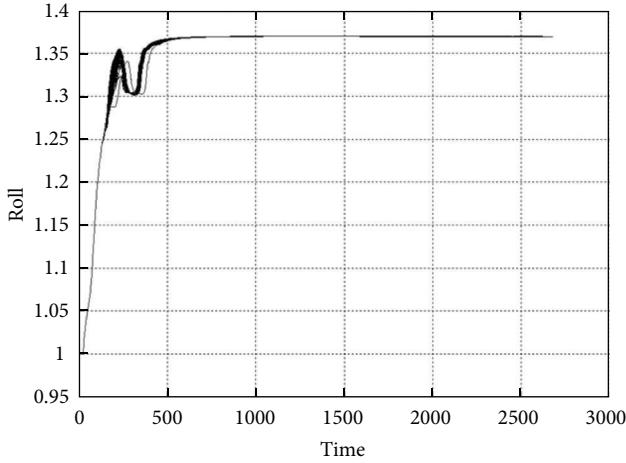


FIGURE 4: Angle of the Roll with respect to time.

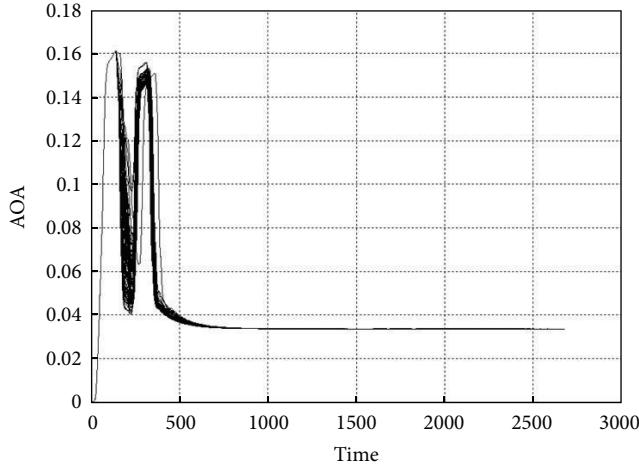


FIGURE 5: Angle of the Attack with respect to time.

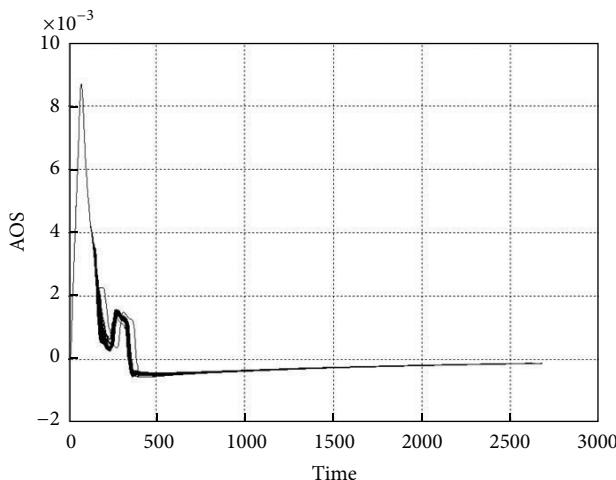


FIGURE 6: Angle of the Sideslip with respect to time.

Therefore, the individual direction of movement at a certain time is not only relative to its direction one moment before, but also relative to the directions of its surrounding individuals' movements. The influence of all the individuals to the individual i can be described as the following equation:

$$\frac{1}{\sum 1/p_{ij}} \sum_{j=1}^n \frac{\theta_j(t)}{p_{ij}}. \quad (30)$$

Then, the speed direction of the UAV i at time $(t + 1)$ can be modified as the following equation:

$$\theta_i(t+1) = \alpha \theta_i(t) + \beta \frac{1}{\sum 1/p_{ij}} \sum_{j=1}^n \frac{\theta_j(t)}{p_{ij}} + \gamma \arctan \frac{y_0 - y_1}{x_0 - x_1}. \quad (31)$$

We consider the swarms of 100 UAVs with six degrees of freedom. The weights of the cost function are set to $\alpha = 0.3$, $\beta = 0.5$, $\gamma = 0.2$, $V = [200, 0, 0]$ m/s and $m = 25$ kg. Direction is the rand variable from $-2 * \pi$ to $2 * \pi$. The position of the UAVs is the rand variable. The factors of the influence on the flight are wind and airstream. The results for the case of 100 UAVs are shown in Figures 1 and 2.

From Figure 1, at $t = 19$ s, the velocities of the swarms achieve consensus at $\theta = 3.2$.

Figure 2 describes the trajectories with respect to time. The UAV swarms system will eventually be able to form a stable distance between each individual and the same velocity vectors. Collision between individuals is thus avoided.

Figures 3, 4, and 5 show the Pitch, Roll, and Attack with respect to time. From the simulation results, we can conclude that the UAVs based on the method successfully fly after the adjustment at the initial stage.

Figure 6 shows the Sideslip with respect to time. Through the analysis of the Sideslip Angle, we can find that the Angle of the Sideslip is less than 0.5 degrees and tends to zero, to ensure the turning flight control.

5. Conclusion

This paper analyzed current researches and existent problems of UAV swarms. Afterwards, by the theory of stochastic process and supplemented variables, a differential-integral model was established. The existence and uniqueness of the solution of the system were discussed. The flocking control law is given based on artificial potential with system consensus. At last, we analyzed the stability of the proposed flocking control algorithm based on the Lyapunov approach and proved the conclusion that the system in 28 s can converge to the consensus direction of the velocity. And we performed simulation tests to verify the conclusion.

Acknowledgment

This paper is supported by The National Defense Pre-Research Foundation of China (Grant no. B22201XXXX).

References

- [1] H. B. Duan, Q. N. Luo, and G. J. Ma, "Hybrid particle swarm optimization and genetic algorithm for multi-UAV formation reconfiguration," *IEEE Computational Intelligence Magazine*, vol. 8, pp. 16–27, 2013.
- [2] W. Yi, M. B. Blake, and R. G. Madey, "An operation-time simulation framework for UAV swarm configuration and mission planning," *Procedia Computer Science*, vol. 18, pp. 1949–1958, 2013.
- [3] Y. Tang, H. J. Gao, J. Kurths, and J.-A. Fang, "Evolutionary pinning control and its application in UAV coordination," *IEEE Transactions on Industrial Informatics*, vol. 8, no. 4, pp. 828–838, 2012.
- [4] G. B. Lamont, J. N. Slear, and K. Melendez, "UAV swarm mission planning and routing using multi-objective evolutionary algorithms," in *Proceedings of the 1st IEEE Symposium of Computational Intelligence in Multicriteria Decision Making (MCDM '07)*, pp. 10–20, April 2007.
- [5] E. Besada-Portas, L. de la Torre, J. M. de la Cruz, and B. de Andrés-Toro, "Evolutionary trajectory planner for multiple UAVs in realistic scenarios," *IEEE Transactions on Robotics*, vol. 26, no. 4, pp. 619–634, 2010.
- [6] S. Kanchanavally, R. Ordóñez, and C. J. Schumacher, "Path planning in three dimensional environment using feedback linearization," in *Proceedings of the American Control Conference (ACC '06)*, pp. 3545–3550, Minneapolis, Minn, USA, June 2006.
- [7] M. Shanmugavel, A. Tsourdos, R. Zbikowski, and B. A. White, "3D path planning for multiple UAVs using pythagorean hodograph curves," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference and Exhibit*, pp. 1576–1589, Hilton Head, SC, USA, August 2007.
- [8] I. Hasircioglu, H. R. Topcuoglu, and M. Ermis, "3-D path planning for the navigation of unmanned aerial vehicles by using evolutionary algorithms," in *Proceedings of the 10th Annual Genetic and Evolutionary Computation Conference (GECCO '08)*, pp. 1499–1506, July 2008.
- [9] P. Vincent and I. Rubin, "A framework and analysis for cooperative search using UAV swarms," in *Proceedings of the ACM Symposium on Applied Computing*, pp. 79–86, 2004.
- [10] G. Varela, P. Caamamño, F. Orjales, Á. Deibe, F. López-Peña, and R. J. Duro, "Swarm intelligence based approach for real time UAV team coordination in search operations," in *Proceedings of the 3rd World Congress on Nature and Biologically Inspired Computing (NaBIC '11)*, pp. 365–370, October 2011.
- [11] Y. I. Yang, *Cooperative search by uninhabited air vehicles in dynamic environment [Ph.D. thesis]*, University of Cincinnati, Cincinnati, Ohio, USA, 2005.
- [12] P. Dasgupta, "A multiagent swarming system for distributed automatic target recognition using unmanned aerial vehicles," *IEEE Transactions on Systems, Man, and Cybernetics A*, vol. 38, no. 3, pp. 549–563, 2008.
- [13] M. Yunhong, J. Zhe, and Z. Deyun, "A faster pruning optimization algorithm for task assignment," *Journal of Northwestern Polytechnical University*, vol. 31, pp. 40–43, 2013.
- [14] B. Di, R. Zhou, and Q.-X. Ding, "Distributed coordinated heterogeneous task allocation for unmanned aerial vehicles," *Control and Decision*, vol. 28, pp. 274–278, 2013.
- [15] W. You, Sh. Wang, and J. Tao, "Multi-UAV dynamic task assignment by ISODATA restrained clustering," *Electronics Optics & Control*, vol. 17, pp. 22–26, 2010.
- [16] D. Dionne and C. A. Rabbath, "Multi-UAV decentralized task allocation with intermittent communications: the DTC algorithm," in *Proceedings of the American Control Conference (ACC '07)*, pp. 5406–5411, July 2007.
- [17] P. Dasgupta and M. Hoeing, "Dynamic pricing algorithms for task allocation in multi-agent swarms," in *Massively Multi-Agent Technology*, N. Jamali, P. Scerri, and T. Sugawara, Eds., vol. 5043 of *Lecture Notes in Computer Science*, pp. 64–79, 2008.
- [18] P. Gaudiano, B. Shargel, and E. Bonabeau, *Swarm Intelligence: A New C2 Paradigm with an Application to Control Swarms of UAVs*, Icosystem, Cambridge, Mass, USA, 2003.
- [19] J. Finke, K. M. Passino, S. Ganapathy, and A. Sparks, "Modeling and analysis of cooperative control systems for uninhabited autonomous vehicles," in *Cooperative Control*, V. Kumar, N. Leonard, and A. S. Morse, Eds., vol. 309 of *Lecture Notes in Control and Information Science*, pp. 79–102, Springer, New York, NY, USA, 2005.
- [20] T. McLain, R. Beard, and J. Kelsey, "Experimental demonstration of multiple robot cooperative target intercept," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, AIAA-2002-4678, 2002.
- [21] A. Moitra, R. Szczecina, V. Didomizio, L. Hoebel, R. Mattheyses, and B. Yamrom, "A novel approach for the coordination of multi-vehicle teams," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, pp. 608–618, Monterey, Calif, USA, 2001.
- [22] P. Vincent and I. Rubin, "A framework and analysis for cooperative search using UAV swarms," in *Proceedings of the ACM Applied Computing*, pp. 79–86, Nicosia, Cyprus, 2004.
- [23] H. Hexmoor, B. McLaughlan, and M. Baker, "Swarm control in unmanned aerial vehicles," in *Proceedings of the International Conference on Artificial Intelligence (ICAI '05)*, pp. 911–917, June 2005.
- [24] R. Garcia and L. Barnes, "Multi-UAV simulator utilizing x-plane," *Journal of Intelligent and Robotic Systems*, vol. 57, no. 1–4, pp. 393–406, 2010.
- [25] M. A. Russell, G. B. Lamont, and K. Melendez, "On using SPEEDES as a platform for a parallel swarm simulation," in *Proceedings of the Winter Simulation Conference*, pp. 1129–1137, December 2005.
- [26] S. Luke, C. Cioffi-Revilla, L. Panait, K. Sullivan, and G. Balan, "MASON: a multiagent simulation environment," *Simulation*, vol. 81, no. 7, pp. 517–527, 2005.
- [27] S. J. Rasmussen, J. W. Mitchell, P. R. Chandler, C. J. Schumacher, and A. L. Smith, "Introduction to the Multi-UAV2 simulation and its application to cooperative control research," in *Proceedings of the American Control Conference (ACC '05)*, pp. 4490–4501, June 2005.
- [28] T. McLain, R. Beard, and J. Kelsey, "Experimental demonstration of multiple robot cooperative target intercept," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, AIAA-2002-4678, Monterey, Calif, USA, 2002.
- [29] H. G. Tanner, A. Jadbabaie, and G. J. Pappas, "Stable flocking of mobile agents, part I: fixed topology," in *Proceedings of the 42nd IEEE Conference on Decision and Control*, pp. 2010–2015, December 2003.
- [30] H. G. Tanner, A. Jadbabaie, and G. J. Pappas, "Stable flocking of mobile agents part II: dynamic topology," in *Proceedings of the 42nd IEEE Conference on Decision and Control*, pp. 2016–2021, December 2003.

- [31] M. J. Mataric, *Interaction and intelligent behavior [Ph.D. thesis]*, Massachusetts Institute of Technology, Cambridge, Mass, USA, 1994.
- [32] I. Kelly and D. Keating, “Flocking by the fusion of sonar and active infrared sensors on physical autonomous robots,” in *Proceedings of the Conference on Mechatronics and Machine Vision in Practice*, pp. 14–17, 1996.
- [33] A. T. Hayes and P. Dormiani-Tabatabaei, “Self-organized flocking with agent failure: off-line optimization and demonstration with real robots,” in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 3900–3905, May 2002.
- [34] O. Holland, J. Woods, R. de Nardi, and A. Clark, “Beyond swarm intelligence: the ultraswarm,” in *Proceedings of the IEEE Swarm Intelligence Symposium (SIS '05)*, pp. 217–224, June 2005.
- [35] E. Ferrante, A. E. Turgut, N. Mathews, M. Birattari, and M. Dorigo, “Flocking in stationary and non-stationary environments: a novel communication strategy for heading alignment,” in *Parallel Problem Solving from Nature—PPSN XI*, R. Schaefer, C. Cotta, J. Kolodziej, and G. Rudolph, Eds., vol. 6239 of *Lecture Notes in Computer Science*, pp. 331–340, 2010.
- [36] A. Stranieri, E. Ferrante, A. E. Turgut et al., “Self-organized flocking with a heterogeneous mobile robot swarm,” Tech. Rep., 2011.
- [37] B. Li, “Stochastic process model of the multi-UAVs collaborative system based on state transition,” in *Proceedings of Conference on Modeling, Identification and Control*, pp. 757–761, 2012.
- [38] Z. X. Chen, *Partial Differential Equations*, Science Press, Beijing, China, 2002.
- [39] L. Guo, H. Xu, C. Gao, and G. Zhu, “Stability analysis of a new kind series system,” *IMA Journal of Applied Mathematics*, vol. 75, no. 3, pp. 439–460, 2010.
- [40] D. p. Gaver, “Time to failure and availability of paralleled system with repair,” *IEEE Transactions on Reliability*, vol. 12, pp. 30–38, 1963.
- [41] Q. J. Fan, *Key techniques research of cooperative formation biomimetic flight control for multi-UAV [Ph.D. thesis]*, Nanjing University of Aeronautics and Astronautics, Nanjing, China, 2008.
- [42] C. Yancai, *Research on distributed cooperative control for swarm UAVs [Ph.D. thesis]*, Nanjing University of Aeronautics and Astronautics, Nanjing, China, 2011.

Research Article

A Microbial Fuel Cell Modified with Carbon Nanomaterials for Organic Removal and Denitrification

Njud S. Alharbi

Biological Sciences Department, King Abdulaziz University, P.O. Box 80204, Jeddah 21589, Saudi Arabia

Correspondence should be addressed to Njud S. Alharbi; njud.alharbi@yahoo.com

Received 26 July 2013; Accepted 23 September 2013

Academic Editor: Bashir Ahmad

Copyright © 2013 Njud S. Alharbi. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper investigated microbial denitrification using electrochemical sources to replace organic matter as reductant. The work also involved developing a system that could be optimised for nitrate removal in applied situations such as water processing in fish farming or drinking water, where high nitrate levels represent a potential health problem. Consequently, the study examined a range of developments for the removal of nitrate from water based on the development of electrochemical biotransformation systems for nitrate removal. This also offers considerable scope for the potential application of these systems in broader bionanotechnology based processes. Furthermore, the work discussed the context of improved microbial fuel cell (MFC) performance, potential analytic applications, and further innovations using a bionanotechnology approach to analyse cell-electrode interactions. High nitrate removal rate of more than 95% was successfully achieved by using a MFC system modified with carbon nanomaterials.

1. Introduction

A biological nitrate removal using microbial fuel cells (MFCs) has attracted great attention due to its ability to directly generate electricity, while accomplishing water and wastewater treatment. MFCs are bioelectrochemical systems (BESs) that use electrochemically active microorganisms. The microorganisms act as catalyst for the electrochemical oxidation of the organic material, and the electrode is therefore referred to as a microbial bioanode [1]. This oxidation occurs in an anaerobic environment, resulting in producing electrons, protons, and CO_2 . The protons, which are created at the anode to maintain a charge balance, typically migrate through the solution to the cathode. In contrast, the electrons flow through an external electrical circuit with a load resistance to the cathode and in turn combine with protons and an oxidant to generate electricity. In a conventional MFC, the cathode is abiotic, due to the use of expensive and sustainable catalysts. Microbial biocathodes have been shown as greatly promising alternative, as inexpensive and sustainable electrode materials can be used [2]. In addition, anaerobic biocathodes can offer the advantages of having a MFC system

with both anaerobic anode and cathode chambers. This helps minimise the risk of oxygen leaking in the anode chamber, thus increasing the efficiency of its reaction, and also helps reduce the cost of the catalyst used. The first development of a biocathode was achieved by Clauwaert et al. in [3], demonstrating that a complete denitrification can be performed using microorganisms in the cathode with electrons supplied by microorganisms oxidising acetate in the anode. Simultaneous organic removal, power production, and full denitrification were achieved without relying on H_2 -formation or external power. It was also reported that a bacterial culture enriched in *Geobacter* species could reduce nitrate to nitrite by using the cathode as the terminal electron donor [4]. In 2008, Lefebvre and coworkers proposed a novel type of two-chambered MFC, where the costly catalyst on the cathode surface was replaced by an autoheterotrophic denitrifying biofilm [5]. Denitrification was performed by microorganisms using electrons supplied by bacteria oxidising domestic wastewater and with acetate as the substrate in the anode chamber.

The MFC performance can be improved through a proper design of the MFC reactor and the selection of

appropriate materials. A wide variety of carbon-based materials, including carbon paper, carbon cloth, carbon felt, and graphite granules (GGs), have been examined as electrode in MFC [6–9]. Carbon based materials such as graphite felt (GF), carbon fibres, and carbon cloth have been widely used as electrode materials for electrochemical properties due to their stability, high surface area, and availability at a reasonable cost. GF electrodes have been shown to be attractive materials due to their interesting characteristics in applications, such as electrosynthesis and metal recovery. However, GF electrodes have low electrochemical activity due to poor kinetics and reversibility that limit their use as active electrode materials. Therefore, great attention has been paid to the modification of such electrode materials in order to improve their electrochemical properties. The modification of the electrode surface area may also improve such aspects as bacterial adhesion, hence increasing the electron transfer from bacteria to the electrode surface. *Geobacter* sp. can make an electrical connection with graphite electrodes and can accept electrons from an electrode when the electrode is poised at a negative potential [4]. It was reported that the electrochemical activity of the GF can be improved by treating the GF with concentrated sulphuric acid [10]. Alternative novel modification techniques were also developed to enhance the GF electrochemical activity material, as shown in [11]. Furthermore, it has been demonstrated that carbon nanomaterials have the ability to facilitate the electron transfer process during the electroreduction and electrooxidation of electroactive species such as NADH and hydrogen peroxide [12, 13] and during the enzyme-substrate interaction [14]. Additionally, the porous structure of the carbon nanomaterials can give the electrode better wetting properties [15]. This can allow the analyte to diffuse into the carbon nanomaterial bundles with lower friction, as mentioned in [16].

In this study, a mediatorless H-shaped MFC is constructed using bacteria both in the anode and cathode in order to perform biological nitrate removal. A mixed bacterial culture in the cathode can perform denitrification through the use of electrons supplied by a mixed bacterial culture oxidising acetate in the anode. This configuration represents an anodic (oxidative) reaction in water heavily contaminated with organics, while the reduction for nitrate is carried in relative clean environment, so avoiding the addition of organic matter required for conventional denitrification systems. Furthermore, in an attempt to improve the MFC, modifications to the electrochemical properties of the electrode were investigated through the use of a cyclic voltammetry using carbon nanomaterials to coat the graphite felts electrodes. Among all the nanomaterials used in this study, graphitised carbon nanofibers (GCNFs) were selected for further investigation, as they offered the best electrochemical performance and were thought to provide the largest active surface area. The performance of the MFC system coupled with the GCNFs modified electrodes was evaluated and significant improvements were observed. The highest voltage output achieved was about 41 mV with over 95% nitrate removal.

2. Materials and Methods

2.1. MFC Design and Operation

2.1.1. MFC Construction. An H-shape MFC, which consisted of two separated chambers joined with a glass tube containing a 4.7 cm × 4.5 cm diameter proton exchange membrane (PEM), was constructed. The volume of each chamber was approximately 200 mL with a 50 mL headspace. Anodic and cathodic chambers were operated in anaerobic conditions, where the top of each chamber was sealed with a rubber stopper. A platinum wire was introduced from the top of each chamber through the rubber stopper to solder one end of a rectangular prism shaped graphite felt electrode (GFE), having a surface area of 40 cm² and a weight of 0.719 g. Given that the rectangular prism-shaped electrode (SA) had a length (*L*) of 4 cm, a width (*W*) of 0.5 cm, and a height (*H*) of 4 cm, the outer surface area was calculated as

$$\begin{aligned} SA &= 2LW + (2L + 2W)H \\ SA &= 2 \times 4 \times 0.5 + (2 \times 4 + 2 \times 0.5) \times 4 = 40 \text{ cm}^2. \end{aligned} \quad (1)$$

Each chamber has two side ports to allow the provision of fresh substrate, as well as the purge of nitrogen, and to allow the removal of the treated one.

2.1.2. MFC Inoculation and Operation. A soil sample was collected from a depth of 3 cm using a clean spatula and bag and was then delivered to a laboratory within 5 min. An inoculum solution was prepared by adding 10 g soil to a 250 mL flask containing 70 mL of an autoclaved solution. The latter was prepared with 1 g CH₃-COONa-3H₂O and 3 g KNO₃ in 500 mL distilled water and then autoclaved at 121°C for 15 min. The inoculum solution was shaken and kept for 20 min before being used. The mixed culture of soil inoculum was simultaneously inoculated into each chamber during the MFC start-up. Both anode and cathode chambers were similarly filled with an artificial wastewater medium. This artificial wastewater medium, which was prepared according to [17], contained inorganic salts dissolved in 990 mL of 5 mM phosphate buffer and 10 mL of a trace mineral solution. The phosphate buffer solution used to form the basis of the salt solution was prepared as a mixture of KH₂PO₄ and K₂HPO₄. The salt solution was autoclaved at 121°C for 15 min and left to cool. Nitrogen gas was then purged for 30 min to remove oxygen. The solution's pH was checked and adjusted to 7 in each chamber through the addition of 1 M HCl or 1 M NaOH. The former (HCl) was used to reduce pH, while the latter (NaOH) was used to increase pH.

In order to maximise the growth of the biofilm-forming organisms, enrichment was conducted in three different modes for three months, following the procedure given in [18]. The MFC reactor was first operated in a fed-batch mode under a closed circuit condition through 10 KΩ in the first month. This essential growth mode allowed the microorganisms to grow at a constant load using acetate as a carbon source in the anodic chamber and nitrate as an electron acceptor in the cathodic chamber. This was facilitated through the weekly addition of 2 mM (169.5 mg/L) sodium

acetate $\text{CH}_3\text{-COONa-3H}_2\text{O}$ and 8.4 mM (847.9 mg/L) potassium nitrate (KNO_3) into the anodic and cathodic chambers, respectively. The second mode was performed as starvation mode for a month, where the addition of a carbon source was stopped and the nitrate fed was carried out weekly. The main goal of this mode was to enable consumption of carbon source added in the first mode. The mode was also given a selection of organisms that were capable of either using residual acetate or internally storing carbon to be subsequently used for cellular maintenance needs. This purpose was achieved by replacing the medium solution in both chambers after two weeks of the starvation mode period, which helped avoid sodium and nitrite accumulations in the anode and cathode chambers, respectively. Furthermore, the weekly addition of the carbon source was restarted and the external load was reduced from $10 \text{ K}\Omega$ to 500Ω to allow a higher current to flow between the electrodes for the rest of the enrichment period. This resulted in higher availability of an electron acceptor in the cathode and offered the opportunity of the growth of exoelectrogenic organisms. During the enrichment process, both chambers were purged with gaseous nitrogen for 10 min after each fuel addition in order to obtain anaerobic conditions. Moreover, the solution in each chamber was continuously mixed using a magnetic stirrer to enhance mass transfer. The experiment was conducted at room temperature.

The enrichment process was completed and the MFC operation was then started, where the power generated is computed as the production of cell voltage across an external resistance due to the current flow through the resistor. The MFC system is operated at a steady state when the power generated equals the power consumed for an extended time. In steady state MFC systems, sustainable power can be generated as the product of a steady current passing through a fixed load and a constant voltage drop across this load. Due to the possibility of many steady conditions in a MFC system, it is important to define the condition in which the MFC produces the maximum sustainable current, as well as computing the maximum sustainable power. In order to obtain a steady state condition, the MFC system was initially conducted through an external resistance of 500Ω in several batch modes using acetate as the carbon source and nitrate as the electron acceptor under anaerobic conditions. The cell voltage was measured every hour using a digital multimeter connected to a personal computer through a data acquisition system (34405A, Agilent). The nutrient medium was completely replaced at the operation start-up and when the voltage dropped to less than 5 mV as an end of batch. Previous study [19] found that a stable MFC system was achieved when the voltage output was reproducible after replacing the medium at least twice. In this study, a stable voltage generation (sustainable voltage) of 30 mV approximately was produced after three batches. Furthermore, in order to define the steady state that provides the maximum power output, a polarization curve was obtained by measuring the stable voltage generated at various external resistances. In addition, a series of batch-mode MFC tests were performed to investigate the effect of nitrate and acetate concentrations on the MFC's performance and its denitrification activities. The operation of these

tests was carried out under closed circuit with an external resistance of 500Ω . The effect of external resistance on the denitrification process was studied by operating the MFC under three different loads. Nitrate reduction and nitrite accumulation rates were studied and evaluated throughout the tests. Nitrate and nitrite concentrations were measured using an ultraviolet spectrophotometric screening method proposed in [20] and a development spectrophotometric method proposed in [21], respectively.

2.2. MFC Developments and Modifications

2.2.1. Investigation of the Electrochemical Properties of the Electrode. This section aims to investigate the electrochemical properties of a GFE using cyclic voltammetry experiments. The cyclic voltammetry is an active technique that involves the application of potential to an electrode and monitoring of the current response through the electrochemical cell over a period of time. The applied potential induces a change in the concentration of the electroactive element at the electrode surface through electrochemical oxidation or reduction. This technique can either be used for organic or inorganic substances including studies of adsorption processes on surfaces, electron transfer and reaction mechanisms, kinetics of electron transfer processes, and transport of species in solution. Furthermore, the electrochemical experiments were conducted using Epsilon electrochemistry and a three-electrode arrangement. The three electrodes were comprised of a working electrode at which the redox reaction takes place, a reference electrode through which no current flows, and a counter electrode (an auxiliary electrode) which completes the circuit. The working electrode was graphite felt, the auxiliary electrode was platinum (Pt), and the reference electrode was a silver/silver chloride electrode (Ag/AgCl). These electrodes are connected to a potentiostat that provides the desired potential. The potentiostat was linked up with a personal computer controlled electrochemical application that was used for collecting and calculating the data. The electrochemical experiments were carried out in a one compartment electrochemical cell with a volume of about 25 mL at room temperature and in an oxygen free environment (by bubbling nitrogen through the solution). The electrolyte used in the electrochemical experiments was 1 mM methyl viologen (MV^{+2}) in a 0.1 M phosphate buffer solution (PBS), which was composed of a mixture of KH_2PO_4 and K_2HPO_4 and with a pH of 7. The methyl viologen (MV^{+2}) has been considered as the simplest redox system in which MV^{+2} reduced to the relatively stable cation radical MV^+ [22, 23]. The MV^+ is relatively stable in oxygen free solutions, resulting in a deep blue colour, which is evidence of the formation of MV^+ in the solution, in contrast to the colourless MV^{+2} [24, 25]. The methyl viologen can act as a good electron transfer mediator for a biological system [23] and can also be electroreduced on a surface of various electrodes [26, 27]. Additionally, the effect of the pH value on electrochemical behaviour was investigated through the use of a range of pH values from 4 to 8, in the steps of 1. The interaction between the concentration and the peak current obtained

was also studied, where methyl viologen concentrations of 0.25 mM, 0.5 mM, 1 mM, 2.5 mM, and 5 mM were used. Moreover, the effect of the PBS concentration was studied, and PBS concentrations of 0.2 M, 0.1 M, 0.05 M, 0.01 M, 0.005 M, and 0.001 M were used. In addition, in order to study the correlation between the redox mechanism and the surface area used, the graphite (carbon) felt electrode was cut into multiple pieces with different surface areas. The electrochemical experiments were performed as follows.

- (1) A solution of 0.1 M phosphate buffer in 25 mL distilled water of pH 7 was prepared, and a 1 mM methyl viologen (MV^{+2}) concentration was then made in the 0.1 M phosphate buffer solution (i.e., using the phosphate buffer solution as diluent). The resultant solution was poured into an electrochemical cell with a volume of about 25 mL.
- (2) The three electrodes were carefully connected to an external cell box in the faraday cage and were then fitted into the electrochemical cell, by making sure that all electrodes were submerged but not touching the cell bottom.
- (3) The computer programme was set to the following conditions: an initial potential (IP) of -200 mV, a final potential (FP) of -200 mV, a switching potential (SP) of -900 mV, and a scan rate (SR) of 20, 50, 100, 200, or 250 mVs^{-1} at room temperature.
- (4) The solution in the cell was purged with nitrogen for 2 min, while stirring through the use of a small magnetic stirrer to achieve anoxic conditions.
- (5) The experiment was run and a voltammogram was taken.
- (6) Steps 1 to 5 were repeated with different pH values of 4, 5, 6, and 8.
- (7) Steps 1 to 5 were repeated with different methyl viologen concentrations of 0.25 mM, 0.5 mM, 2.5 mM, and 5 mM, while using a pH value of 7.
- (8) Steps 1 to 5 were repeated with different phosphate buffer concentrations of 0.2 M, 0.05 M, 0.01 M, 0.005 M, and 0.001 M, while using a 1 mM methyl viologen concentration and adjusting the pH value to 7 in all the phosphate buffer concentrations.
- (9) Steps 1 to 5 were repeated with different surface areas of the graphite (carbon) felt electrode. All the pieces of the graphite (carbon) felt electrodes were washed with distilled water before being used.

2.2.2. Electrode Modifications Based on Carbon Nanomaterials. Graphite felt has a large specific surface area and good stability [28]. However, it has a lower electrochemical activity that leads to a limited voltage efficiency and a lower power density, compared to the other materials used in the MFC system. This work aims to improve the electrochemical activity of the GF electrode through the use of carbon nanomaterials. The carbon electrode surface area can be enlarged by dispersing the carbon nanomaterials on the surface of the electrode

to form a randomly dispersed array of high surface area. Four types of carbon nanomaterials were used in this study to modify the GF electrode. These carbon nanomaterials include single-walled carbon nanotubes (SWCNTs), graphitised carbon black (GCB), carbon nanofibres (CNFs), and graphitised carbon nanofibres (GCNFs). The electrochemical performance of the modified GF electrodes was investigated and evaluated through cyclic voltammetry. The experiments were performed as follows.

- (1) The tests were performed in 1 mM methyl viologen (MV^{+2}) in a 0.1 M phosphate buffer solution of pH 7 at a room temperature.
- (2) Four graphite felt electrodes were cut into a similar size of $4\text{ mm} \times 4\text{ mm} \times 2\text{ mm}$ and washed with distilled water before treating.
- (3) Suspensions of carbon nanomaterials were prepared by mixing 0.7 mg of each nanomaterial with $700\text{ }\mu\text{L}$ of N,N-Dimethylformamide (DMF) as the dispersing agent and agitating the mixture using a sonicating tip (Ultrasonic Processor, Sonics Vibra Cell) for 1 h. Close to $10\text{ }\mu\text{L}$ of the resultant solution (i.e., the nanomaterial and DMF solution) was dropped directly onto the GF electrode surface and was allowed to dry at 40°C for 45 min to evaporate the solvent.
- (4) The electrodes were tested following steps 1 to 5 described in the procedure given in Section 2.2.1.

3. Results and Discussions

3.1. Electrochemical Properties. The maximum power density was evaluated through the examination of a polarization curve, which characterises voltage as a function of current. The power production over a range of current densities was obtained by changing the external resistance R_{ext} using a resistor box, when the voltage production became stable. The MFC reactor was initially operated under an open circuit condition. Once the reactor achieved a stable voltage output of 0.349 V, the resistor box was switched on and the external resistance was varied from 10Ω to $10\text{ k}\Omega$ in steps of 250Ω every 10 min and the cell voltage was measured at each resistance. Current and power levels were calculated with the voltage and resistance based on Ohm's law. Current and power densities were calculated by normalising the current and voltage through an electrode surface area. Polarization and power density curves are displayed in Figure 1(a). The polarization curves illustrating the three characteristic regions of voltage drop in the MFC are shown in Figure 1(b). These regions include a rapid voltage decrease due to the flow of current through high external resistance, an almost constant decrease in voltage, and a second significant voltage drop at high current densities. The decrease in the cell voltage is a consequence of electrode overpotentials (activation, bacterial metabolic, and mass transfer losses) and ohmic losses. The maximum power density obtained was 1.26 mW/m^2 at a current density of 10.23 mA/m^2 . This is slightly lower than that reported in [29], where the highest power density achieved was 1.7 mW/m^2 at a current density of 15 mA/m^2 .

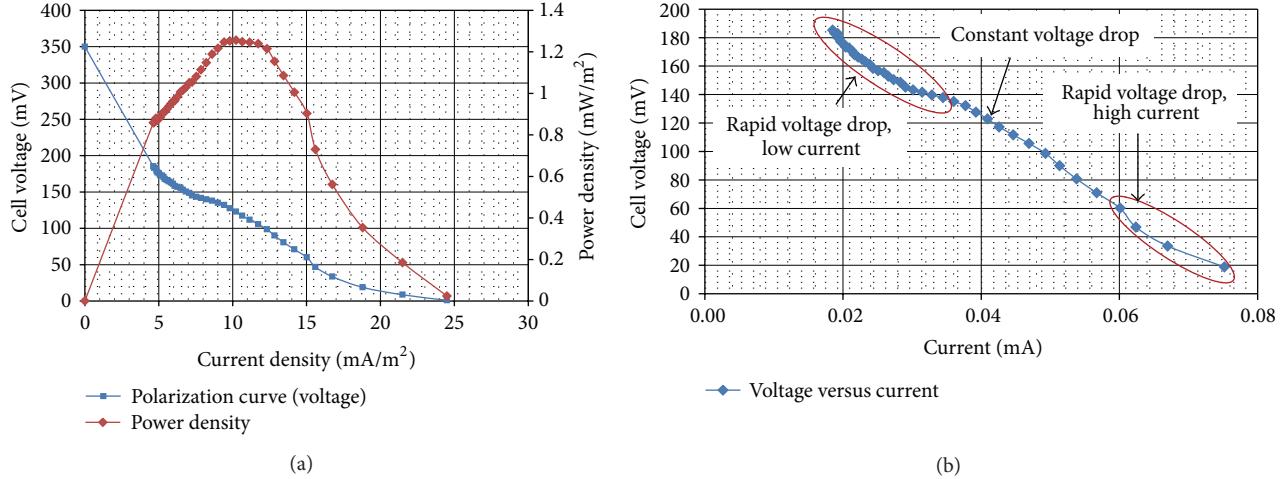


FIGURE 1: (a) Polarization curves. (b) The three characteristic regions of voltage drop.

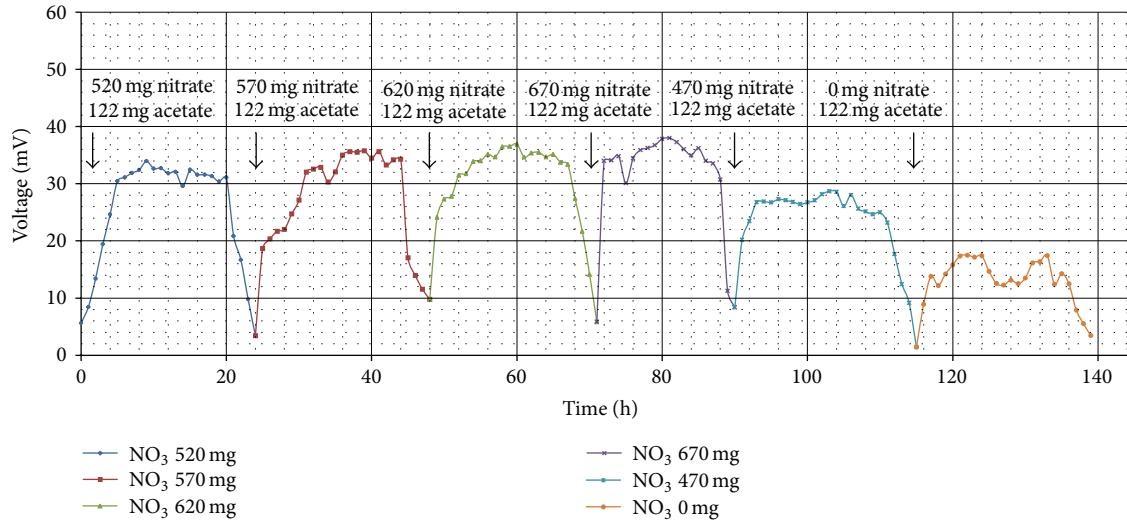


FIGURE 2: Effect of nitrate concentrations on voltage generation.

This is a result of two factors: firstly, [29] used a graphite-Mn(IV) and a graphite-Fe(III) electrode as an anode and a cathode, respectively; however, a pure GF electrode is used here in both chambers. Modification of GF electrodes using nanomaterials and the improvements observed are discussed in Section 3.5. Furthermore, the results achieved in this study were based on the enrichment of soil inoculum only; however, anaerobic digester sludge was used in [29]. Based on the slope of the linear region of the polarization curve, an internal resistance R_{int} of 2893Ω could be determined. It is also shown in Figure 1(a) that the power output was maximal when $R_{\text{int}} = R_{\text{ext}} = 3000 \Omega$.

3.2. Effects of Nitrate Concentration on MFC Performance.

The effects of different nitrate concentrations on denitrification activity and voltage output were investigated at a fixed

external resistance of 500Ω . The MFC system was operated with a number of batches at a range of nitrate concentrations between 470 mg/L and 670 mg/L in steps of 50 mg/L in the cathodic chamber, while a fixed acetate concentration of 122 mg/L was added in the anodic chamber at the beginning of each batch. An illustration of typical profiles of cell voltages produced is shown in Figure 2. The denitrification activity was examined through the measurement of nitrate and nitrite concentrations at the end of each batch. The denitrification activity (including nitrate reduction and nitrite accumulation rates) and current generation achieved are outlined in Table 1.

The results showed that both nitrate removal and nitrite accumulation rates exhibited an increase with the increase of nitrate concentration. Increasing the nitrate concentration from 470 mg/L to 670 mg/L improved the denitrification activity by almost 2% (from 87.32% to 89.3%). However, an increase in nitrite accumulation by a factor of approximately

TABLE 1: Nitrate removal and nitrite accumulation of several closed batches fed with 122 mg/L acetate and different nitrate concentrations.

Nitrate concentration (mg/L)	Nitrate removal mg/L	Nitrate removal %	Nitrite accumulation mg/L	Average current (μ A)
470	410.39	87.32	5.153	46
520	455.12	87.524	6.626	52.4
570	500.3	87.8	10.31	54
620	547.26	88.27	12.52	61
670	598.21	89.3	16.93	65

3 (from 5.15 mg/L to 16.93 mg/L) was incurred. Previous work [29] suggested that the current production was dependent on glucose and nitrate. The experimental results obtained here confirmed this suggestion and also demonstrated that the denitrification rate was supported by the current production (which was dependent on nitrate). However, the system was also operated with the absence of nitrate as an electron acceptor in the cathodic chamber, and a low electrical current was observed.

3.3. Effects of Acetate Concentration on Denitrification Activity. To investigate the effect of acetate on the MFC performance, the fuel cell was fed with several batches at different acetate concentrations in the anodic chamber. Acetate concentrations of 72, 122, 172, 222, and 272 mg/L were used. Each batch was performed with an external resistance of 500 Ω at a nitrate concentration of 520 mg/L in the cathodic chamber. Cell voltage profiles produced are shown in Figure 3. The denitrification rate and current production obtained are given in Table 2. The addition of sodium acetate can allow the bacteria to provide more electrons and to increase the voltage output. This was observed as an increase in the cell voltage when the concentration of acetate was increased. Meanwhile, increasing acetate concentration resulted in an increase in the current generation. In the cathodic chamber, an increase in the nitrate removal from 85.6% to 92.23%, together with a decrease in the nitrite accumulation from 7.4 mg/L to 5.2 mg/L, was observed when the acetate concentration increased from 72 mg/L to 272 mg/L. This indicates that when acetate concentration of 272 mg/L was used, 1.08% of the nitrate removed was turned into nitrite. The remaining 91.15% (474.97 mg/L) of the nitrate removed was possibly converted into nitrogen gas. The experimental results proved that the sodium acetate highly supports the denitrification activity. However, the MFC performance was evaluated in the absence of sodium acetate and a denitrification rate of 71.7% was achieved while producing low cell current. This is due to the bacteria making use of the carbon stored internally by the cells from the previous batches, where carbon sources are used. It has to be noted that operating the MFC system in carbon starvation mode (i.e., without using a carbon source)

during the enrichment forced the bacteria to store carbon internally.

3.4. Effects of External Resistance on Denitrification Activity.

To investigate the effect of electricity generation on the denitrification process, further experiments were conducted with an MFC under different external resistances of 500, 5000, and 10000 Ω . Sodium acetate and nitrate concentrations of 122 mg/L and 520 mg/L, respectively, were used. Average currents of 52 μ A (500 Ω), 22 μ A (5 K Ω), and 14 μ A (10 K Ω) were produced, as shown in Table 3. The results confirmed that the denitrification rate is strongly dependent on the cell current produced, which was varied here by external resistance. It was observed that using higher resistance led to low cell current, resulting in lower nitrate removal rates and higher accumulation rates. This was possibly due to insufficient electron donors being available on the cathode.

3.5. MFC Performance with Modified GF Electrodes. The interaction between the bacterial biofilm and the electrode surface area can affect the MFC's performance. The reactants transport, including substrate, electrons, and electron accepting species from the bulk to the electrode surface, and the reaction kinetics on the electrodes surfaces can also influence the MFC's performance. The reaction kinetics on the electrode surface and the mass transfer are affected by the electrode materials [30], surface chemical properties of the electrodes [31], size and shape of the electrodes [32], and biofilm condition [33]. The electrode should provide a good environment for the bacteria to attach to and transport electrons, a large surface area, and a high conductivity [30]. The internal resistance of an MFC consists of two parts: nonohmic and ohmic resistances [34]. The former comprises a charge transfer resistance and a diffusion resistance [35], and these can be reduced by increasing the electrode surface area as well as selecting electrodes with good catalytic abilities. Ohmic resistance can be decreased by arranging the electrodes closely, using solutions with high conductivity and using a membrane with low resistivity.

Modification of the GF electrode surface to enhance its reaction kinetics and mass transfer is a good way to improve the MFCs performance. The GF electrode was modified with carbon nanomaterials, including SWCNTs, GCB, CNFs, and GCNFs, which are promising materials that can provide great stability and high conductivity. The electrochemical behaviour of the modified GF electrodes was investigated by using cyclic voltammetry. Cyclic voltammograms of methyl viologen redox reactions on different modified GF electrodes were recorded and compared. Comparisons of the kinetic methyl viologen redox reactions on the unmodified GF electrode were also considered. The results indicated an improvement in the electrochemical activity of the methyl viologen on the modified electrodes compared to the unmodified electrode. The GCNFs modified electrode exhibited the best electrochemical activity among the other modified electrodes, due to its having the largest surface area which greatly increased the rate of electron transfer. Therefore it was chosen to enhance the MFC's performance. In an

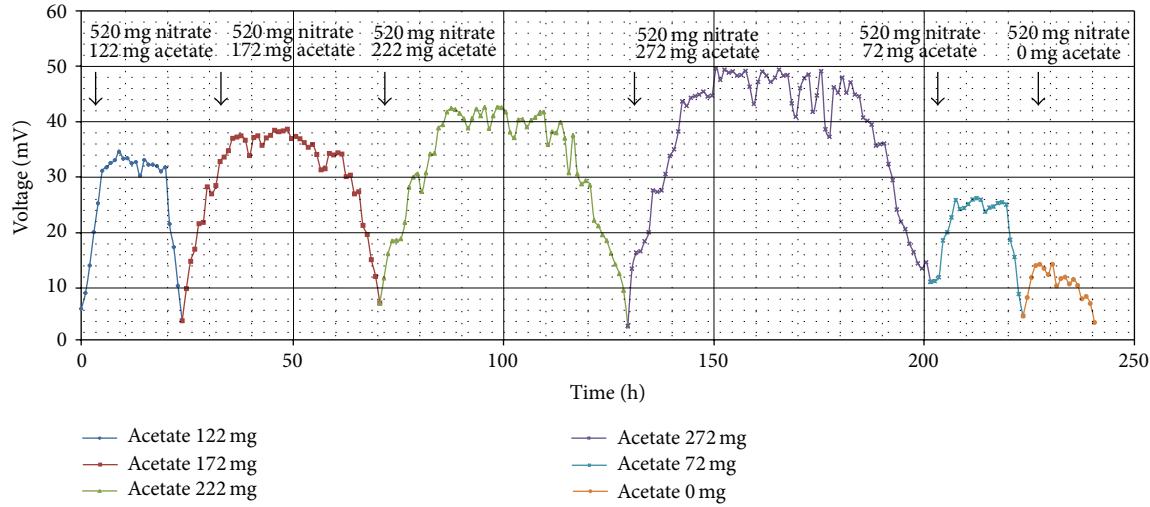


FIGURE 3: Voltage generation using acetate at different concentrations.

TABLE 2: Nitrate removal and nitrite accumulation of several closed batches fed with 122 mg/L acetate and different nitrate concentrations.

Acetate concentration (mg/L)	Nitrate removal mg/L	Nitrate accumulation mg/L	Average current (μ A)
	%		
72	445.01	85.6	7.4
122	455.12	87.5	52.4
172	466.8	89.8	59
222	472.3	90.83	63
272	480.17	92.2	74

MFC reactor, both anode and cathode were provided by GCNFs modified electrodes. A modified MFC system was enriched following the procedures described in Section 2.1.2 for three months, and its performance was evaluated in terms of power generation and nitrate removal. The results demonstrated that the GCNFs modified MFC system offered about 8% nitrate reduction rate higher than that achieved using unmodified electrodes (the unmodified MFC system removed 87.5% of the nitrate). This is due to the long term stability provided by the GCNFs modified MFC system, where an average of 35 mV was obtained over a period of 44 h; see Figure 4.

4. Conclusions

Nitrogen and nitrate have important environmental impact and there are many problems associated with their release into the environment. Nitrate can be removed from water through the use of physicochemical methods. However, these processes are not viable and their use is problematic. An alternative promising and versatile approach that can be used for nitrate removal is biological denitrification. However a major drawback is the potential bacterial contamination of treated water; thus additional filtration and disinfection are

TABLE 3: Nitrate removal, nitrite accumulation, average current, and CE of several closed batches fed with 520 mg/L nitrate and 122 mg/L acetate concentrations under different external resistances: 500, 5000, and 10000 Ω .

External resistance (Ω)	Nitrate removal mg/L	Nitrite accumulation mg/L	Average current (μ A)
	%		
500	455.12	87.5	6.6
5,000	434.98	83.65	22.2
10,000	412.5	79.33	14

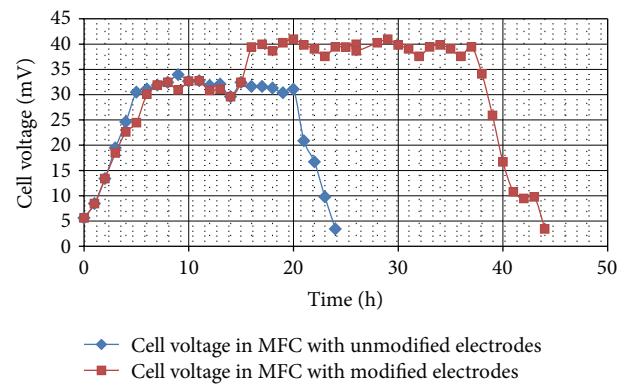


FIGURE 4: Voltage generation in MFC with modified and unmodified electrodes.

potentially needed to meet current drinking water standards. The stated aims and objectives of this study were to investigate the electrochemical removal of nitrate from water and associated technical problems. The results of the work showed that these aims were successfully achieved setting a good platform for future work in this area.

In order to perform biological nitrate removal, a mediatorless H-shaped MFC was constructed using bacteria both in the anode and cathode. A mixed bacterial culture in

the cathode performed denitrification through the use of electrons supplied by a mixed bacterial culture oxidising acetate in the anode. The effects of acetate/nitrate on current generation and nitrate removal and the denitrification activity as a function of external resistance were also studied. The results demonstrated that the denitrification rate is highly dependent on current production, which is influenced by external resistance and acetate and nitrate concentrations. An increase in the nitrate reduction rate was observed when the nitrate concentration increased. However, an increase in the nitrite accumulation rate was also induced which can be problematic due to its toxicity. In contrast, increasing the acetate concentration improved nitrate removal rates and reduced nitrite accumulation, with these results indicating the importance of the feed ratios for the two compartments. Higher external resistances were shown to inhibit the denitrification activity due to lower electrical currents produced. In addition, a polarization curve was obtained by changing the external resistance using a resistor box, and a maximum power density of 1.26 mW/m^2 was achieved at a current density of 10.23 mA/m^2 . This was also compared well with previous findings.

The electrochemical behaviours of a graphite felt electrode were investigated through the use of the cyclic voltammetry technique. Enhanced redox behaviours were achieved by modifying the GF electrode surface using carbon nanomaterials, such as SWCNTs, GCB, CNFs, and GCNFs. The electrochemical properties associated with the modified GF electrodes were studied, evaluated, and compared with those of the unmodified GF electrode. The GCNFs modified electrode offered the best electrochemical activity compared to the other modified electrodes, due to the large surface area provided and the improved electron transfer rate. Using these types of electrodes should improve the interaction with the microbes as these particles are much smaller than the felt surfaces giving the microbe more surface to directly interact with. Enhancing the reaction kinetics and mass transfer of the GF electrode through the GCNFs material helped improve the MFC's performance. The maximum voltage obtained was 40.94 mV and more than 95% of nitrate was removed, as compared with the unmodified electrode (33.95 mV maximum voltage, 87.5% nitrate removal). In this study only a few conditions of electrode modification were investigated. There are many possibilities that need to be investigated to see if the electrode's performance can be improved further. The optimisation study could have a strong impact on potential development of MFC technology.

References

- [1] B. Cohen, "The bacteria culture as an electrical half-cell," *Journal of Bacteriology*, vol. 21, pp. 18–19, 1931.
- [2] Z. He and L. T. Angenent, "Application of bacterial biocathodes in microbial fuel cells," *Electroanalysis*, vol. 18, no. 19–20, pp. 2009–2015, 2006.
- [3] P. Clauwaert, K. Rabaey, P. Aelterman et al., "Biological denitrification in microbial fuel cells," *Environmental Science and Technology*, vol. 41, no. 9, pp. 3354–3360, 2007.
- [4] K. B. Gregory, D. R. Bond, and D. R. Lovley, "Graphite electrodes as electron donors for anaerobic respiration," *Environmental Microbiology*, vol. 6, no. 6, pp. 596–604, 2004.
- [5] O. Lefebvre, A. Al-Mamun, and H. Y. Ng, "A microbial fuel cell equipped with a biocathode for organic removal and denitrification," *Water Science and Technology*, vol. 58, no. 4, pp. 881–885, 2008.
- [6] H. Liu, S. Cheng, and B. E. Logan, "Power generation in fed-batch microbial fuel cells as a function of ionic strength, temperature, and reactor configuration," *Environmental Science and Technology*, vol. 39, no. 14, pp. 5488–5493, 2005.
- [7] B. Min, S. Cheng, and B. E. Logan, "Electricity generation using membrane and salt bridge microbial fuel cells," *Water Research*, vol. 39, no. 9, pp. 1675–1686, 2005.
- [8] K. Rabaey, P. Clauwaert, P. Aelterman, and W. Verstraete, "Tubular microbial fuel cells for efficient electricity generation," *Environmental Science and Technology*, vol. 39, no. 20, pp. 8077–8082, 2005.
- [9] S. You, Q. Zhao, J. Zhang, H. Liu, J. Jiang, and S. Zhao, "Increased sustainable electricity generation in up-flow air-cathode microbial fuel cells," *Biosensors and Bioelectronics*, vol. 23, no. 7, pp. 1157–1160, 2008.
- [10] B. Sun and M. Skyllas-Kazacos, "Chemical modification and electrochemical behaviour of graphite fibre in acidic vanadium solution," *Electrochimica Acta*, vol. 36, no. 3–4, pp. 513–517, 1991.
- [11] B. Sun and M. Skyllas-Kazacos, "Chemical modification of graphite electrode materials for vanadium redox flow battery application—part II. Acid treatments," *Electrochimica Acta*, vol. 37, no. 13, pp. 2459–2465, 1992.
- [12] S. Hrapovic, Y. Liu, K. B. Male, and J. H. T. Luong, "Electrochemical biosensing platforms using platinum nanoparticles and carbon nanotubes," *Analytical Chemistry*, vol. 76, no. 4, pp. 1083–1088, 2003.
- [13] J. Wang and M. Musameh, "Enzyme-dispersed carbon-nanotube electrodes: a needle microsensor for monitoring glucose," *Analyst*, vol. 128, no. 11, pp. 1382–1385, 2003.
- [14] J. J. Gooding, R. Wibowo, J. Liu et al., "Protein electrochemistry using aligned carbon nanotube arrays," *Journal of the American Chemical Society*, vol. 125, no. 30, pp. 9006–9007, 2003.
- [15] J. M. Nugent, K. S. V. Santhanam, A. Rubio, and P. M. Ajayan, "Fast electron transfer kinetics on multiwalled carbon nanotube microbundle electrodes," *Nano Letters*, vol. 1, no. 2, pp. 87–91, 2001.
- [16] H. Verweij, M. C. Schillo, and J. Li, "Fast mass transport through carbon nanotube membranes," *Small*, vol. 3, no. 12, pp. 1996–2004, 2007.
- [17] J. Lee, N. T. Phung, I. S. Chang, B. H. Kim, and H. C. Sung, "Use of acetate for enrichment of electrochemically active microorganisms and their 16S rDNA analyses," *FEMS Microbiology Letters*, vol. 223, no. 2, pp. 185–191, 2003.
- [18] A. P. Borole, C. Y. Hamilton, T. A. Vishnivetskaya et al., "Integrating engineering design improvements with exoelectrogen enrichment process to increase power output from microbial fuel cells," *Journal of Power Sources*, vol. 191, no. 2, pp. 520–527, 2009.
- [19] C. Sukkasem, S. Xu, S. Park, P. Boonsawang, and H. Liu, "Effect of nitrate on the performance of single chamber air cathode microbial fuel cells," *Water Research*, vol. 42, no. 19, pp. 4743–4750, 2008.
- [20] L. S. Clesceri, A. E. Greenberg, and A. D. Eaton, "Standard methods for the examination of water and wastewater," in

- Proceedings of the 20th American Water Works Association*, Washington, DC, USA, 1999.
- [21] L. Merino, "Development and validation of a method for determination of residual nitrite/nitrate in foodstuffs and water after zinc reduction," *Food Analytical Methods*, vol. 2, no. 3, pp. 212–220, 2009.
 - [22] S. Alehashem, F. Chambers, J. W. Strojek, G. M. Swain, and R. Ramesham, "Cyclic voltammetric studies of charge transfer reactions at highly boron-doped polycrystalline diamond thin-film electrodes," *Analytical Chemistry*, vol. 67, no. 17, pp. 2812–2821, 1995.
 - [23] E. Steckhan and T. Kuwana, "Spectroelectrochemical study of mediators I. Bipyridinium salts and their electron transfer rates to cytochrome c," *Berichte der Bunsengesellschaft für physikalische Chemie*, vol. 78, no. 3, pp. 253–259, 1974.
 - [24] E. M. Kosower and J. L. Cotter, "Stable free radicals. II. The reduction of 1-methyl-4-cyanopyridinium ion to methylviologen cation radical," *Journal of the American Chemical Society*, vol. 86, no. 24, pp. 5524–5527, 1964.
 - [25] H. T. van Dam and J. J. Ponjee, "Electrochemically generated colored films of insoluble viologen radical compounds," *Journal of the Electrochemical Society*, vol. 121, no. 12, pp. 1555–1558, 1974.
 - [26] R. R. Lilienthal and D. K. Smith, "Solvent effects on the redox-dependent binding properties of a viologen-based receptor for neutral organic molecules," *Analytical Chemistry*, vol. 67, no. 20, pp. 3733–3739, 1995.
 - [27] H.-H. Yang and R. L. McCreery, "Effects of surface monolayers on the electron-transfer kinetics and adsorption of methyl viologen and phenothiazine derivatives on glassy carbon electrodes," *Analytical Chemistry*, vol. 71, no. 18, pp. 4081–4087, 1999.
 - [28] X.-G. Li, K.-L. Huang, S.-Q. Liu, N. Tan, and L.-Q. Chen, "Reaction mechanism of V(IV)/V(V) redox couple at graphite felt composite electrode bonded with conductive carbon plastic," *Journal of Electrochemistry*, vol. 12, pp. 368–372, 2006.
 - [29] Y.-H. Jia, H.-T. Tran, D.-H. Kim et al., "Simultaneous organics removal and bio-electrochemical denitrification in microbial fuel cells," *Bioprocess and Biosystems Engineering*, vol. 31, no. 4, pp. 315–321, 2008.
 - [30] B. E. Logan, *Microbial Fuel Cells*, John Wiley & Sons, Hoboken, NJ, USA, 2008.
 - [31] V. Debabov, "Electricity from microorganisms," *Microbiology*, vol. 77, no. 2, pp. 123–131, 2008.
 - [32] P. Aelterman, M. Verschaele, M. Marzorati, N. Boon, and W. Verstraete, "Loading rate and external resistance control the electricity generation of microbial fuel cells with different three-dimensional anodes," *Bioresource Technology*, vol. 99, no. 18, pp. 8895–8902, 2008.
 - [33] K. Y. Cheng, G. Ho, and R. Cord-Ruwisch, "Affinity of microbial fuel cell biofilm for the anodic potential," *Environmental Science and Technology*, vol. 42, no. 10, pp. 3828–3834, 2008.
 - [34] Y. Fan, E. Sharbrough, and H. Liu, "Quantification of the internal resistance distribution of microbial fuel cells," *Environmental Science and Technology*, vol. 42, no. 21, pp. 8101–8107, 2008.
 - [35] J. Larminie and A. Dicks, *Fuel Cell Systems Explained*, John Wiley & Sons, Chichester, UK, 2000.

Research Article

Study of Intelligent Photovoltaic System Fault Diagnostic Scheme Based on Chaotic Signal Synchronization

Chin-Tsung Hsieh and Jen Shiu

Department of Electrical Engineering, National Chin-Yi University of Technology, No. 57, Section 2, Zhongshan Road, Taiping District, Taichung 41170, Taiwan

Correspondence should be addressed to Chin-Tsung Hsieh; fred@ncut.edu.tw

Received 14 July 2013; Revised 10 September 2013; Accepted 18 September 2013

Academic Editor: Dumitru Baleanu

Copyright © 2013 C.-T. Hsieh and J. Shiu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As the photovoltaic system consists of many equipment components, manual inspection will be very costly. This study proposes the photovoltaic system fault diagnosis based on chaotic signal synchronization. First, MATLAB was used to simulate the fault conditions of solar system, and the maximum power point tracking (MPPT) was used to ensure the system's stable power and capture and record the system fault feature signals. The dynamic errors of various fault signals were extracted by chaotic signal synchronization, and the dynamic error data of various fault signals were recorded completely. In the photovoltaic system, the captured output voltage signal was used as the characteristic values for fault recognition, and the extension theory was used to create the joint domain and classical domain of various fault conditions according to the collected feature data. The matter-element model of extension engineering was constructed. Finally, the whole fault diagnosis system is only needed to capture the voltage signal of the solar photovoltaic system, so as to know the exact fault condition effectively and rapidly. The proposed fault diagnostor can be implemented by embedded system and be combined with ZigBee wireless network module in the future, thus reducing labor cost and building a complete portable renewable energy system fault diagnostor.

1. Introduction

The rapid development of technology has led to over dependence on energy. Although oil is the most important energy source at present, it will be in shortage in the next 50 years. In order to solve this problem, humans have begun to seek alternative energy sources which will not harm the earth, such as renewable energy sources. The renewable energy sources are the energy sources derived from the nature, such as solar energy, wind power, tidal power, and geothermal heat, which are inexhaustible energy sources. Fossil energy is exhaustible, sunlight can meet 2,850 times of global energy needs, wind energy can meet 200 times of global energy needs, hydraulic power can meet 3 times of global energy needs, biomass energy can meet 20 times of global energy needs, and geothermal heat can meet 5 times of global energy needs.

This study focused on solar photovoltaic power generation. In the regions with abundant light supply (e.g., space exposed to the sun, ocean, coast, and open rocky area), solar

energy is supplied continuously, and the production process does not pollute the environment or consume other earth resources or cause greenhouse effect. Solar energy is derived from the sun inexhaustibly and is as good as water power or wind power. Many countries are developing solar energy actively. The solar energy facilities can be constructed facilities, such as wind energy facilities, protecting many lands and ecologies. The solar module can be mounted on buildings, which is called photovoltaic integrated buildings. The solar panel can generate electric power in sunlight, insulate heat, reduce the temperature inside buildings effectively, and reduce the building energy consumption. Solar energy is a renewable energy that can be conveniently obtained. As long as the solar panel is placed in an open area with long sunshine, the solar energy can be obtained effectively. However, as the equipment is set in open area for a long time, it is likely to be destroyed by natural disasters and to age. These factors will reduce the efficiency of solar photovoltaic system. Once a fault occurs in the solar module, it must be inspected, but the

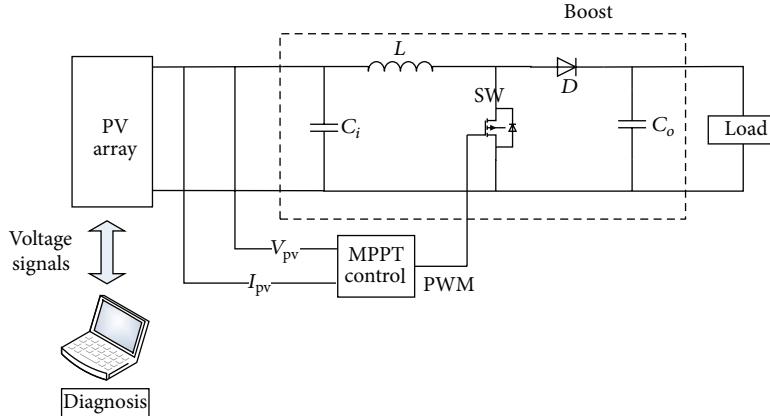


FIGURE 1: Solar photovoltaic system structure.

faults may be difficult to detect with naked eye. The detection is labor and cost consuming. This study thus proposes an intelligent solar photovoltaic system fault diagnositor. When small solar power systems are popularized and combined with commercial power, the household solar power system is likely to have faults; thus, the solar photovoltaic system fault diagnositor will be an important tool. The research and development of solar energy fault diagnositor will be very helpful to solar power systems, not only shortening manual detection time and reducing risks but also detecting the fault point effectively.

Most current studies on the fault diagnosis for solar photovoltaic system use software simulated or measured power generation data of solar photovoltaic system as the fault diagnosis data. As the output power of solar photovoltaic system is influenced by the sunshine intensity and module surface temperature significantly and the solar power system data are nonlinear system, many studies use software for simulation. Several diagnostic methods for solar photovoltaic system are introduced as follows. Literature [1] proposed a diagnostic method based on error backpropagation algorithm (BP), combined with the unique characteristics of calculator and artificial neural network, such as rapid development of complexity, association, memory, adaptability, self-learning, and processing complex multimode. The artificial neural network [1, 2] is applicable to fault diagnosis, and BP neural network is the most used one. The input signal is transmitted from the input layer and hidden layer to the output layer. The output end generates the output signal. The weight of network is fixed in initial propagation, and the state of neuron of a layer only influences the state of neuron of next layer. When the output signal is not as good as the output layer expected, the error backpropagation is triggered. In the error signal backpropagation, the weight of network can adjust the error signal. However, the defect in neural network is that the optimal number of neurons of hidden layer should be selected through multiple experiments, and a long training is required in order to converge to the system. Literature [3] used MATLAB software to simulate the characteristics of solar photovoltaic system. Equivalent mathematical equation

was used to simulate the characteristic curves of photovoltaic system in sunlight and in shadow. The simulation system sets the series-parallel connection of modules and photovoltaic module shadow coverage area, whether or not to mount bypass diode and blocking diode. The simulation effect was very close to the actual result, but the equivalent parameter changed when a fault occurred. Thus, the defect in the above method is that the fault parameter is difficult to be calculated when the photovoltaic system has other faults. Literature [4, 5] used PSIM software to build a solar photovoltaic system. The characteristic curve of simulation result was similar to the actual operation curve, and the extension theory was applied to diagnose the fault category, so the diagnosis effect was very good. However, the curve of simulated photovoltaic system fault in [4, 5] was slightly different from the actual faulty system. The simulation system did not consider the influence of bypass diode and blocking diode on the photovoltaic system, and the actual photovoltaic system was equipped with bypass diode and blocking diode. Therefore, there was a pit in IV curve when the photovoltaic system had a fault, so that the maximum power point shifted, resulting in multiple peaks of PV characteristic curve. This paper proposes a fault diagnosis method based on chaotic synchronization signal, combined with the intelligent classification of extension theory. The chaotic system is very sensitive to the change in system parameter, and the system parameter of system must be changed if the solar photovoltaic system has a fault. Therefore, this method only needs to import the signal of the photovoltaic system with faults into the chaotic synchronization system according to the variance in system parameter and to capture the variance in dynamic error. It is imported into the extension theory to distinguish the fault state effectively, and the extension theory does not need learning time, so the diagnosis is rapid.

2. Photovoltaic System

The main function of solar photovoltaic system is to convert the solar irradiation into electrical energy by solar photovoltaic cell and an electric power converter. The equipped

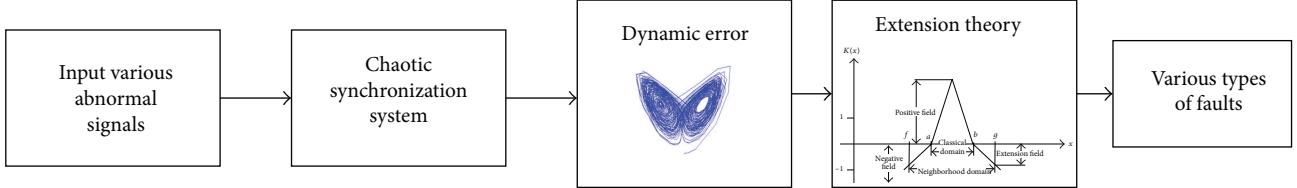


FIGURE 2: Chaotic signal synchronization and extension diagnostic system flow chart.

electric power converter can stabilize, increase and reduce the output voltage, or convert the frequency according to the requirement of output load. The output power depends on the area of solar photovoltaic cell, conversion efficiency, solar illumination, ambient temperature, and effect of load. The solar photovoltaic array used in this study was a $10 * 2$ array, equipped with a boost converter and MPPT algorithm, as shown in Figure 1. The MPPT algorithm guarantees the maximum power output when the solar photovoltaic array has faults. The voltage variation is measured in the system operation to distinguish the fault.

The power characteristic generated by solar photovoltaic cell is not linear, and the power of solar photovoltaic cell varies with current sunshine intensity and ambient temperature. In order to keep the output power at the maximum value of the characteristic curve [6], the solar photovoltaic system must be equipped with MPPT control to guarantee the maximum output power, so as to maximize the system output efficiency. At present, the common incremental conductance MPPT method [7] is used in this paper.

3. Proposed Fault Diagnosis Method

3.1. Chaotic Synchronization System. As the chaotic system is very sensitive to the change in system parameter [8–14], this study used this characteristic to capture the solar photovoltaic system fault voltage signal for chaotic synchronization transformation. The kinematic trajectories of dynamic error were extracted, and these kinematic trajectories were the fault features; thus, the extension theory was used to recognize fault conditions. The Lorenz chaotic synchronization system is expressed as follows [8]:

$$\text{Master: } \begin{cases} \dot{x}_1 = \alpha(x_2 - x_1), \\ \dot{x}_2 = \beta x_1 - x_1 x_3 - x_2, \\ \dot{x}_3 = x_1 x_2 - \gamma x_3, \end{cases} \quad (1)$$

$$\text{Slave: } \begin{cases} \dot{y}_1 = \alpha(y_2 - y_1), \\ \dot{y}_2 = \beta y_1 - y_1 y_3 - y_2, \\ \dot{y}_3 = y_1 y_2 - \gamma y_3. \end{cases} \quad (2)$$

Dynamic error of system is as follows:

$$\begin{bmatrix} \dot{e}_1 \\ \dot{e}_2 \\ \dot{e}_3 \end{bmatrix} = \begin{bmatrix} -\alpha & \alpha & 0 \\ \beta & -1 & 0 \\ 0 & 0 & -\gamma \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} + \begin{bmatrix} 0 \\ -y_1 y_3 + x_1 x_3 \\ y_1 y_2 - x_1 x_2 \end{bmatrix}, \quad (3)$$

TABLE 1: Solar photovoltaic system fault category.

Fault category	Fault conditions
PF1	Normal operation
PF2	Two series branches, one fault in any branch
PF3	Two series branches, two faults in any branch
PF4	Two series branches, three faults in any branch
PF5	Two series branches, one fault in each branch
PF6	Two series branches, two faults in each branch
PF7	Two series branches, one fault in one branch, two faults in the other branch
PF8	Two series branches, one fault in one branch, three faults in the other branch
PF9	Two series branches, two faults in one branch, three faults in the other branch

where α , β , and γ are positive constants. According to the condition of chaos theory [8], if the eigenvalues of system are negative, the system state is determined as stable, so that the chaotic attractor can be generated. The kinematic trajectory of chaotic attractor is used for various studies. This paper observes the kinematic trajectory of chaotic attractor to distinguish the fault type of solar photovoltaic system.

3.2. Extension Theory. The extension is the real number extended from fuzzy set range $[0, 1]$ to $[-\infty, \infty]$, representing the degree of a property of things. First, the maximum and minimum extents of characteristics of real number are determined. The maximum extent is set as a' and the minimum extent is b' , and the real domain $[a', b']$ is obtained. The classical domain range $[a, b]$ can be worked out by (4) and (5), where α is the extend factor of classical domain, aiming at extension considering light error resulted from human or irresistible factors in actual measurement. This region is called classical domain [15].

Consider the following:

$$a = (1 - \alpha) a', \quad b = (1 + \alpha) b', \quad (4)$$

$$f = (1 - \beta) a, \quad g = (1 + \beta) b. \quad (5)$$

In order to avoid omitting important features outside the classical domain range $[a, b]$, the extension rate β is proposed to obtain the joint domain. The joint domain $[f, g]$ is obtained by (6). Finally, the obtained classical domain $X_o = [a, b]$ and joint domain $X = [f, g]$ are two intervals in the real number

TABLE 2: Matter-element model.

Simulations				Experiments													
Neighborhood Domain	$c_1 \langle 200, 520 \rangle$	$c_2 \langle 200, 520 \rangle$		PF1 PF2 PF3 PF4 PF5 PF6 PF7 PF8 PF9	$c_1 \langle 485, 490 \rangle$ $c_2 \langle 485, 490 \rangle$	$c_1 \langle 335, 345 \rangle$ $c_2 \langle 335, 345 \rangle$	$c_1 \langle 335, 345 \rangle$ $c_2 \langle 335, 345 \rangle$	Neighborhood Domain	$c_1 \langle 200, 400 \rangle$	$c_2 \langle 200, 400 \rangle$		PF1 PF2 PF3 PF4 PF5 PF6 PF7 PF8 PF9	$c_1 \langle 378, 386 \rangle$ $c_2 \langle 375, 390 \rangle$	$c_1 \langle 349, 325 \rangle$ $c_2 \langle 323, 330 \rangle$	$c_1 \langle 340, 345 \rangle$ $c_2 \langle 335, 350 \rangle$	$c_1 \langle 312, 317 \rangle$ $c_2 \langle 310, 315 \rangle$	$c_1 \langle 255, 261.5 \rangle$ $c_2 \langle 252, 259 \rangle$
PF6 PF8	$c_1 \langle 286, 289 \rangle$	$c_2 \langle 286.5, 288.5 \rangle$		PF7 PF9	$c_1 \langle 313, 316 \rangle$ $c_2 \langle 313, 316 \rangle$	$c_1 \langle 215, 217 \rangle$ $c_2 \langle 215, 217 \rangle$		PF6 PF8	$c_1 \langle 395, 301 \rangle$ $c_2 \langle 292, 306 \rangle$	$c_1 \langle 262.5, 268 \rangle$ $c_2 \langle 266, 272 \rangle$		PF7 PF9	$c_1 \langle 312, 317 \rangle$ $c_2 \langle 310, 315 \rangle$	$c_1 \langle 255, 261.5 \rangle$ $c_2 \langle 252, 259 \rangle$			
PF8	$c_1 \langle 227.5, 228.5 \rangle$	$c_2 \langle 227.5, 228.5 \rangle$															

TABLE 3: Simulated fault diagnosis results.

Qty	ED _{PF1}	ED _{PF2}	ED _{PF3}	ED _{PF4}	ED _{PF5}	ED _{PF6}	ED _{PF7}	ED _{PF8}	ED _{PF9}	Results
1	1.000	-0.585	-0.839	-0.981	-0.722	-0.931	-0.893	-0.991	-1.000	PF1
2	1.000	-0.585	-0.839	-0.981	-0.722	-0.931	-0.983	-0.991	-1.000	PF1
3	-0.186	1.000	-0.224	-0.910	-0.337	-0.670	-0.485	-0.958	-1.000	PF2
4	-0.176	1.000	-0.224	-0.910	0.337	-0.670	-0.485	-0.958	-1.000	PF2
5	-1.000	-0.357	1.000	-0.549	-0.004	0.054	0.564	-0.662	-0.760	PF3
6	-1.000	-0.357	1.000	-0.549	-0.004	0.054	0.565	-0.662	-0.759	PF3
7	-1.000	-0.885	-0.642	1.000	-0.822	-0.258	-0.510	0.982	0.465	PF4
8	-1.000	-0.885	-0.642	1.000	-0.822	-0.258	-0.510	0.987	0.468	PF4
9	-0.584	0.406	0.366	-0.843	1.000	-0.420	-0.093	-0.926	-1.000	PF5
10	-0.584	0.406	0.369	-0.842	1.000	-0.419	-0.091	-0.926	-1.000	PF5
11	-1.000	-0.633	0.140	-0.007	-0.432	1.000	0.562	-0.162	-0.289	PF6
12	-1.000	-0.631	0.147	-0.007	-0.429	1.000	0.571	-0.161	-0.289	PF6
13	-1.000	-0.530	0.459	-0.170	-0.273	0.648	1.000	-0.308	-0.424	PF7
14	-1.000	-0.530	0.459	-0.176	-0.273	0.636	1.000	-0.313	-0.429	PF7
15	-1.000	-0.893	-0.668	0.855	-0.834	-0.312	-0.545	1.000	0.455	PF8
16	-1.000	-0.895	-0.668	0.855	-0.834	-0.322	-0.552	1.000	0.450	PF8
17	-1.000	-0.934	-0.796	0.136	-0.898	-0.578	-0.721	0.711	1.000	PF9
18	-1.000	-0.932	-0.791	0.169	-0.896	-0.566	-0.713	0.761	1.000	PF9

field $[-\infty, \infty]$. If x is a point of real number, the correlation function is defined as

$$K(x) = \begin{cases} -\rho(x, X_o), & x \in X_o, \\ \frac{\rho(x, X_o)}{\rho(x, X) - \rho(x, X_o)}, & x \notin X_o. \end{cases} \quad (6)$$

Figure 2 is the chaotic signal synchronization and extension diagnostic system flow chart. First, the measured voltage of solar photovoltaic system is recorded, and then the recorded voltage signal to be measured is imported into the slave system of chaotic signal synchronization system. The chaotic synchronization system generates the chaotic dynamic error signal after the subtraction between master system and slave system. The kinematic trajectory formed of chaotic dynamic

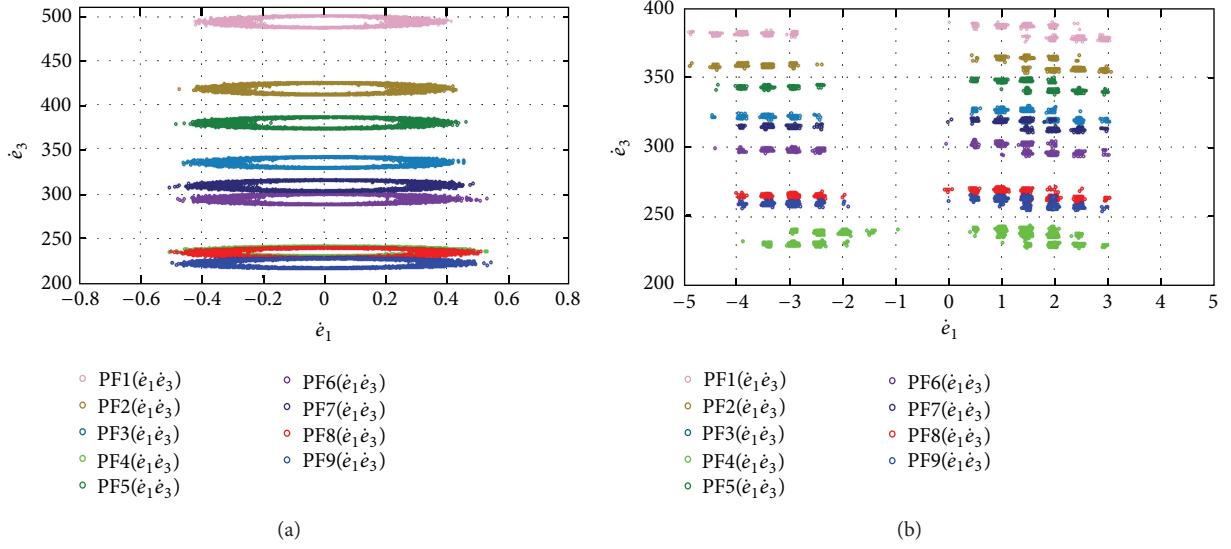
error signal is the main basis of extension diagnosis, which is the fault feature of solar photovoltaic system. Finally, as long as the chaotic dynamic error signal is imported into the finished extension matter-element model, the fault category can be identified rapidly and accurately by extension diagnosis of fault conditions.

4. Simulation and Experimental Results

The solar photovoltaic system was simulated by using MATLAB 2009 a. The solar photovoltaic array for experiment was Agilent Technologies E4360A modularized solar array simulator platform. The solar panel was SM 1611. The specifications are as follows: open-circuit voltage is 3 V, short-circuit current is 0.8 A, maximum power point voltage is 2.36 V, maximum power point current is 0.72 A, solar illumination is

TABLE 4: Actual fault diagnosis results.

Qty	ED _{PF1}	ED _{PF2}	ED _{PF3}	ED _{PF4}	ED _{PF5}	ED _{PF6}	ED _{PF7}	ED _{PF8}	ED _{PF9}	Results
1	1.000	-0.391	-0.790	-1.000	-0.641	-0.887	-0.843	-0.963	-0.978	PF1
2	1.000	-0.391	-0.790	-1.000	-0.641	-0.887	-0.843	-0.963	-0.978	PF1
3	-0.222	1.000	-0.476	-1.000	-0.103	-0.719	-0.610	-0.908	-0.945	PF2
4	-0.222	1.000	-0.476	-1.000	-0.103	-0.719	-0.610	-0.907	-0.945	PF2
5	-0.743	-0.391	1.000	-1.000	-0.042	-0.115	0.226	-0.710	-0.829	PF3
6	-0.743	-0.392	1.000	-1.000	-0.044	-0.116	0.224	-0.710	-0.829	PF3
7	-1.000	-0.964	-0.890	1.000	-0.931	-0.787	-0.858	-0.581	-0.448	PF4
8	-0.488	-0.964	-0.890	1.000	-0.931	-0.787	-0.858	-0.581	-0.448	PF4
9	-0.488	-0.074	-0.205	-1.000	1.000	-0.573	-0.408	-0.860	-0.917	PF5
10	-1.000	-0.073	-0.205	-1.000	1.000	-0.573	-0.408	-0.860	-0.917	PF5
11	-1.000	-0.817	-0.440	-0.812	-0.653	1.000	-0.280	-0.517	-0.638	PF6
12	-0.888	-0.817	-0.440	-0.821	-0.653	1.000	-0.280	-0.517	-0.638	PF6
13	-0.888	-0.572	0.145	-1.000	-0.272	-0.017	1.000	-0.678	-0.810	PF7
14	-1.000	-0.572	0.145	-1.000	-0.272	-0.017	1.000	-0.678	-0.810	PF7
15	-1.000	-0.865	-0.589	-0.074	-0.745	-0.205	-0.471	1.000	0.363	PF8
16	-1.000	-0.865	-0.589	-0.074	-0.745	-0.205	-0.472	1.000	0.363	PF8
17	-1.000	-0.889	-0.662	0.015	-0.791	-0.346	-0.566	0.285	1.000	PF9
18	-1.000	-0.889	-0.662	0.015	-0.791	-0.346	-0.565	0.286	1.000	PF9

FIGURE 3: (a) $\dot{e}_1 \dot{e}_3$ dynamic error plane of nine simulated fault conditions. (b) $\dot{e}_1 \dot{e}_3$ dynamic error plane of nine physical fault conditions.

1000 W/m², ambient temperature is 25°C, and the connection mode is 10 series and 2 parallel. The fault types are shown in Table 1. The voltage-current characteristic curve with faults was captured and recorded, and the voltage in nine cases was imported into the chaotic synchronization system, so that the chaotic synchronization system generated dynamic error. Figure 3(a) shows the $\dot{e}_1 \dot{e}_3$ dynamic error plane of nine simulated fault conditions. Figure 3(b) shows the $\dot{e}_1 \dot{e}_3$ dynamic error plane of nine actual fault conditions. The fault features are extracted as the base of extension diagnosis. Finally, the matter-element models are finished in Table 2 for

fault recognition. Finally, 18 data are imported for the system to distinguish. There are 18 data imported into simulated and actual diagnoses, respectively. The diagnostic results are shown in Tables 3 and 4. They show that the proposed scheme has good diagnostic performance.

5. Conclusion

This study diagnosed the faults in solar photovoltaic system based on chaotic signal synchronization and used extension theory to distinguish the fault category. As the chaotic

synchronization allows the system to have obvious nonlinear motion due to the change in system parameter, this obvious kinematic trajectory was used as the characteristic values of faults, and excellent diagnosis effect was obtained by the intelligent recognition of extension. The extension theory does not need iteration or learning, and its diagnosis is very fast in comparison to general neural diagnosis. Accurate, rapid, timesaving, and cost saving fault recognition could be achieved by capturing one voltage signal and two characteristic values of chaos. The cost was much lower than other traditional methods which need many sensors to capture physical quantities for distinguishing faults. As an intelligent fault diagnostor, its effect is excellent.

Conflict of Interests

The authors declare that they have no conflict of interests.

Acknowledgments

The author would like to thank Professor Her-Terng Yau for his valuable comments on chaos synchronization.

References

- [1] Y. Wu, Q. Lan, and Y. Sun, "Application of BP neural network fault diagnosis in solar photovoltaic system," in *Proceedings of the 2009 IEEE International Conference on Mechatronics and Automation (ICMA '09)*, pp. 2581–2585, August 2009.
- [2] M. H. Wang and C. P. Hung, "Extension neural network and its applications," *Neural Networks*, vol. 16, no. 5-6, pp. 779–784, 2003.
- [3] H. Patel and V. Agarwal, "MATLAB-based modeling to study the effects of partial shading on PV array characteristics," *IEEE Transactions on Energy Conversion*, vol. 23, no. 1, pp. 302–310, 2008.
- [4] K. H. Chao, S. H. Ho, and M. H. Wang, "Modeling and fault diagnosis of a photovoltaic system," *Electric Power Systems Research*, vol. 78, no. 1, pp. 97–105, 2008.
- [5] M. Veerachary, "PSIM circuit-oriented simulator model for the nonlinear photovoltaic sources," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 2, pp. 735–740, 2006.
- [6] K. Ding, X. G. Bian, H. H. Liu, and T. Peng, "A MATLAB-simulink-based PV module model and its application under conditions of nonuniform irradiance," *IEEE Transactions on Energy Conversion*, vol. 27, no. 4, pp. 864–872, 2012.
- [7] N. Onat, "Recent developments in maximum power point tracking technologies for photovoltaic systems," *International Journal of Photoenergy*, vol. 2010, Article ID 245316, 11 pages, 2010.
- [8] C. H. Huang, C. H. Lin, and C. L. Kuo, "Chaos synchronization-based detector for power-quality disturbances classification in a power system," *IEEE Transactions on Power Delivery*, vol. 26, no. 2, pp. 944–953, 2011.
- [9] W. H. Deng and C. P. Li, "Chaos synchronization of the fractional Lü system," *Physica A*, vol. 353, no. 1–4, pp. 61–72, 2005.
- [10] C. Li and G. Peng, "Chaos in Chen's system with a fractional order," *Chaos, Solitons and Fractals*, vol. 22, no. 2, pp. 443–450, 2004.
- [11] A. K. Golmankhaneh, R. Arefi, and D. Baleanu, "The proposed modified Liu system with fractional order," *Advances in Mathematical Physics*, vol. 2013, Article ID 186037, 6 pages, 2013.
- [12] A. K. Golmankhaneh, R. Arefi, and D. Baleanu, "Synchronization in a nonidentical fractional order of a proposed modified system," *Journal of Vibration and Control*, 2013.
- [13] D. Baleanu, K. Diethelm, E. Scalas, and J. J. Trujillo, *Fractional Calculus Models and Numerical Methods, Series on Complexity, Nonlinearity and Chaos*, World Scientific, Singapore, 2012.
- [14] S. Y. Xiang, W. Pan, B. Luo et al., "Synchronization of unpredictability-enhanced chaos in VCSELs with variable-polarization optical feedback," *IEEE Journal of Quantum Electronics*, vol. 47, no. 10, pp. 1354–1361, 2011.
- [15] C. T. Hsieh, M. H. Wang, and Y. P. Kuo, "PEM fuel cell diagnostic system using the extension theory," *International Journal of Distributed Sensor Networks*, vol. 2013, Article ID 539250, 7 pages, 2013.

Research Article

Epidemic Modelling by Ripple-Spreading Network and Genetic Algorithm

Jian-Qin Liao,^{1,2} Xiao-Bing Hu,^{1,3} Ming Wang,¹ and Mark S. Leeson³

¹ State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China

² Angel Women's & Children's Hospital, Chengdu 610041, China

³ School of Engineering, University of Warwick, Coventry CV4 7AL, UK

Correspondence should be addressed to Xiao-Bing Hu; dr.xiaobinghu@hotmail.co.uk

Received 20 July 2013; Revised 8 September 2013; Accepted 8 September 2013

Academic Editor: J. A. Tenreiro Machado

Copyright © 2013 Jian-Qin Liao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Mathematical analysis and modelling is central to infectious disease epidemiology. This paper, inspired by the natural ripple-spreading phenomenon, proposes a novel ripple-spreading network model for the study of infectious disease transmission. The new epidemic model naturally has good potential for capturing many spatial and temporal features observed in the outbreak of plagues. In particular, using a stochastic ripple-spreading process simulates the effect of random contacts and movements of individuals on the probability of infection well, which is usually a challenging issue in epidemic modeling. Some ripple-spreading related parameters such as threshold and amplifying factor of nodes are ideal to describe the importance of individuals' physical fitness and immunity. The new model is rich in parameters to incorporate many real factors such as public health service and policies, and it is highly flexible to modifications. A genetic algorithm is used to tune the parameters of the model by referring to historic data of an epidemic. The well-tuned model can then be used for analyzing and forecasting purposes. The effectiveness of the proposed method is illustrated by simulation results.

1. Introduction

Mathematical representation and analysis of infectious diseases has been central to infectious disease epidemiology since its inception as a discipline more than a century ago [1–7]. In recent years, detailed electronic surveillance of infectious diseases has become widespread owing to the advent of improved computing, electronic data management, the ability to share and deposit data over the internet, and rapid diagnostic tests and genetic sequence analysis. These ongoing developments have increased the application of mathematical models to both the generation and testing of basic scientific hypotheses and to the design of practical strategies for disease control. Mathematical analyses and models have provided successful explanations of previously puzzling observations and played a central part in public health strategies in many countries [3, 4].

Fundamental to the growing importance of mathematical epidemiology has been the integration of mathematical models with rigorous statistical methods to estimate

key parameters of these models and test hypotheses using available data. In the absence of reliable data, mathematics can be used to help formulate hypotheses, inform data-collection strategies, and determine sample sizes, which can permit discrimination of competing hypotheses. In this way, mathematics is “no more, but no less, than a way of thinking clearly about the problem in hand” [5]. The extent and quality of available data can be variable. Ideally, data should be analysed using models that adequately describe the observed dynamics and patterns of interest and the mechanisms that generate these observations. Models should be as simple as possible, but not so simple that the conclusions drawn are altered by the consideration of additional realistic complexity. Unnecessary complexity can obscure fundamental results and is almost as undesirable as oversimplification. Indeed, model choice—the process of deciding which model complexities are necessary—is a central part of mathematical modelling of infectious diseases.

The most recent survey on epidemic modelling can be found in [7]. Basically, existing epidemic models can be

classified into two categories [8–20]: (i) top-down models which are deterministic and based on systems of differential equations [20], Markov Chain [8], mean field type equations [9], and (ii) bottom-up models which are stochastic and based on computer simulations [10], agent based methods [11], cellular automata [12], and network theory [13]. For top-down models, diffusive or perfect mixing and random motion are assumptions that are not always fulfilled, at least in the human population. These models also tend to incorporate many parameters to explain reality, which increase their complexity and make them computationally intensive and difficult to analyze. Contrary to what happens to top-down models, the complex systems approach is the foundation of bottom-up models, where it is considered that spatial extended systems are capable of nontrivial collective behaviour—unexpected behaviour which is observed in macroscopic quantities. It is assumed that there are several levels of reality: at a microscopic level, interactions may be described by complicated potentials, but, at a macroscopic level, the properties of the system are dominated by the aggregated effect of all microscopic interactions. Human epidemics are strongly related to the dynamics of populations and to the network of social contacts. In particular, network theory has proven a promising new method applicable to epidemiology. For instance, the influence of small-world and scale-free topologies on the breakout of plagues was investigated in [13, 15]; a random network was used to conduct multiscale analysis on epidemic dynamics [16]; a growing network model was reported to develop a population-level epidemic model [17]. As stated in [18], the combination of network theory and epidemic modelling can deliver an improved understanding of disease dynamics and better public health through effective disease control. However, as pointed out in a recent perspective paper [19], new theories and methods are still needed to study interacting dynamics, amplification, and cascading effects in complex network systems such as epidemic dynamics.

Inspired by the natural ripple-spreading phenomenon on liquid surfaces, this paper reports a novel complex system based bottom-up epidemic model: ripple-spreading epidemic model (RSEM), which is an application-focused extension of our recent work on general ripple-spreading network models [21, 22]. As widely acknowledged, random contacts and movements of individuals impose a big challenge to epidemic modelling. Defining neighbourhoods and/or introducing transport rules at a microscopic level are often measures adopted to simulate social contacts and physical movements of individuals. In contrast, the new model proposed here takes account of the effect of such interactions between individuals via a ripple-spreading process and the reaction of nodes to ripples, whilst all nodes can be fixed without the need for a predefined neighbourhood. Basically, the infection probability is reflected by the point energy of ripples, and the social activeness of individuals can be associated with the threshold and the amplifying factor of nodes. Actually, the proposed ripple-spreading model can intuitively capture many spatial and temporal factors which matter in the outbreak of plagues. Therefore, this new model, when in combination with an effective parameter tuning method such

as genetic algorithms (GAs), possesses excellent potential for studying epidemic dynamics.

The remainder of this paper is organized as follows. Firstly, a general ripple-spreading network model of epidemic is proposed in Section 2. Then a genetic algorithm based method is reported in Section 3 to tune the model, so that it can simulate a specific epidemic. Some simulation results are illustrated in Section 4, and the paper ends with its main conclusions in Section 5.

2. Ripple-Spread Epidemic Model (RSEM)

2.1. The Basic Idea of Ripple-Spread Network Model. The basic natural ripple-spreading phenomenon is as follows. Suppose a collection of stakes is randomly distributed in a quiet pond, and suddenly a stone is thrown into the pond generating an initial ripple from the point where the stone hits the quiet water surface. When the ripple reaches a near stake, a new ripple is generated around the stake due to the reflection effect. Hereafter, for the sake of consistency, we denote such a new ripple as a responding ripple or outgoing ripple and the ripple which triggers the responding ripple as a stimulating ripple or incoming ripple. As the initial stimulating ripple is spreading, more and more responding ripples are stimulated around stakes. However, since the point energy on the initial stimulating ripple decays as it spreads out, those responding ripples triggered at a late phase will hardly be noticed. Let a node in a network stand for a stake in the pond, and an edge will be established between two nodes if a stake's ripple triggers a new ripple around the other stake. Then, after all ripples decay, we will get a network according to which stake's ripple has caused which stake to generate a new ripple. This is the basic idea of the ripple-spreading network model. Figure 1 gives an illustration of the development of a ripple-spreading network. For more details, readers are referred to [21, 22].

We can liken the outbreak of plagues to the above natural ripple-spreading phenomenon. Replace the stakes in the pond with a population of susceptible individuals in a community where an epidemic may break out. An initial infective case in the community is likened to the stone which hits the water surface. The influence of this initial infective case on other individuals is analogous to the initial stimulating ripple. The probability of infection can be related to the point energy of the ripple (but as will be explained later, the point energy is not exactly the probability of infection). The point energy decays as the ripple spreads out, and this can effectively capture the fact that long distance and few contacts often mean a low probability of infection. The process that a stimulating ripple triggers a responding ripple is related to a susceptible individual being infected and then becoming infective. Stakes of different textures may have different reflection effects; for instance, a rigid texture will cause strong refection whilst a soft texture absorbs the most energy of the incoming ripple and therefore has no reflection. This is likened to the difference in the physical fitness, immunity, and social activities of individuals. Therefore, we can set a threshold for each individual. If and only if the point energy of an incoming ripple is above the threshold,

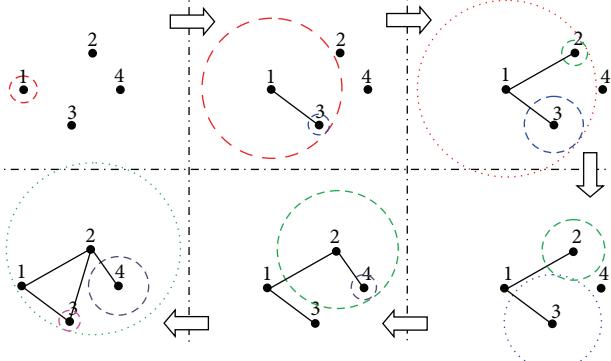


FIGURE 1: An illustration of the development of a ripple-spreading network.

it can be possible for an individual to be infected. Once an individual is infected, how infective s/he will become largely depends on his/her social activities. We can set an amplifying factor for each individual, which will determine the initial point energy of the responding ripple based on the point energy of the stimulating ripple. Therefore, the responding ripple of a socially active individual will have high point energy, which means a high probability of infecting other individuals. A stake can generate new reflection every time when it is reached by an incoming ripple that has enough point energy. This corresponds to the fact that an individual can be infected again when s/he becomes susceptible once again following a period of postrecovery immunity. Which stake's ripple causes which subsequent stake to generate a new ripple is analogous to who infects whom, indicated by an established edge between two nodes in a network. Who has infected whom during an epidemic outbreak can be illustrated by a growing network that is simulated by which stake's ripple has triggered which following stake to generate a new ripple. Now, one may get a feeling that the ripple-spreading network model invented in [21, 22] can be used to simulate the outbreak of plagues. To this end, we first need to, based on the work reported in [21, 22], develop a mathematical ripple-spreading network model of epidemic, which hereafter is called ripple-spreading epidemic model, denoted as RSEM.

2.2. Mathematical Formulation of RSEM. In the proposed RSEM, there are two groups of parameters. The first group is those of the general ripple-spreading network model as reported in [21, 22], but some modifications may be necessary in order to fit them in the scope of epidemiology. In this group, first we have parameters for N_{EISR} epicenters of initial stimulating ripples (EISRs), which are related to those initial cases in an epidemic outbreak. In this study, we only focus on the infectious disease transmission between human hosts, and other hosts such as rats and mosquitoes are not considered. Therefore, each EISR is actually a node in the network, that is, an individual in the community. The i th EISR, $i = 1, 2, \dots, N_{EISR}$, has an initial point energy of $E_{EISR}(i)$, and it is not active, that is, not infective, until time

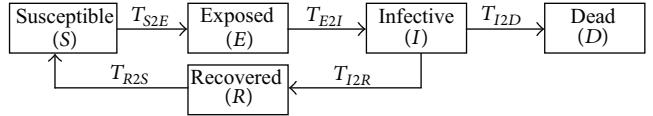


FIGURE 2: The cycle of epidemic states.

instant $T_{EISR}(i)$. Suppose there are N_N nodes in the network. Then, for node i , there is a threshold $\beta(i)$ to determine whether it is possible for the node to be infected by a certain infective node and an amplifying factor $\alpha(i)$ to calculate the initial energy when the node is infected. Basically, we can assume all ripples have the same spreading speed s and the same energy decaying coefficient vector η .

The second parameter group is epidemic specific. Because the nodes of network represent individuals, every node will have a certain epidemic state at each time point. Let $S_N(i, t)$ denote the epidemic state of node i at time t . Table 1 lists the epidemic states used in this study. Figure 2 shows how an individual will go through these epidemic states in an epidemic cycle. Simply speaking, in the epidemic cycle, an individual is initially susceptible. Once s/he is infected, s/he becomes exposed and then s/he develops to be infective; an infected individual may either die or recover from the infection. A recovered individual usually gets postrecovery immunity. If this is temporary, then after a period of time of immunity, s/he will become susceptible again. As one can see from Figure 2, there is a time period for an individual to transfer from one epidemic state to its followingup epidemic state. Except T_{S2E} , all other state transfer times, that is, $T_{E2I}, T_{I2D}, T_{I2R}$, and T_{R2S} , are independent of the ripple-spreading process, being instead mainly determined by the nature of the epidemic, the physical fitness of individuals, and/or relevant public health policies/measures. Basically, from statistical study of historical data, we may get an estimate of each state transfer time distribution (excluding T_{S2E}). In this study, we assume they all have Poisson distributions. For instance, T_{E2I} may have a Poisson distribution with a mean of \bar{T}_{E2I} . Therefore, with \bar{T}_{E2I} as a parameter, we can roughly know how long it will take for an individual to transfer from state "E" to state "I". When running the RSEM, once a node becomes exposed, we then randomly assign a T_{E2I} to the node according to a Poisson distribution with \bar{T}_{E2I} as the mean. Similarly, we deal with other state transfer times except T_{S2E} . Therefore, in the second group, we have four means as parameters, namely, $\bar{T}_{E2I}, \bar{T}_{I2D}, \bar{T}_{I2R}$, and \bar{T}_{R2S} . For the sake of generality, $T_{E2I}, T_{I2D}, T_{I2R}$, and T_{R2S} are allowed to be zero, which means one or more epidemic states may not be experienced by an individual in an epidemic cycle. For example, if $T_{E2I} = 0$ and $T_{R2S} = 0$, then there is no latency period or postrecovery immunity.

Besides the two groups of parameters, there are three kinds of dynamics in the RSEM: the ripple-spreading process, the reaction of nodes to ripples, and the state transfer of nodes. The first two originate from the general work reported in [21, 22], but some minor epidemic-specific modifications

are introduced. The dynamics of state transfer is a brand-new concept for the general ripple-spreading network model.

The ripple-spreading process in the RSEM is mathematically described as follows. Suppose ripple i is associated with the i th node, $i = 1, 2, \dots, N_N$. Then, let $E_N(i)$ be the initial point energy of the ripple i , $e_N(i, t)$ the point energy of the ripple i at time t , and $r_N(i, t)$ the radius of ripple i . Initialize $E_N(i) = 0$, $e_N(i, 0) = 0$ and $r_N(i, 0) = 0$. Once ripple i is triggered at a certain time point, $E_N(i)$ will be given a value larger than zero. If $E_N(i) > 0$ and $S_N(i, t) = "I"$ (i.e., the state of node i is "Infective"), then ripple i will spread out as follows:

$$r_N(i, t) = r_N(i, t - 1) + s, \quad (1)$$

$$e_N(i, t) = f_{\text{Decay}}(E_N(i), r_N(i, t)), \quad (2)$$

where f_{Decay} is a function defining how the point energy decays as the ripple spreads out. A typical decaying function may be

$$f_{\text{Decay}}(E, r) = \frac{\eta(k) E}{2\pi r}, \quad (3)$$

and k is an integer index calculated as

$$k = \frac{r_N(i, t)}{s}. \quad (4)$$

Once the $S_N(i, t)$ changes from "I" to "D" or "R", reset $E_N(i) = 0$, $e_N(i, 0) = 0$, and $r_N(i, 0) = 0$.

The reaction of nodes to ripples defines how a node is infected. Suppose node i is reached by ripple j at time t ; that is,

$$r_N(j, t) \geq D_N(i, j) \geq r_N(j, t - 1), \quad (5)$$

where $D_N(i, j)$ is the distance between node i and node j . If $S_N(i, t) = "S"$ and $e_N(j, t) \geq \beta(i)$, that is, the point energy of ripple j is above the threshold of node i , then node i will be infected with a probability which may be defined based on an arctangent function of $e_N(j, t)$ as follows:

$$p_R(i) = \frac{\tan^{-1}(\kappa((e_N(j, t) - \beta(i))/\beta(i)) - \delta) + \tan^{-1}(\delta)}{\pi/2 + \tan^{-1}(\delta)}, \quad (6)$$

where $\kappa > 0$ and δ are coefficients which can adjust the shape and location of the arctangent function. Once node i is infected, we establish a directional link from node j and node i by modifying the adjacency matrix:

$$M_A(j, i) = M_A(j, i) + 1, \quad (7)$$

set $S_N(i, t + 1) = "E"$, set the initial point energy of ripple i as

$$E_N(i) = \alpha(i) e_N(j, t), \quad (8)$$

record the time of becoming exposed $t_E(i) = t$, and assign values to $T_{E2I}(i)$, $T_{I2D}(i)$, $T_{I2R}(i)$, and $T_{R2S}(i)$ according to relevant Poisson distributions defined by \bar{T}_{E2I} , \bar{T}_{I2D} , \bar{T}_{I2R} , and \bar{T}_{R2S} .

Please note that, although ripple i has $E_N(i) > 0$ at time $t_E(i)$, it will not start spreading until the state of node i has become "Infective", which is determined by the following state transfer dynamics: if $S_N(i, t) = "E"$ and $t - t_E(i) \geq T_{E2I}(i)$, then set $S_N(i, t + 1) = "I"$ and start the ripple-spreading process for ripple i ; if $S_N(i, t) = "I"$ and $t - t_E(i) \geq T_{E2I}(i) + T_{I2D}(i)$, randomly decide whether node i will die according to a preset death rate R_D ; if node i has died, set $S_N(i, t + 1) = "D"$, $r_N(i, t) = 0$, and stop the ripple-spreading process of ripple i ; if $S_N(i, t) = "I"$ and $t - t_E(i) \geq T_{E2I}(i) + T_{I2R}(i)$, set $S_N(i, t + 1) = "R"$, $r_N(i, t) = 0$, and $E_N(i) = 0$, and stop the ripple-spreading process of ripple i ; if $S_N(i, t) = "R"$ and $t - t_E(i) \geq T_{E2I}(i) + T_{I2R}(i) + T_{R2S}(i)$, set $S_N(i, t + 1) = "S"$. So, node i has gone through an epidemic cycle.

By integrating the above parameters and dynamics, the proposed RSEM can be finally described as a whole by the following steps.

Step 1. Initialization, that is, set $S_N(i, 0) = "S"$, $E_N(i) = 0$, $r_N(i, 0) = 0$, $t = 0$ and $M_A(j, i) = 0$ for all $i = 1, 2, \dots, N_N$ and $j = 1, 2, \dots, N_N$, randomly choose N_{EISR} nodes as initial cases, set their state as "E", set their T_{E2I} as T_{EISR} , and randomly set their T_{I2D} , T_{I2R} , and T_{R2S} according to the relevant Poisson distributions defined by \bar{T}_{I2D} , \bar{T}_{I2R} , and \bar{T}_{R2S} .

Step 2. While the termination criteria are not satisfied, let $t = t + 1$, do for $i = 1, 2, \dots, N_N$.

Substep 2.1. Let $S_N(i, t) = S_N(i, t - 1)$.

Substep 2.2. If $E_N(i) > 0$ and $S_N(i, t) = "I"$, let ripple i spread for one time step by following the ripple-spreading process described by (1) to (4).

Substep 2.3. If node i has $S_N(i, t) = "S"$ and is reached ripple j according to (5), calculate the reaction of node i as described by (6) to (8).

Substep 2.4. If $S_N(i, t) \neq "S"$, following the state transfer dynamics to process node i for one time step.

When an RSEM run is terminated, a network will appear based on all established directional links which indicate how the infectious disease has transmitted between individuals. Figure 3 illustrates how the infectious disease transmission can be simulated by the proposed RSEM.

2.3. Further Analysis of RSEM. As discussed in the Introduction section, social contacts and physical movements of individuals impose a big challenge to epidemic modeling. Fortunately, thanks to ripple-spreading dynamics, the proposed RSEM can effectively simulate the effect of social contacts and physical movements without actually applying such activities to each individual. Basically, the strength of social contacts and physical movements of an individual is largely reflected by the amplifying factor $\alpha(i)$ and the threshold $\beta(i)$. For a socially active individual, s/he is more

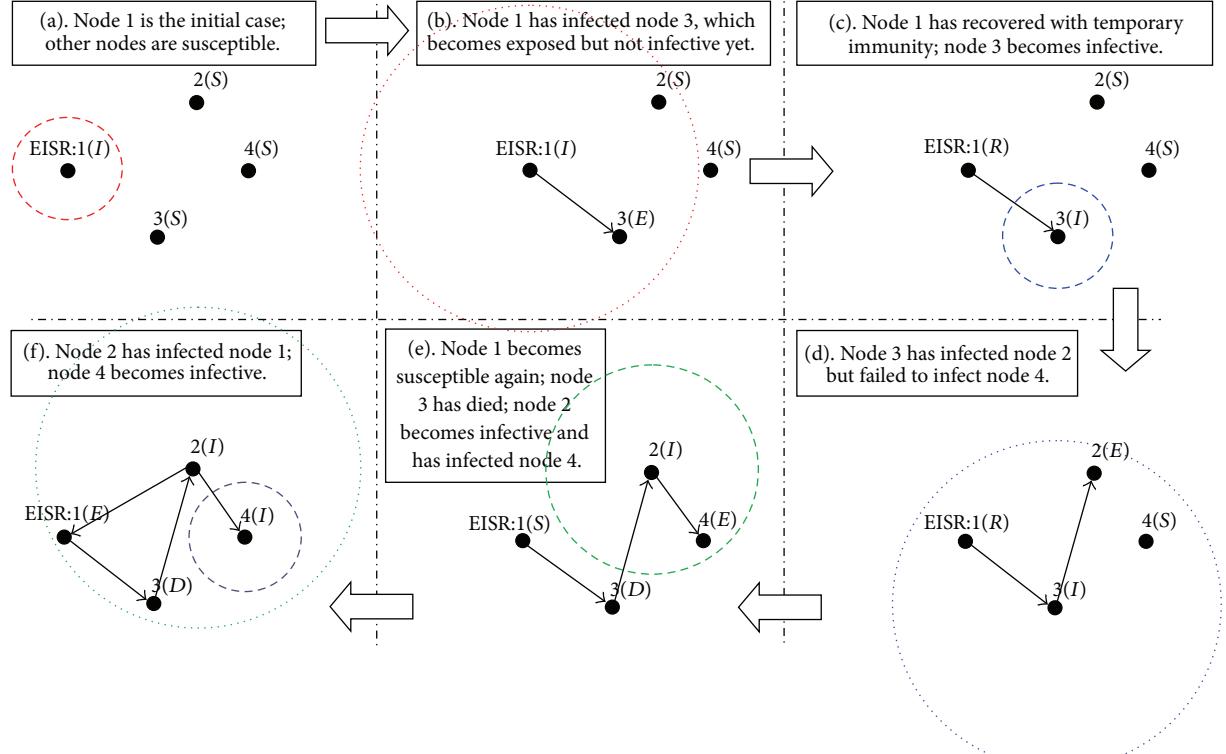


FIGURE 3: An illustration of RSEM.

likely to be infected, so s/he should have a smaller $\beta(i)$. According to (6), one can see that a smaller $\beta(i)$ means a higher probability of being infected. At the same time we set a larger $\alpha(i)$ for a socially active individual, and, from (8), one can see that his/her ripple will then have more point energy to infect others, which is in line with the reality. According to (2) and (3), as a ripple spreads out, its point energy decays, which in general means that the infection probability becomes lower as the distance between two individuals increases. This is also the case in reality: for an individual, usually most of his/her social contacts and physical movements happen within a small group of people (compared with the whole community) and a limited physical space (compared with the entire space). Therefore, most secondary infections caused by him/her will be spatially limited. Now it is clear that, by associating the point energy of ripples to the probability of infection and the amplifying factor and the threshold of nodes to the social contacts and physical movements of individuals, the proposed RSEM provides a new approach to simulate and study the epidemic dynamic.

One may argue that, according to Table 1 and Figure 2, the proposed RSEM misses the state of “Admitted”, which usually means that individuals who are infective are quarantined in hospital and therefore have no chance to transmit the infection to others. Obviously, this state is not an original state in the natural epidemic cycle, but a man-made state due to the public health service. Actually, the effect of the state of “Admitted” can be equivalent to shortening \bar{T}_{I2D} and \bar{T}_{I2R} , and increasing \bar{T}_{R2S} . Therefore, in this study, we do

not consider the state of “Admitted” explicitly. This will be further discussed later in the simulation section.

It should be noted that, in the ripple-spreading process, (3) uses an energy decaying coefficient vector rather than a single coefficient scalar. This is because the infectiousness over time after infection is not a constant, as shown by the examples in Figure 4. Therefore, we can use a piecewise function defined by the vector η to describe the time-varying infectiousness of a given epidemic.

In the reaction of node to ripples, according to (6), the curve of the probability $p_R(i)$ for node i to be infected has a shape of an arctangent function like the solid line in Figure 5. Basically, once $e_N(j, t) \geq \beta(i)$, the larger the point energy $e_N(j, t)$ is, the higher the probability for node i to be infected by node j . The probability $p_R(i)$ does not go up linearly as $e_N(j, t)$ increases. Actually, $p_R(i)$ increases very slowly before $e_N(j, t)$ reaches a certain critical value, around which $p_R(i)$ goes up sharply and then gets almost saturated no matter how large $e_N(j, t)$ is. This is reasonable, because, according to (3), the value of $e_N(j, t)$ largely reflects the distance (spatial, social, or both) between two nodes, and in reality infection occurs mainly within a certain range of spatial and/or social distance, whilst beyond that range the probability of infection is quite low. The coefficient κ determines how sharply the probability curve changes around the critical value of $e_N(j, t)$: the larger the value of κ , the sharper the probability curve. The coefficient δ determines how likely it is that an $e_N(j, t)$ will cause a large probability of infection: the smaller the value of δ , the more likely a large probability. To simplify

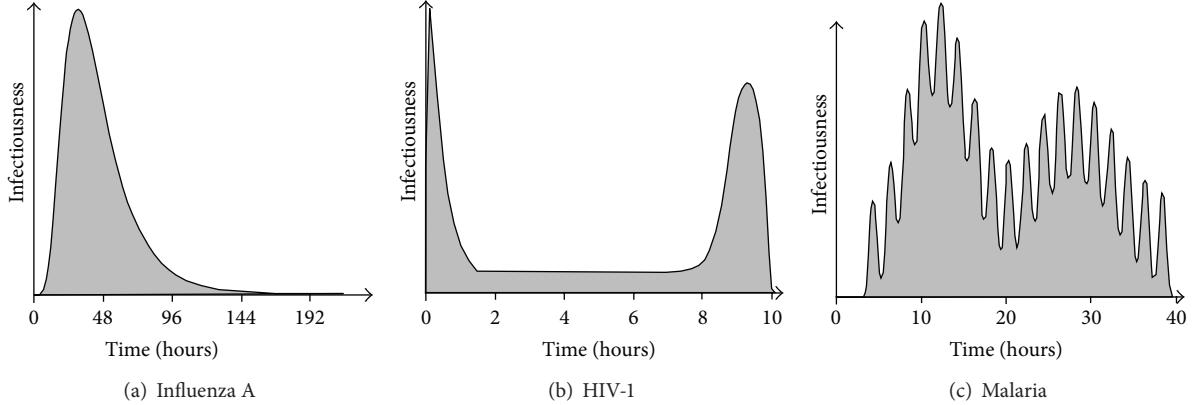


FIGURE 4: Biological infectiousness over time after infection for three different human pathogens [14].

TABLE 1: States used in epidemic model.

Symbol	Epidemic state	Description
S	Susceptible	Healthy Individuals who could potentially develop the disease.
E	Exposed	Individuals who have been infected with the disease, but who are still in the latent period (with or without symptoms of the disease) and who cannot transmit the disease to others.
I	Infective	Individuals who are infected with the disease (with or without symptoms of the disease) and who are capable of transmitting the infection to others.
R	Recovered	Individuals who have recovered from infection thereby acquiring immunity (temporary or permanent) from infection.
D	Dead	Individuals who have died from infection

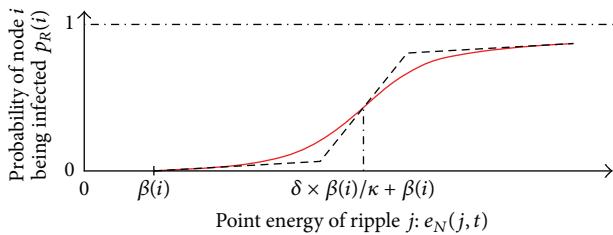


FIGURE 5: Probability of infection and point energy of ripple.

the RSEM, one may sometimes use a piecewise probability function, such as the dotted line in Figure 5, to approximate the arctangent function.

From the description above, it is apparent that the proposed RSEM is more complicated than many existing

epidemic models. This also means that it is more flexible to modifications. For instance, the space where nodes distribute and ripples spread out is not necessarily the real geographic space. Instead, it can be an artificial multidimensional space which includes social relationships, shopping patterns, and habits. The nodes may not always be fixed; in other words, by referring to relevant statistic studies of long-distance travelling patterns, some randomly chosen nodes may occasionally jump in the space. The proposed RSEM can easily incorporate such modifications, but this is beyond the scope of this paper.

3. Genetic Algorithm for Tuning Model Parameters

Like all other models, the proposed RSEM is supposed to be able to (i) simulate historic epidemic outbreaks, (ii) analyze epidemic dynamics/mechanism and health policies/measures, and (iii) make forecasts to some extent. Simulating historic epidemic outbreaks is mainly used for model verification building confidence so that the model will be employed for analysis and forecasting. In the simulation of historic epidemic outbreaks, the output of a model should match the relevant historic data as closely as possible. Basically, this can be assessed by comparing the number of daily infections $N_{DI}(t)$, the number of daily deaths $N_{DD}(t)$, the number of daily recoveries $N_{DR}(t)$, and the basic reproduction ratio R_0 , which is defined as the average number of secondary infections produced when one infected individual is introduced into a population where everyone else is susceptible [14]. In short, the quantity R_0 governs whether an infection can spread and be sustained within a population. If R_0 is greater than one, then the number of infections in a susceptible population will increase and the infection will be sustained, whereas if R_0 is less than one, the infection should fail to take hold and will die out very quickly.

For a simple deterministic model with a few parameters such as those based differential equations, it is possible to tune the parameters by hand based on experience, in order to match relevant historic data. However, for a complicated stochastic model with many parameters, which is the case for the proposed RSEM, hand tuning is very difficult, if

not impossible. Actually, the output of the RSEM does not simply rely on the values of parameters but is more determined by whether such values together satisfy certain conditions, known or unknown. For instance, [21] derived some conditions which involve all parameters to determine whether all nodes will be connected together in the general ripple-spreading network model. Such complicated relationships between parameters reflect the fact that in reality it is usual for many factors to work together in a complex, intermixed fashion. Tuning the parameters for a complicated model/system is often far beyond the capability of even human experts. Fortunately, we can turn to methods such as genetic algorithms (GAs) for help. A GA is a large-scale parallel stochastic searching and optimizing method inspired by the biological mechanisms of evolution and heredity. In recent years, GAs have been widely used for resolving various complex problems, particularly those including parameter optimization [23–26]. The basic idea of a GA is that, given a population of chromosomes, the environmental pressure causes natural selection (survival of the fittest), and hereby the fitness of the population grows. It is easy to see such a process as optimization. Given an objective function to be maximized, we can randomly create a set of candidate solutions (chromosomes) and use the objective function as an abstract fitness measure (the higher the better). Based on this fitness, some of the better chromosomes are chosen to seed the next generation by applying crossover and/or mutation. Crossover is applied to two selected chromosomes, the so-called parents, and results in one or two new chromosomes, the children. Mutation is applied to one chromosome and results in one new chromosome. Applying crossover and mutation leads to a set of new chromosomes, the offspring. Based on their fitness, these offsprings compete with the old chromosomes for a place in the next generation (in some GA implementations, the population is all replaced by the offspring). This process can be iterated until a solution is found or a previously set time limit is reached. Many components of such an evolutionary process are stochastic. According to Darwin, the emergence of new species, adapted to their environment, is a consequence of the interaction between the survival of the fittest mechanism and undirected variations. Variation operators must be stochastic, the choice of which pieces of information will be exchanged during crossover, as well as the changes in a chromosome during mutation, is random. On the other hand, selection operators can be either deterministic or stochastic. In the latter case fitter chromosomes have a higher chance of being selected than less fit ones, but typically even the weak chromosomes have a chance to become a parent or to survive. For more theoretical details of GAs, readers may refer to [23–26].

To apply a GA in this study, firstly, we need to decide, in the RSEM, which parameters can be set up by hand and which ones need to be tuned by the GA. Basically, parameters such as the number of initial cases N_{EISR} , the means of state transfer times \bar{T}_{E2I} , \bar{T}_{I2D} , \bar{T}_{I2R} , and \bar{T}_{R2S} , energy decaying coefficient vector η , and death rate R_D can be easily set up according to relevant historic data and statistic studies. Coefficients κ and δ can also be tuned by hand mainly to

change the overall shape of the infection probability curve as shown in Figure 5. Most ripple-spreading related parameters are purpose-designed and have no real-world meanings or references, and therefore they should be tuned by the GA. These parameters specifically include the initial point energy $E_{EISR}(i)$ of initial cases, the ripple spreading speed s , the threshold $\beta(i)$, and the amplifying factor $\alpha(i)$. It should be noted that we do not need to tune $\beta(i)$ and $\alpha(i)$ for every node, but we just need to tune their means, that is, $\bar{\beta}$ and $\bar{\alpha}$, and then generate $\beta(i)$ and $\alpha(i)$ according to a Poisson distribution with $\bar{\beta}$ and $\bar{\alpha}$ as the means, respectively.

Then, we construct a fitness function based on $N_{DI}(t)$, $N_{DD}(t)$, $N_{DR}(t)$, and R_0 as follows:

$$\begin{aligned} J = & w_1 |\bar{R}_0 - R_0^*| + w_2 \sum |\bar{N}_{DI}(t) - N_{DI}^*(t)| \\ & + w_3 \sum |\bar{N}_{DD}(t) - N_{DD}^*(t)| \\ & + w_4 \sum |\bar{N}_{DR}(t) - N_{DR}^*(t)|, \end{aligned} \quad (9)$$

where w_1 to w_4 are weights, the variables with “*” are the approximated values of relevant historical data, and the over barred variables are the relevant mean results of a number of random RSEM runs with a given set of parameter values. Usually, the curve of historical data is not smooth and nor is the curve of a single RSEM run. If we directly compare the historical curve with the curve of any single RSEM run, the summed error will usually be very large, even though the two curves are actually similar to each other. Therefore, we need to use approximated values and average values, both of which are much smoother for accurate comparison. Another reason for utilizing the average results of many RSEM runs is that the proposed RSEM is stochastic in nature, which means that the model will deliver different results in different runs even for the same set of parameter values. Therefore, to better assess the effect of a given set of parameter values, a sufficient number of RSEM runs need to be conducted for a parameter value set.

Thus, we can use GA to tune the parameters for the RSEM as follows.

Step 1. Initialize a population of sets of parameter values.

Step 2. For a set of parameter values, conduct a number of RSEM runs, and then calculate the fitness based on the average results. Repeat the same for every set of parameter values in the population.

Step 3. If the termination criterion is satisfied, go to *Step 4*. Otherwise, perform selection, crossover, and mutation to generate a new population of sets of parameter values. Go to *Step 2*.

Step 4. Output the best set of parameter values, and stop.

Once the RSEM is well tuned by the GA, we can then use it to analyze epidemic details and policy effects, as well as forecasting new trend or future development of an epidemic. Some relevant technical details concerning GA design and implementation may be found in [22]. It should be noted,

however, that [22] focuses on applying a GA to optimize the basic RSNM of [21] so as to generate small-world and scale-free network topologies. As discussed in Section 2, the basic RSNM is somewhat different from the RSEM proposed in this paper. Nevertheless, the use of a GA to optimize model parameters in [22] is helpful to tune the RSEM in this paper.

4. Illustrative Simulation Results

As explained in Section 3, the first step in applying the proposed RSEM is to use some historical data to tune the parameter values in order to simulate the associated real epidemic outbreak. Here we consider the March 2003 Severe Acute Respiratory Syndrome (SARS) outbreak in Hong Kong. From the empirical data for Hong Kong, we know that a severe localized epidemic occurred in Amoy Gardens. Through contact tracing, it has been estimated that 332 cases were involved in this outbreak, a large proportion of which could be traced to a single person. Outbreaks involving the direct spread of the disease from one primary case to a very large number of secondary cases are often referred to as “Super-Spreading Events”, which are often ideally used to test epidemic models. As one can see from Figures 6, 7, and 8, the real historical data form no smooth curves. Because the SARS case in Amoy Gardens is a unique single case, we cannot get any smooth curve of average historical data, which makes it difficult to calculate the fitness of chromosome as defined in (9). Thus, to be able to apply a GA to tune the reported RSEM, we adopt the differential equations based model in [20] to generate an approximated smooth curve of average historical data. With 3 smooth curves (of daily infections, death, and recovery, resp.) generated by the model in [20], we can calculate the chromosome fitness according to (9) and then apply GA to tune the RSEM parameters. It should be noted that there are actually no daily infection data, and we hence use the daily admissions data in [20] to approximate them. In this study, the weights w_1 to w_4 in (9) are set to 100, 1, 10, and 2, respectively. These values for w_1 to w_4 are chosen largely by referring to the data of basic reproduction ratio, daily admissions, daily death, and daily recovery in [20], so that the four items in (9) will have roughly equal contributions to the fitness function. In the test, all SARS-specific disease related parameters are set up according to [20], and the number of nodes is $N_N = 2396$ (i.e., the susceptible population in Amoy Gardens). For the GA, the crossover probability is set to 0.6, the mutation probability is set to 0.1, the population size is 200, and the number of evolving generations is 500. For each chromosome in a population, we use the parameter values represented by it to run the RSEM for 100 times, and then use the average results to calculate its fitness according to (9). Finally, we use the fittest chromosome in the last generation of GA to set the RSEM, in order to simulate the SARS case in Amoy Gardens.

Figures 6, 7, and 8 plot the final simulation results of applying RSEM to study the SARS case in Amoy Gardens; the standard deviations (SDs) of the results of the model in [20] and the RSEM from the historical data are given in Table 2. From Figures 6, 7, and 8 and Table 2, one can see that (i) the average curves of RSEM are similar to those

of [20]; (ii) the SDs of the RSEM are close to those of the model in [20]; (iii) the results of a single RSEM run have patterns (peaks and valleys) fairly similar to those of historical data; (iv) the ranges and tendencies of RSEM data all match those of historical data and [20]. Therefore, Figures 6, 7, and 8 and Table 2 illustrate that the proposed RSEM and GA can deliver a fairly satisfactory performance in the case study of SARS and therefore can be useful to model epidemic dynamics. However, due to lack of relevant data on the physical fitness, immunity, and social activities of the population in Amoy Gardens, at the moment it is still difficult to explain or make sense of the values of RSEM parameters in this SARS case. Therefore, more cases studies based on different historical data and social survey data are still needed, which is beyond the scope of this paper and demands effort in future research.

Due to the early stage of this work, we have not collected sufficient historical data for parameter tuning purposes. However, we can still investigate the model based on synthetic data sets, in order to understand the properties of the model itself. Here, we can give some simulation results to see how the parameters can affect the output of the model. Now we set $N_N = 250$; that is, the population of original susceptible individuals is 250. We assume there are 8 initial cases randomly distributed in the population. Firstly, we suppose that the population is uniformly randomly distributed in a square area defined by coordinates $(-1000, -1000)$ and $(1000, 1000)$, and all initial cases have the same initial point energy $E_{EISR} = 10000$, the Poisson means for generating $T_{E2I}(i)$, $T_{I2D}(i)$, $T_{I2R}(i)$ and $T_{R2S}(i)$, $\beta(i)$ and $\alpha(i)$ are $\bar{T}_{E2I} = 10$, $\bar{T}_{I2D} = 15$, $\bar{T}_{I2R} = 15$ and $\bar{T}_{R2S} = 20$, $\bar{\beta} = 5$, and $\bar{\alpha} = 1000$, and coefficients $\eta = 1$, $\kappa = 1$, and $\delta = 10$. Then, in the simulation, we want to study the influence of three sets of parameters: (i) the square area (which is related to the density of population), (ii) the pair of $\bar{\beta}$ and $\bar{\alpha}$ (which are related to the overall social activeness of the population), and (iii) the triple of \bar{T}_{I2D} , \bar{T}_{I2R} , and \bar{T}_{R2S} (which are related to the general public health service standard). Each time we only change the value(s) of one set of parameters by 20% of the associated default values, while all the other parameters have the default values as given above. Please note that $\bar{\beta}$ and $\bar{\alpha}$ change in opposite directions ($\bar{\beta}$ increases whilst $\bar{\alpha}$ decreases) and \bar{T}_{R2S} and the pair of $(\bar{T}_{I2D}, \bar{T}_{I2R})$ change in opposite directions (\bar{T}_{R2S} increases whilst the other two decrease). For each given whole set of parameter values, we conduct 10 RSEM runs. Then, we only check the number of total infections, and the average results are given in Table 3, from which we have the following observations.

- (i) The default values for parameters result in a network where all nodes are connected, which means the whole population is eventually infected. This default case is used as a benchmark to assess the influence of some model parameters.
- (ii) The density of population plays an important role in the transmission of infectious diseases. A high density will contribute to the outbreak of plagues.

TABLE 2: Comparative results between RSEM and the model in [20].

Standard deviations	Number of infections		Number of death		Number of recovery	
	Daily	Cumulative	Daily	Cumulative	Daily	Cumulative
Model in [20]	12.7012	26.2656	1.5375	4.2721	7.9375	37.0366
RSEM	12.7217	27.2463	1.5573	3.5244	8.0686	40.9718

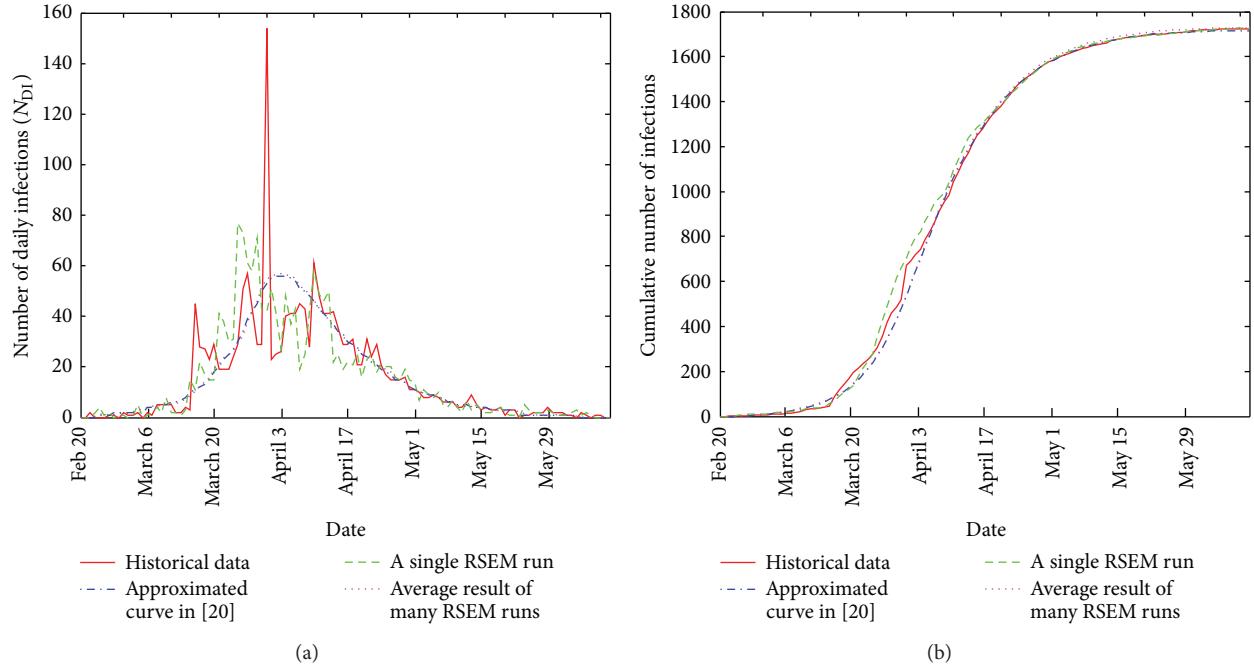


FIGURE 6: Simulation results for Amoy Gardens, Hong Kong: number of infections.

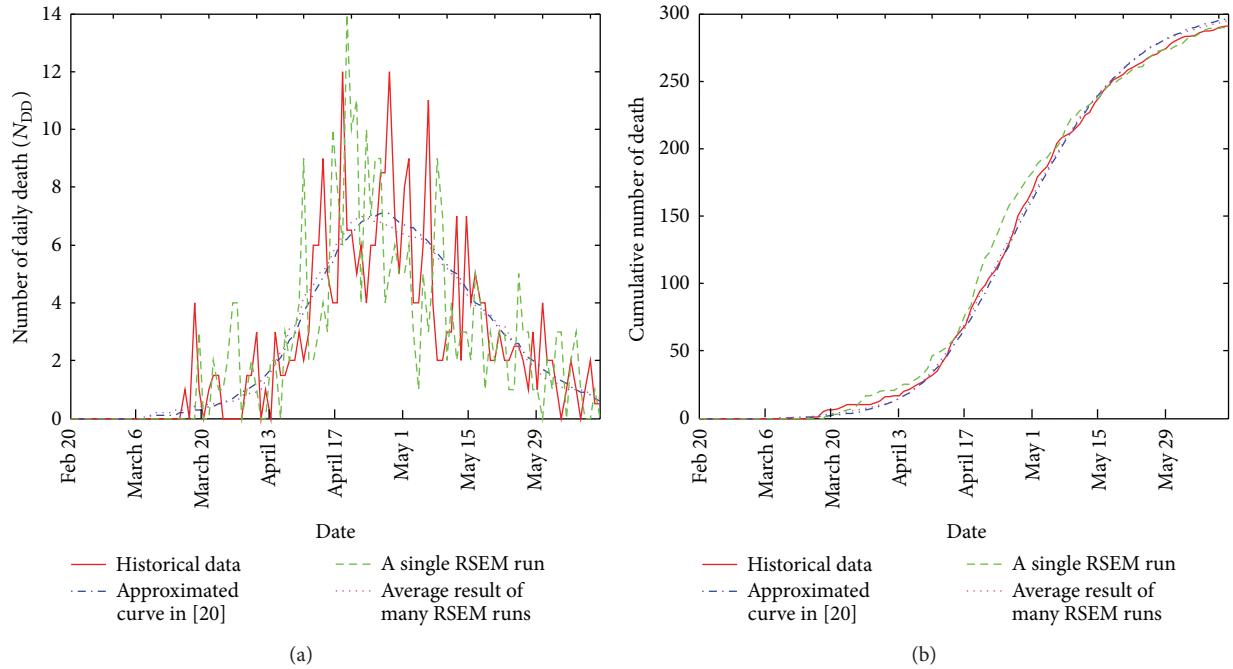


FIGURE 7: Simulation results for Amoy Gardens, Hong Kong: number of deaths.

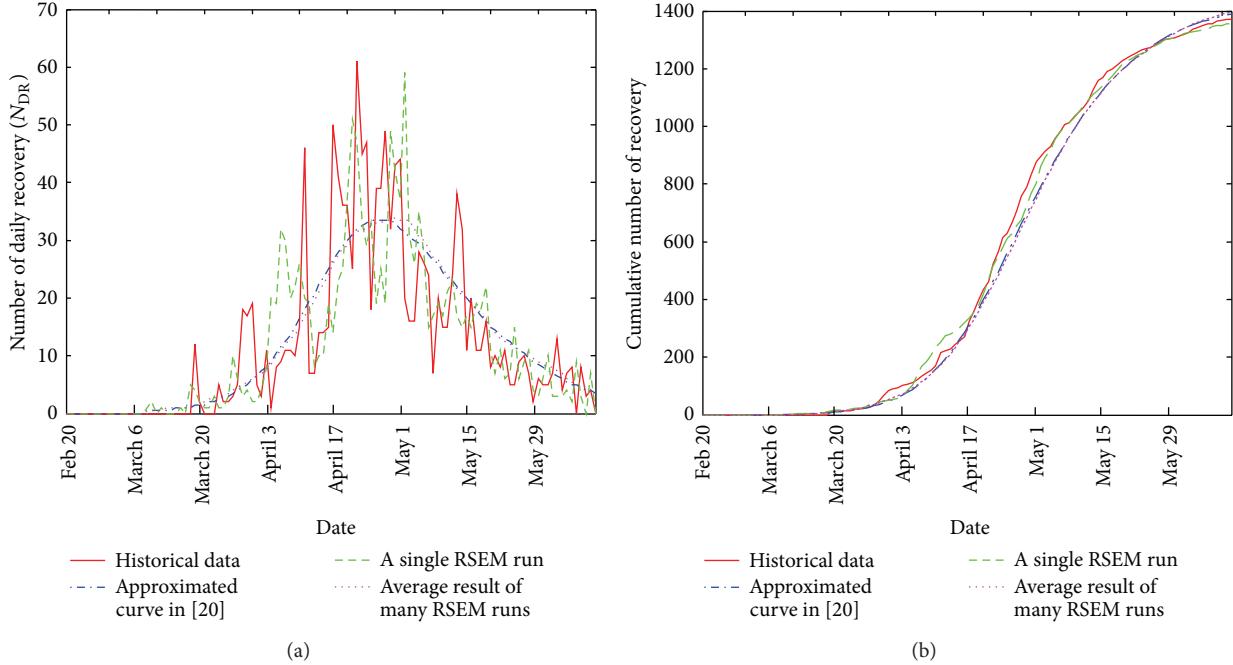


FIGURE 8: Simulation results for Amoy Gardens, Hong Kong: number of recovery.

- (iii) Large threshold and small amplifying factor have a negative influence on the transmission of infectious diseases. This means a community where the overall social activeness is low will have a better resistance to epidemic outbreaks. When an 80% change applies to the default values, usually only the initial cases will cause a few infections, and the epidemic will die out soon.
- (iv) Similarly, the general public health service standard plays a crucial role in the control of epidemic outbreaks. If infected individuals can be quarantined in a timely manner, say, if \bar{T}_{I2D} and \bar{T}_{I2R} are reduced by 80%, then infections are largely limited to the initial cases and the nearby.
- (v) The above simulation results clearly reveal the trends of the influence of different RSEM parameters. However, again, to fully understand the performance of RSEM, extensive efforts are still needed based on various real epidemics data in future research.

5. Conclusions and Future Work

This paper reports a novel ripple-spreading network model for the study of infectious disease transmission. We term the new model the ripple-spreading epidemic model (RSEM). By mimicking the ripple-spreading phenomenon on a calm liquid surface, the new model has many natural advantages to describe spatial and temporal factors in the outbreak of plagues. The effect of individuals' social activities, such as contacts and movements, which are often difficult for many existing models to simulate, can be effectively described by

TABLE 3: Influence of model parameters on total infections.

Change % default values	Square area	$(\bar{\beta}, \bar{\alpha})$	$(\bar{T}_{R2S}, \bar{T}_{I2D}, \bar{T}_{I2R})$
0%	250.0	250.0	250.0
20%	228.7	219.3	232.5
40%	205.7	185.8	191.6
60%	136.4	108.2	93.8
80%	51.9	16.5	21.9

the amplifying factor of nodes and the point energy of ripples, which largely determines the probability of infection. As the point energy of a ripple decays as the ripple spreads out, the probability of infection decreases. The threshold of nodes is another important parameter determining the dynamic process of infectious disease transmission, and basically this parameter is well associated with the physical fitness and immunity of individuals. The proposed model is highly flexible to modifications and is therefore capable of incorporating new real factors which matter in the study of epidemic. The general model can be tuned by genetic algorithm according to the historic data of a specific plague, and the well-tuned model can then be used to analyze and forecast the associated infectious disease. The effectiveness of the proposed model and method is illustrated by some simulation results.

Future work will include (i) carrying out more theoretical analysis of the proposed model, (ii) comparing with other models, and (iii) modifying and tuning the proposed model in order to study some real outbreaks of plagues. In particular,

as a new epidemic model, the RSEM needs relevant data on the physical fitness, immunity, and social activities of the population in a community/society, in order to better set up and understand the RSRPs (ripple-spreading related parameters). To get such data, a purpose-designed social survey needs to be conducted within a certain time window after a plague breaks out. Therefore, conducting relevant social survey for a recent plague event to collect sufficient data will be a crucial part of future work. Once such data are collected for a specific plague event, more comparisons with different relevant existing models will be made. Such application-oriented study needs to be conducted for different diseases, in order to explore the full potential of the proposed RSEM.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported in part by the China “973” Project under Grant 2012CB955404, the project Grant 2012-RC-02 from Beijing Normal University, China, and the Seventh Framework Programme (FP7) of the European Union under Grant PIOF-GA-2011-299725. Some early results of this paper were presented at the 2012 the 5th International Conference on BioMedical Engineering and Informatics (BMEI 2012), October 16–18, 2012, Chongqing, China.

References

- [1] H. Heesterbeek, “The law of mass-action in epidemiology: a historical perspective,” in *Ecological Paradigms Lost: Routes of Theory Change*, K. Cuddington and B. Beisner, Eds., pp. 81–105, Elsevier, Burlington, Mass, USA, 2005.
- [2] K. Dietz and J. A. P. Heesterbeek, “Daniel Bernoulli’s epidemiological model revisited,” *Mathematical Biosciences*, vol. 180, pp. 1–21, 2002.
- [3] R. M. Anderson and R. M. May, *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, 1991.
- [4] J. Glasser, M. Meltzer, and B. Levin, “Mathematical modeling and public policy: responding to health crises,” *Emerging Infectious Diseases*, vol. 10, no. 11, pp. 2050–2051, 2004.
- [5] R. M. May, “Uses and Abuses of Mathematics in Biology,” *Science*, vol. 303, no. 5659, pp. 790–793, 2004.
- [6] N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases and Its Applications*, Hafner Press, New York, NY, USA, 2nd edition, 1975.
- [7] C. I. Siettos and L. Russo, “Mathematical modeling of infectious disease dynamics,” *Virulence*, vol. 4, no. 4, pp. 295–306, 2013.
- [8] P. D. O’Neill, “A tutorial introduction to Bayesian inference for stochastic epidemic models using Markov chain Monte Carlo methods,” *Mathematical Biosciences*, vol. 180, pp. 103–114, 2002.
- [9] A. Kleczkowski and B. T. Grenfell, “Mean-field-type equations for spread of epidemics: the “small world” model,” *Physica A*, vol. 274, no. 1, pp. 355–360, 1999.
- [10] “Computer simulation model provides tactics to deal with flu epidemic,” <http://www.healthjockey.com/2009/05/04/computer-simulation-model-provides-tactics-to-deal-with-flu-epidemic/>.
- [11] R. Connell, P. Dawson, and A. Skvortsov, “Comparison of an agent-based model of disease propagation with the generalised SIR epidemic model,” Tech. Rep. DSTO-TR-2342, 2009.
- [12] M. Toffoli, *Cellular Automata Machines*, Scientific Computation, MIT Press, Cambridge, Mass, USA, 1991.
- [13] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, “Complex networks: structure and dynamics,” *Physics Reports*, vol. 424, no. 4–5, pp. 175–308, 2006.
- [14] N. C. Grassly and C. Fraser, “Mathematical models of infectious disease transmission,” *Nature Reviews Microbiology*, vol. 6, no. 6, pp. 477–487, 2008.
- [15] R. Pastor-Satorras and A. Vespignani, “Epidemic spreading in scale-free networks,” *Physical Review Letters*, vol. 86, no. 14, pp. 3200–3203, 2001.
- [16] A. I. Reppas, K. G. Spiliotis, and C. I. Siettos, “Epidemionics: from the host-host interactions to the systematic analysis of the emergent macroscopic dynamics of epidemic networks,” *Virulence*, vol. 1, no. 4, pp. 338–349, 2010.
- [17] R. Breban, R. Vardavas, and S. Blower, “Linking population-level models with growing networks: a class of epidemic models,” *Physical Review E*, vol. 72, no. 4, part 2, Article ID 046110, p. 8, 2005.
- [18] M. J. Keeling and K. T. D. Eames, “Networks and epidemic models,” *Journal of the Royal Society Interface*, vol. 2, no. 4, pp. 295–307, 2005.
- [19] D. Helbing, “Globally networked risks and how to respond,” *Nature*, vol. 497, pp. 51–59, 2013.
- [20] N. Jia and L. Tsui, “Epidemic modelling using SARS as a case study,” *North American Actuarial Journal*, vol. 9, no. 4, pp. 28–42, 2005.
- [21] X.-B. Hu, M. Wang, M. S. Leeson, E. L. Hines, and E. Di Paolo, “Deterministic ripple-spreading model for complex networks,” *Physical Review E*, vol. 83, no. 4, Article ID 046123, 2011.
- [22] X. B. Hu, M. Wang, and M. S. Leeson, “Ripple-spreading network model optimization by genetic algorithm,” *Mathematical Problems in Engineering*, vol. 2013, Article ID 176206, 15 pages, 2013.
- [23] J. H. Holland, *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, Mich, USA, 1975.
- [24] E. Goldberg, *Genetic Algorithms in Search Optimization & Machine Learning*, Addison-Wesley, Reading, Mass, USA, 1989.
- [25] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, Springer, Berlin, Germany, 2nd edition, 1994.
- [26] A. E. Eiben and J. E. Smith, *Introduction to Evolutionary Computing*, Natural Computing Series, Springer, Berlin, Germany, 2003.

Research Article

On Classification of PDZ Domains: A Computational Study

Wasim Aftab,¹ Adnan Memic,² and Dumitru Baleanu^{3,4,5}

¹ Department of Electrical and Computer Engineering, King Abdulaziz University, Jeddah, Saudi Arabia

² Centre of Nanotechnology, King Abdulaziz University, Jeddah, Saudi Arabia

³ Department of Mathematics and Computer Sciences, Faculty of Arts and Sciences, Cankaya University, 06530 Ankara, Turkey

⁴ Chemical and Materials Engineering Department, Faculty of Engineering, King Abdulaziz University, P.O. Box 80204, Jeddah 21589, Saudi Arabia

⁵ Institute of Space Sciences, Magurele-Bucharest, 077125, Romania

Correspondence should be addressed to Adnan Memic; amemic@kau.edu.sa and Dumitru Baleanu; dumitru@cankaya.edu.tr

Received 3 July 2013; Revised 31 July 2013; Accepted 31 July 2013

Academic Editor: J. A. Tenreiro Machado

Copyright © 2013 Wasim Aftab et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Our goal in this present study is to introduce new wavelet based methods for differentiating and classifying Class I and Class II PDZ domains and compare the resulting signals. PDZ domains represent one of the most common protein homology regions playing key roles in several diseases. To perform the classification, we developed two methods. The first of our methods was comparable to the standard wavelet approaches while the second one surpasses it in recognition accuracy. Our models exhibited interesting results, and we anticipate that it can be used as a computational technique to screen out the misfit candidates and to reduce the search space, while achieving high classification and accuracy.

1. Introduction

PDZs are structural domains contained in many proteins. They have been shown to act as key players involved in numerous diseases mediated with the PDZ domain interactions they contain [1–3]. PDZs represent one of the most common protein domains found in the human genome that are made up of approximately 90 amino acids. They were initially identified based on the homology of three different proteins from which their name is derived. One of the main functions that PDZ domains have been recognized for is their scaffolding or mediator role [4–6] in the assembly of receptors at the cellular membrane interface. These reactions are governed by PDZ domain binding of C-termini of their ligand protein, more specifically the last four to six amino acid residues [7–9]. Overall, there are several PDZ domain classifications; however, the two most dominant recognition motifs are $X-S/T-X-\Phi$ for Class I PDZ domains and $X-\Phi-X-\Phi$ for Class II PDZ domains [10]. Most PDZ domains share a similar 3D globular module [11] as shown by NMR and crystallography structures. In general, PDZs are composed of

six beta strands and two alpha helices; however, predicting peptide-PDZ domain interaction can be difficult. Our goal was to generate a computational model that could aid in interaction classifications. Given that PDZ domains often exhibit strong binding towards certain peptide ligands, they can be classified based on the recognition sequence of interacting peptides. However, with more recent reports, it is clear that PDZ domains display significant promiscuity in binding [12], proving to be a challenge for their classification. By extrapolating the relevant sequence information and using several wavelet transformation, we intended to develop a method that would be time saving and offer an alternative to expensive experiments. Work on this topic by Kalyoncu et al. [13] was based on using a statistical learning model for automated prediction of PDZ domains. Chen et al. [14] used a statistical model to perform prediction studies based on domain-peptide interactions. Another interesting classification technique according to the two critical positions of PDZ domains has been introduced by Bezprozvanny and Maximov [15]. Chen et al. [14] used a multidomain selectivity model to predict PDZ domain-peptide interactions across

TABLE 1: Comparisons between our methods and MODWT.

Methods	Sensitivity (%)			Specificity (%)			Positive predictive value (%)			Negative predictive value (%)		
	WAD-1	WAD-2	MODWT	WAD-1	WAD-2	MODWT	WAD-1	WAD-2	MODWT	WAD-1	WAD-2	MODWT
Parametric classifier	81.82	91.67	81.82	66.67	100	55.56	75	100	69.23	75	90	71.43
K-nearest neighbors	72.73	75	63.64	33.33	77.78	88.89	57.14	81.82	87.5	50	70	66.67

Sensitivity gives true positive rate, or the recall rate of prediction algorithm. See (10).

Specificity gives true negative rate of prediction algorithm. See (11).

Positive predictive value confirms that a correct prediction is actually correct. See (12).

Negative predictive value confirms that a false prediction is actually false. See (13).

Parametric classifier and K-nearest neighbors are popular pattern classification algorithms. See [27].

the mouse proteome. However, in this study, we have converted the biological prediction problem into an engineering challenge guided by mapping protein domain sequences into signals based on the 7 key physiochemical properties of amino acids (hydrophobicity [16], electronic [17], isoelectric point [18], polarity [19], volume [19], composition [20], and molecular weight [20]).

In this paper, the problem statement is as follows: given any PDZ sequence $S = \text{GTRITLEEITLERA}$, predict to which class of PDZ (I or II) S belongs.

The core of our study is the feature extraction from the amino acid sequences. In order to extract features from our dataset, we have used wavelets-based signal processing methodologies because it retains more abundant information of sequence order in frequency domain and time domain [21], which is a key to our method. In this study, we were able to develop two novel methods for feature extraction and classification. Feature extraction in the first method named WAD-1 is achieved by first smoothing the signal incorporating empirical mode decomposition (EMD) [1] and second treating normalized signals with maximum overlap discrete wavelet transform (MODWT) [22].

Our second method named WAD-2 uses trigram frequencies of amino acids followed by MODWT for feature extraction. The idea of amino acid frequency has been used popularly in many past [23, 24] and recent [13, 25, 26] proteomics analysis with various mathematical and statistical models.

Finally, we used different classifiers, namely, parametric classifier [27] and K-nearest neighbors [27], to classify these features. The performance of current methods showed improvement compared with wavelets only method in Table 1.

2. Methods

2.1. Empirical Mode Decomposition (EMD). According to Huang et al. [28], the empirical mode decomposition (EMD) [28] is a technique used to decompose a given signal into a set of elementary signals called “intrinsic mode functions” (IMFs). The algorithm operates by removing IMFs in every iteration. Rato et al. [29] has proposed an improved EMD algorithm, which is as follows from their paper.

Presenting any signal, $\chi(t)$, the IMFs are generated by an iterative activity titled sifting operations, which consist of the following steps.

- (i) Generate each and every localized maxima, M_x , $x = 1, 2, \dots$, and so forth, and minima, m_y , $y = 1, 2, \dots$ and so on, in $\chi(t)$.
- (ii) Generate the interpolating signals for maxima as $M(t) = f_M(M_x, t)$ and for minima as $m(t) = f_m(m_y, t)$. The interpolating signals thus computed contributes to the upper and lower envelopes of the signal.
- (iii) Consider $e(t) = (M(t) + m(t))/2$.
- (iv) Remove $e(t)$ from the original signal $\chi(t)$, and update it as $\chi(t) = \chi(t) - e(t)$.
- (v) Go back to step (i) and continue until $\chi(t)$ produced in step (iv) remains almost unvaried.
- (vi) After obtaining an IMF $\varphi(t)$, subtract it from updated $\chi(t)$ as $\chi(t) = \chi(t) - \varphi(t)$, and jump to step (i) if multiple extremum (a maxima or minima) for $\chi(t)$ is noticed.

This IMF can be mathematically defined as [29]

$$\begin{aligned} \varphi(t) &= \operatorname{Re} \left\{ |y(t)| e^{j\arg(y(t))} \right\}, \\ \varphi(t) &= |y(t)| \cos [\theta(t)], \\ \varphi(t) &= |y(t)| \cos [\arg(y(t))]. \end{aligned} \quad (1)$$

We also compute a ratio,

$$R = \frac{E(\chi(t))}{E(e(t))}. \quad (2)$$

Here, $E(\chi(t))$ and $E(e(t))$, respectively, indicate the energy of the original signal before sifting and the average energy of the upper and lower envelopes. However, if R crosses-over an allowed threshold of resolution τ , then the IMF computation is stopped.

2.2. Maximum Overlap Discrete Wavelet Transform (MODWT). MODWT was described by Percival and Walden [22]; it is a special case among the currently available wavelet-based techniques for the analysis of arbitrary-length discrete time series. MODWT is different from DWT in a sense that it is a circular shift-invariant transform [6]; the idea here is that, if a circular shift operation is applied to the real time

series data, then it generates identically shifted MODWT coefficients. The MODWT is well suited for any sequence length N , whereas for a complete decomposition of J levels the DWT requires N to be a multiple of 2^J .

Moreover, we can interpret the MODWT as a cyclic version of the DWT, and it is achieved by averaging over all nonredundant DWTs of shifted versions of the original series. However as this operation smoothes the original DWT signals, it also increases the running time of the computer program. If we apply DWT on an n -point time series, it requires $O(n)$ multiplications whereas MODWT for the same series requires $O(n \log 2n)$ multiplications [30].

It is well suited for our study because, once the protein sequence is translated to a numerical sequence, it becomes a time-series sequence $\{X_t \mid t = 0, 1, 2, \dots, N - 1\}$. In order to filter the signals at each level of $\{X_t\}$, MODWT treats the time series as a periodical. The MODWT coefficients are given by [31]

$$\tilde{W}_{j,t} = \sum_{l=0}^{Lj-1} \tilde{h}_{j,l} X_{t-l \bmod N}, \quad (3a)$$

$$\tilde{V}_{j,t} = \sum_{l=0}^{Lj-1} \tilde{g}_{j,l} X_{t-l \bmod N}. \quad (3b)$$

Here, $\tilde{h}_{j,l}$, $\tilde{g}_{j,l}$ are the high and low pass filters, respectively, of level j . It is also evident from the above equations that the unseen values $X_{-1}, X_{-2}, X_{-3}, \dots, X_{-N+1}, X_{-N}$ are the same as the observed values $X_{N-1}, X_{N-2}, X_{N-3}, \dots, X_1, X_0$, which indicates that MODWT induces “cyclic boundary conditions” during wavelet transformation.

The cyclic boundary condition can be difficult to implement in case of nonperiodic signals that exhibit discontinuities between start and end times [31]; therefore, common adoption [31] is “reflection boundary conditions”, in which the time series is extended to $2N$ instead of N . Due to reflection symmetry, the unseen values $X_{-1}, X_{-2}, X_{-3}, \dots, X_{-N+1}, X_{-N}$ are assigned to the seen values $X_0, X_1, X_2, \dots, X_{N-2}, X_{N-1}$. Therefore, (3a) and (3b) can be rewritten as in [31]:

$$\tilde{W}_{j,t} = \sum_{l=0}^{Lj-1} \tilde{h}_{j,l} \dot{X}_{t-l \bmod 2N}, \quad (4a)$$

$$\tilde{V}_{j,t} = \sum_{l=0}^{Lj-1} \tilde{g}_{j,l} \dot{X}_{t-l \bmod 2N}. \quad (4b)$$

Here, $\{\dot{X}_t\}$ is the extension of $\{X_t\}$.

3. Prediction with WAD-1

The improved EMD normalizes the signal to a unit power [29], which was not the case with original EMD [28]. Hence, improved EMD is more suitable for low amplitude biomedical signals [29] since it keeps all the components of a signal and does not decompose into different levels of resolution. The improved EMD algorithm also does not reduce the feature

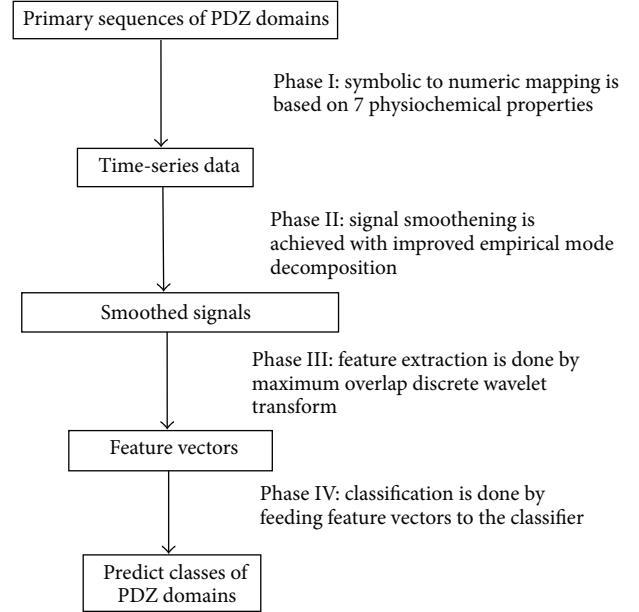


FIGURE 1: Block diagram depicting flowchart of method WAD-1.

space. Feature space reduction is an important part of our job because high dimensional feature space increases the running time of prediction algorithms.

Therefore, we used a combination of MODWT and improved EMD for our feature extraction part because improved EMD preprocesses a signal by interpolating all its maxima and minima, which in turn enhances the signal fit. The signal thus generated is fed to MODWT for feature extraction because it preserves more abundant information of sequence order in both frequency as well as time domains, and at the same time it reduces the feature space, which in turn improves the run time complexity of the classifiers. In essence, with our combination, we achieved reducing the feature space without losing sequence order information.

Our dataset is primarily based on the dataset of Tonikian et al. [32]. We have extracted only human PDZ domains for our work from their dataset.

Our first algorithm WAD-1 operates in four phases as shown in Figure 1. It can be seen from Figure 1 that, in the first phase, mapping from symbolic domain to numeric equivalent is performed. For WAD-1 method this mapping is based on the 7 physiochemical properties of amino acids; that is, every amino acid in each sequence is converted to a real number. With this operation, we obtain time-series data. We smoothen this signal by applying EMD. Let $\hat{S} = \{\hat{s}_0, \hat{s}_1, \dots, \hat{s}_{N-1}\}$ indicate a smoothed signal, where \hat{s}_i indicates the smoothed signal value corresponding to the i th position in the sequence.

We then applied MODWT operation of level J on \hat{S} as follows from [33]:

$$W = \tilde{W}\hat{S}. \quad (5)$$

Here, W is a nonorthogonal real matrix of dimension $(J+1)N \times N$. The MODWT coefficient vector from (1) can be resolved into $(J+1)$ vector as in [33]:

$$\widetilde{W} = [\widetilde{W}_1, \widetilde{W}_2, \dots, \widetilde{W}_J, \widetilde{V}_J]. \quad (6)$$

Here, \widetilde{W}_j ($j = 1, 2, \dots, J$) implies the length of $N/2^j$ vector of wavelet coefficients associated with the change on scale of length $L = 2^{j-1}$, and \widetilde{V}_J implies the length of $N/2^J$ vector of scaling coefficients associated with averages on scale of length 2^J .

From [22, 34], it is evident that MODWT transformation preserves energy of the signal. The following equation proposed from Gupta et al. [33] is shown here:

$$\begin{aligned} \|\widehat{S}\|^2 &= \|W\|^2 \\ \|\widehat{S}\|^2 &= \sum_{j=1}^J \|\widetilde{W}_j\|^2 + \|\widetilde{V}_J\|^2. \end{aligned} \quad (7)$$

The variance of \widehat{S} can be written similarly as in Gupta et al. [33]:

$$\sigma^2 = \frac{1}{N} \|\widehat{S}\|^2 - \bar{\widehat{S}}^2. \quad (8)$$

The feature vector generated through WAD-1 is influenced by the work of Gupta et al. with G protein coupled receptors (GPCRs) [33], and it is made up of the variances of several physicochemical properties of the amino acids in the PDZ sequence. For example, using (8), WAD-1 computes the part of the feature vector for a PDZ sequence based on EIIP values as

$$F_{EIIP} = [\sigma_{EIIP}^2(1), \sigma_{EIIP}^2(2), \dots, \sigma_{EIIP}^2(3)]. \quad (9)$$

Here, J indicates upper limit of the level of decomposition. The complete feature vector is obtained by concatenating 7 such feature vectors computed for 7 physicochemical properties of amino acids.

In this study, we chose the least asymmetric filter (LA8) for the wavelet transformation phase because it has a filter width short enough that any impact caused by boundary conditions stays within the tolerance limit.

We have used ANOVA in the feature extraction phase as this analysis is useful in comparing multiple variables for statistical significance. Gupta et al. [33] have used this method to extract features from amino acid sequences in the classification study of GPCRs.

This feature extraction is the most important phase in the overall classification and analysis since it yields the feature vector (FV), whose dimension is function of J , the number of levels chosen for decomposing the signal. For $J = 5$, (as in our case) the dimension of the feature vector is only $7 * 5 = 35$. This is a great reduction that improves the run time of the classification program. In the final phase, these FVs are fed to the classifier for prediction of PDZ domain.

For MODWT calculations in the 3rd phase, we have used the WMTSA wavelet toolkit [31] from MATLAB. With WAD-1 we are able to reduce the dimensionality of the feature

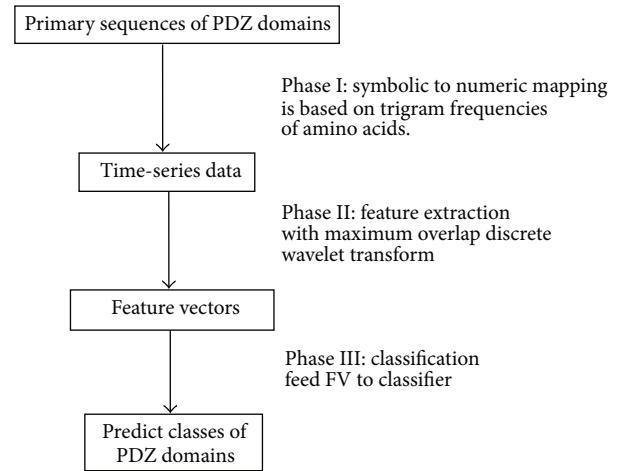


FIGURE 2: Block diagram depicting flowchart of method WAD-2.

space and at the same time retain the positional information of the amino acids in the original sequence. Table 1 clearly demonstrates the consistency of new WAD-1 when compared with popular MODWT.

4. Improved Prediction with WAD-2

In this method, we extract the trigram frequencies from the amino acids as previously reported [13]. For an amino acid sequence such as EITLERG, it has the following trigrams: EIT, ITL, TLE, LER, and ERG.

We calculate trigram frequency of amino acids of every PDZ sequence in the following way.

- (i) First, count the number of times a trigram appears in the sequence.
- (ii) Then, divide this number by the total number of trigrams in the sequence; in fact, it is found to be $(L - 2)$, where L denotes length of a PDZ sequence.
- (iii) Repeat steps (i) and (ii) for every possible trigram for the PDZ sequence.

We reduce the dimensions of the features by applying the idea introduced by Kalyoncu et al. [13], where 20 amino acids are grouped into 7 distinct classes based on dipoles and volumes.

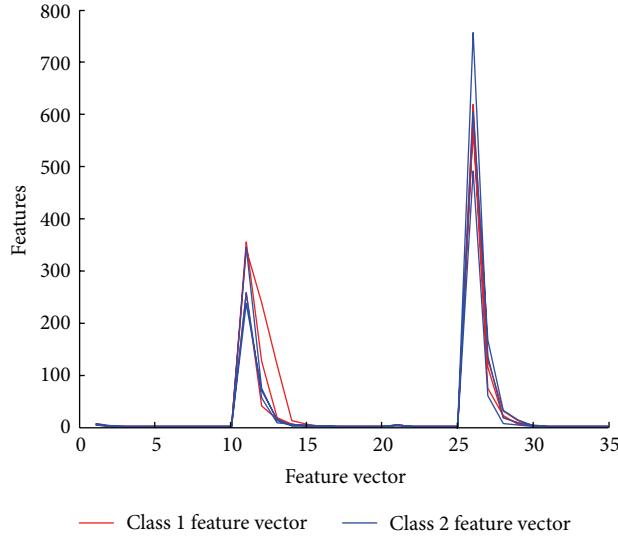
The block diagram of this method is shown in Figure 2, and we call it WAD-2. It operates in 3 phases.

Once the symbolic to numeric mapping is done, we applied MODWT as in WAD-1 for feature extraction.

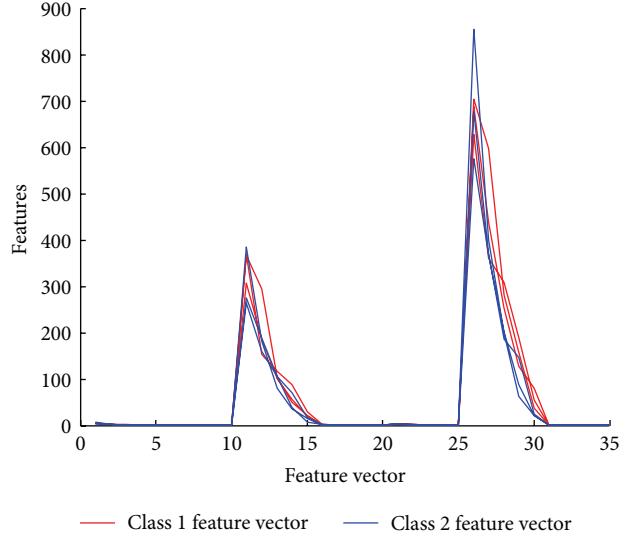
We have used “PCP: a program for supervised classification of gene expression profiles” [27] in our classification phases for both WAD-1 and WAD-2.

5. Results and Discussions

In this paper, we show proof-of-principle application for our algorithms. We evaluated our approaches by 2 different classifiers and also compared them to a wavelet method,



(a) Feature vectors of both classes of PDZ computed using WAD-1 (EMD smoothening followed by MODWT)



(b) Feature vectors of both classes of PDZ computed using MODWT alone

FIGURE 3: Comparison of Signals generated after Feature Extraction with WAD-1 and MODWT for classes I and II of PDZ domains. WAD-1 clearly smoothes the signal more when compared to MODWT.

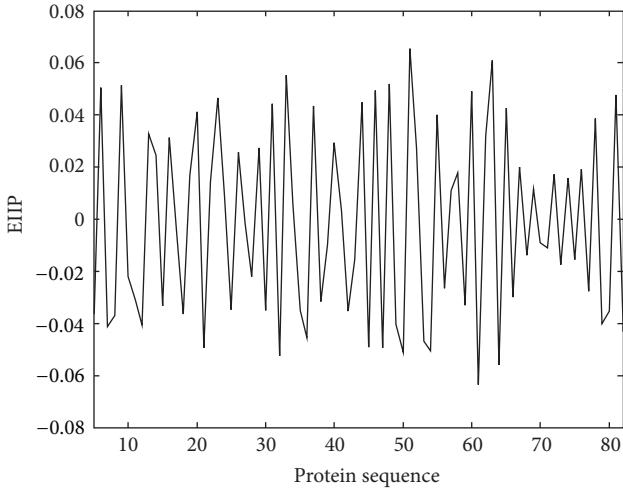


FIGURE 4: Mapped protein sequence (organism: human, PDZ name: DLG1, and length: 84 amino acids) using EIIP scale. This is the first step for WAD-1; it involves the transformation of amino acid sequences into a numerical sequence.

MODWT. We observed that our results are comparable with the wavelet method, and while our second method WAD-2 has shown promising results, the first method WAD-1 is able to produce consistent results (see Table 1).

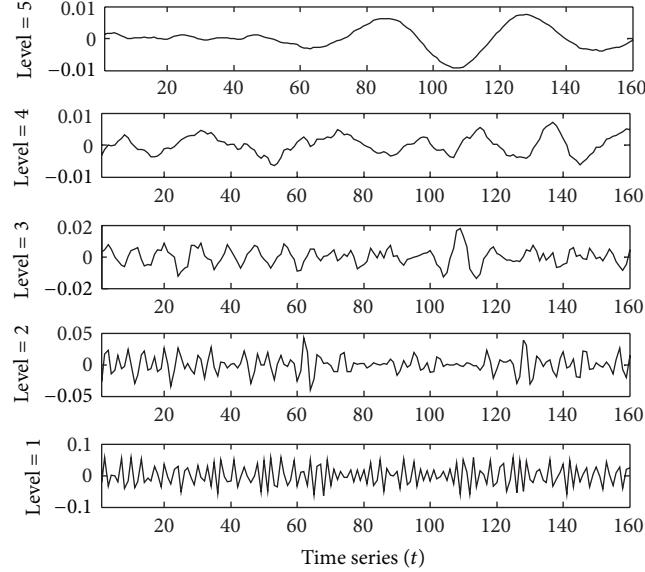
Figure 4 shows the signal based on EIIP values. We show the resulting signals decomposed up to 5 levels by our first method in Figure 5(a). It shows a cumulative abstraction of the variations in the data over regions progressive to the wavelets scales with coefficients at higher levels; therefore, it depicts features at broader time intervals with the increase in the values of j .

We show the feature vector obtained for a PDZ sequence by our first method based on EIIP, isoelectric, composition, polarity, volume, and molecular weight values in Figures 5(b)–5(g), respectively. It is evident from these figures that wavelet variance value diminishes to almost zero for higher levels (after 3rd level) for almost all the physicochemical values. This clearly indicates that this wavelet variance vector has 3 principal components. The comparison between two feature extractions is depicted in Figure 3, from which it is clear that our signals are smoother with sharper peaks when compared to MODWT signals.

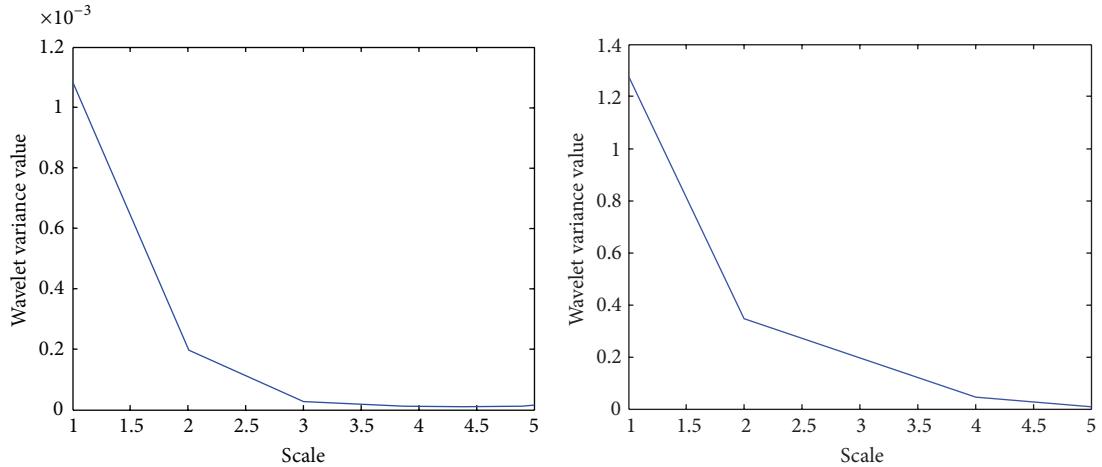
Table 1 shows the comparative predictive performance of WAD-1, WAD-2, and MODWT in terms of four statistical measures: sensitivity, specificity, negative predictive value, and positive predictive value.

Sensitivity and specificity [35] are statistical measures, which help to evaluate the performance of a binary classification method. In order to understand the significance of sensitivity and specificity in our work, we need to comprehend a few more terms like true positive, false positive, true negative, false negative, which are defined as

- (a) true positive: the sample belongs to Class 1, and the algorithm also recognizes it,
- (b) false positive: the sample belongs to Class 2, but the algorithm recognizes it otherwise,
- (c) true negative: the sample belongs to Class 2, and the algorithm also recognizes it,
- (d) false negative: the sample belongs to Class 1, but the algorithm recognizes it otherwise.

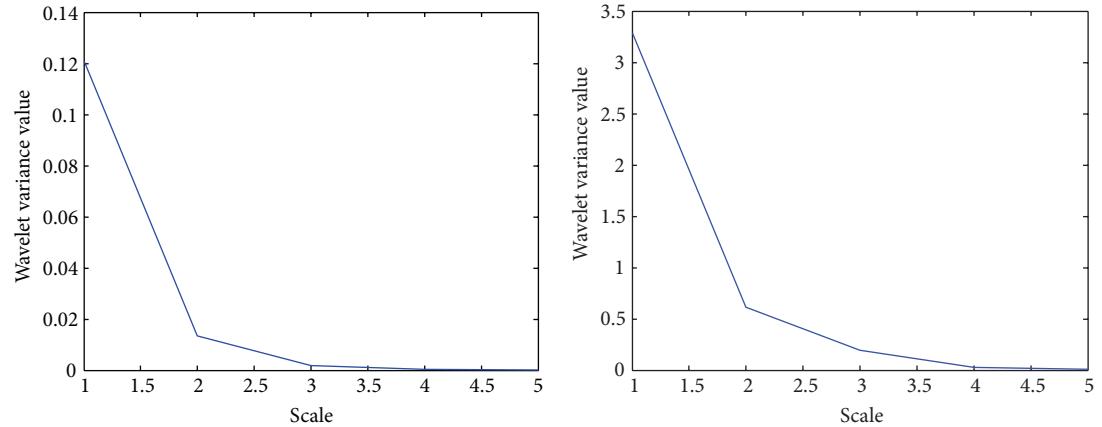


(a) Wavelet coefficients of the mapped protein sequence up to level 5 for EIIP value



(b) Wavelet variance vector, V_{EIIP} for EIIP value, obtained for the input protein sequence

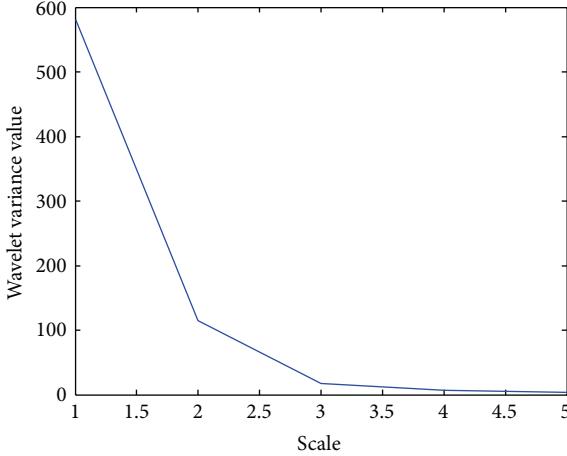
(c) Wavelet variance vector, $V_{ISOELECTRIC}$ for isoelectric value, obtained for the input protein sequence



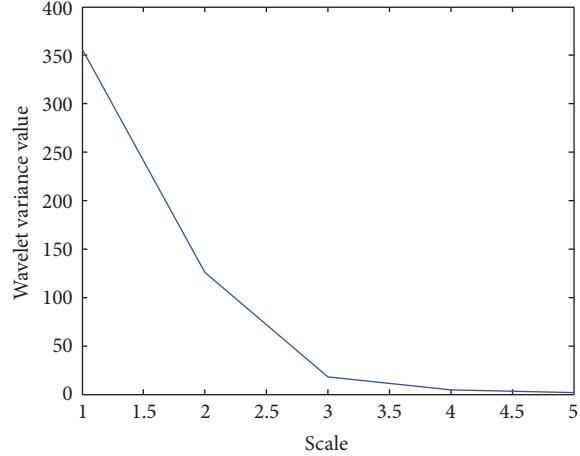
(d) Wavelet variance vector, $V_{COMPOSITION}$ for composition value, obtained for the input protein sequence

(e) Wavelet variance vector, $V_{POLARITY}$ for polarity value, obtained for the input protein sequence

FIGURE 5: Continued.



(f) Wavelet variance vector, V_{VOLUME} for volume value, obtained for the input protein sequence



(g) Wavelet variance vector, $V_{\text{MOL-WT}}$ for molecular weight value obtained for the input protein sequence

FIGURE 5: (a) Wavelet coefficients and (b)–(g): wavelet variance vectors for different physicochemical values. Wavelet variance vector is computed by concatenating the wavelet variances for all the levels from $j = 1$ to J . It is computed for all the 7 physicochemical values of amino acids. And the final feature vector for any PDZ sequence is a concatenation of such wavelet variance vectors. From Figures 5(b)–5(g), it is evident that, for most of the physicochemical properties, the wavelet variance vector has 3 principal components.

Therefore, mathematically, sensitivity and specificity are defined as

$$\text{sensitivity} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}, \quad (10)$$

$$\text{specificity} = \frac{\text{true negatives}}{\text{true negatives} + \text{false negatives}}. \quad (11)$$

According to Akobeng [36], sensitivity and specificity are important measures of the uninflected statement of an assay but cannot be used to categorize the future developments of the sickness in an individual. But other measures like positive and negative predictive values unveil the prospects regarding the probability of a prediction being correct.

Positive predictive value [37] can be described mathematically as

$$\text{positive predictive value (PPV)} = \frac{\alpha}{\alpha + \alpha'}. \quad (12)$$

Here, α = number of samples that actually belongs to Class 1 and α' = number of samples that does not actually belong to Class 1.

Similarly, negative predictive value [37] is defined mathematically as

$$\text{negative predictive value (NPV)} = \frac{\beta}{\beta + \beta'}. \quad (13)$$

Here, β = number of samples that actually belongs to Class 2 and β' = number of samples that does not actually belong to Class 2.

Both NPV and PPV are crucial estimations of the performance of a prediction algorithm because PPV estimates the probability that a true prediction is actually true, and NPV confirms that a false prediction is actually false.

From the context of medical diagnostics, if we consider Class 1 samples as positives and Class 2 samples as negatives, then a high negative predictive value means that the method very rarely recognizes a negative sample (Class 2) as a healthy one (Class 1). In this regard, our second method performs very well (see Table 1).

However, NPV does not consider cases when a positive sample is identified as a negative one, or in our case it, does not give any idea on the chances of misclassification of sample from Class 1 to Class 2.

Therefore, we calculated PPV which complements this drawback of NPV. Similar arguments also apply in case of predictive drawbacks of PPV.

From Table 1, we observe that WAD-1 performed considerably well and WAD-2 surprised us in recognition accuracy. Kalyoncu et al. [13] have done a sophisticated computational study on interaction prediction and classification of PDZ domains, where they achieved an excellent maximum sensitivity of 90.7%. With our WAD-2 method, we are able to achieve similar results with the highest sensitivity of 91.67%.

6. Conclusions

In this work, we have successfully classified PDZ domains of Classes I and II only. We introduced a new method for feature extraction: an EMD smoothing of signals followed by a MODWT transform. We further introduced WAD-2 method based on the trigram frequency of amino acids and found that the results were improved in terms of sensitivity and accuracy. We note that our second WAD-2 method performed better in recognition accuracy than the WAD-1 method. As mentioned earlier, our work in this paper is meant as a proof-of-principle application for our algorithms. We are enthusiastic that further improvement of our algorithm can lead to even

better accuracy and predicting of the promiscuity of PDZ domains.

Abbreviations

EMD:	Empirical mode decomposition
MODWT:	Maximum overlap discrete wavelet transform
IMFs:	Intrinsic mode functions
GPRC:	G protein coupled receptors
FV:	Feature vector
ANOVA:	Analysis of variance
PPV:	Positive predictive value
NPV:	Negative predictive value.

Acknowledgments

This study was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, Saudi Arabia under Grant no. (903-004-D1434). The authors thank the DSR for the technical and financial support.

References

- [1] V. Raghuram, D. O. D. Mak, and J. K. Foskett, "Regulation of cystic fibrosis transmembrane conductance regulator single-channel gating by bivalent PDZ-domain-mediated interaction," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 3, pp. 1300–1305, 2001.
- [2] C. P. Ponting, C. Phillips, K. E. Davies, and D. J. Blake, "PDZ domains: targeting signalling molecules to sub-membranous sites," *BioEssays*, vol. 19, no. 6, pp. 469–479, 1997.
- [3] K. K. Dev, "Making protein interactions druggable: targeting PDZ domains," *Nature Reviews Drug Discovery*, vol. 3, no. 12, pp. 1047–1056, 2004.
- [4] B. Z. Harris and W. A. Lim, "Mechanism and role of PDZ domains in signaling complex assembly," *Journal of Cell Science*, vol. 114, no. 18, pp. 3219–3231, 2001.
- [5] A. S. Fanning and J. M. Anderson, "PDZ domains: fundamental building blocks in the organization of protein complexes at the plasma membrane," *Journal of Clinical Investigation*, vol. 103, no. 6, pp. 767–772, 1999.
- [6] S. Tsunoda, J. Sierralta, Y. Sun et al., "A multivalent PDZ-domain protein assembles signalling complexes in a G-protein-coupled cascade," *Nature*, vol. 388, no. 6639, pp. 243–249, 1997.
- [7] D. A. Doyle, A. Lee, J. Lewis, E. Kim, M. Sheng, and R. MacKinnon, "Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ," *Cell*, vol. 85, no. 7, pp. 1067–1076, 1996.
- [8] Z. Songyang, A. S. Fanning, C. Fu et al., "Recognition of unique carboxyl-terminal motifs by distinct PDZ domains," *Science*, vol. 275, no. 5296, pp. 73–77, 1997.
- [9] M. Sheng and C. Sala, "PDZ domains and the organization of supramolecular complexes," *Annual Review of Neuroscience*, vol. 24, pp. 1–29, 2001.
- [10] U. Wiedemann, P. Boisguerin, R. Leben et al., "Quantification of PDZ domain specificity, prediction of ligand affinity and rational design of super-binding peptides," *Journal of Molecular Biology*, vol. 343, no. 3, pp. 703–718, 2004.
- [11] H. Tochio, F. Hung, M. Li, D. S. Bredt, and M. Zhang, "Solution structure and backbone dynamics of the second PDZ domain of postsynaptic density-95," *Journal of Molecular Biology*, vol. 295, no. 2, pp. 225–237, 2000.
- [12] H. J. Lee and J. J. Zheng, "PDZ domains and their binding partners: structure, specificity, and modification," *Cell Communication and Signaling*, vol. 8, article 8, 2010.
- [13] S. Kalyoncu, O. Keskin, and A. Gursoy, "Interaction prediction and classification of PDZ domains," *BMC Bioinformatics*, vol. 11, article 357, 2010.
- [14] J. R. Chen, B. H. Chang, J. E. Allen, M. A. Stiffler, and G. MacBeath, "Predicting PDZ domain-peptide interactions from primary sequences," *Nature Biotechnology*, vol. 26, no. 9, pp. 1041–1045, 2008.
- [15] I. Bezprozvanny and A. Maximov, "Classification of PDZ domains," *FEBS Letters*, vol. 509, no. 3, pp. 457–462, 2001.
- [16] J. Kyte and R. F. Doolittle, "A simple method for displaying the hydrophobic character of a protein," *Journal of Molecular Biology*, vol. 157, no. 1, pp. 105–132, 1982.
- [17] I. Cosic, "Macromolecular bioactivity: is it resonant interaction between macromolecules? Theory and applications," *IEEE Transactions on Biomedical Engineering*, vol. 41, no. 12, pp. 1101–1114, 1994.
- [18] J. M. Zimmerman, N. Eliezer, and R. Simha, "The characterization of amino acid sequences in proteins by statistical methods," *Journal of Theoretical Biology*, vol. 21, no. 2, pp. 170–201, 1968.
- [19] R. Grantham, "Amino acid difference formula to help explain protein evolution," *Science*, vol. 185, no. 4154, pp. 862–864, 1974.
- [20] http://pages.pomona.edu/~ac044747/aroc/Genetic_Code.swf.
- [21] Z. C. Li, X. B. Zhou, Z. Dai, and X. Y. Zou, "Prediction of protein structural classes by Chou's pseudo amino acid composition: approached using continuous wavelet transform and principal component analysis," *Amino Acids*, vol. 37, no. 2, pp. 415–425, 2009.
- [22] D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis*, Cambridge University Press, Cambridge, UK, 2002.
- [23] J. Cedano, P. Aloy, J. A. Perez-Pons, and E. Querol, "Relation between amino acid composition and cellular location of proteins," *Journal of Molecular Biology*, vol. 266, no. 3, pp. 594–600, 1997.
- [24] P. C. Ng and S. Henikoff, "Predicting deleterious amino acid substitutions," *Genome Research*, vol. 11, no. 5, pp. 863–874, 2001.
- [25] J. A. Tenreiro Machado, "Shannon information and power law analysis of the chromosome code," *Abstract and Applied Analysis*, vol. 2012, Article ID 439089, 13 pages, 2012.
- [26] J. A. Tenreiro Machado, A. C. Costa, and M. D. Quelhas, "Can power laws help us understand gene and proteome information?" *Advances in Mathematical Physics*, vol. 2013, Article ID 917153, 10 pages, 2013.
- [27] L. J. Buturović, "PCP: a program for supervised classification of gene expression profiles," *Bioinformatics*, vol. 22, no. 2, pp. 245–247, 2006.
- [28] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society A*, vol. 454, no. 1971, pp. 903–995, 1998.
- [29] R. T. Rato, M. D. Ortigueira, and A. G. Batista, "On the HHT, its problems, and some solutions," *Mechanical Systems and Signal Processing*, vol. 22, no. 6, pp. 1374–1394, 2008.
- [30] "Nonlinear denoising via wavelet shrinkage," <http://rss.acs.unt.edu/Rdoc/library/wmtsa/html/wavShrink.html>.

- [31] C. R. Cornish, C. S. Bretherton, and D. B. Percival, "Maximal overlap wavelet statistical analysis with application to atmospheric turbulence," *Boundary-Layer Meteorology*, vol. 119, no. 2, pp. 339–374, 2006.
- [32] R. Tonikian, Y. Zhang, S. L. Sazinsky et al., "A specificity map for the PDZ domain family," *PLoS Biology*, vol. 6, no. 9, article e239, 2008.
- [33] R. Gupta, A. Mittal, K. Singh, V. Narang, and S. Roy, "Time-series approach to protein classification problem," *IEEE Engineering in Medicine and Biology Magazine*, vol. 28, no. 4, pp. 32–37, 2009.
- [34] D. P. Percival, "On estimation of the wavelet variance," *Biometrika*, vol. 82, no. 3, pp. 619–631, 1995.
- [35] D. G. Altman and J. M. Bland, "Diagnostic tests 1: sensitivity and specificity," *British Medical Journal*, vol. 308, no. 6943, article 1552, 1994.
- [36] A. K. Akobeng, "Understanding diagnostic tests 1: sensitivity, specificity and predictive values," *Acta Paediatrica*, vol. 96, no. 3, pp. 338–341, 2007.
- [37] D. G. Altman and J. M. Bland, "Diagnostic tests 2: predictive values," *British Medical Journal*, vol. 309, no. 6947, article 102, 1994.

Research Article

Vibration Suppression of an Axially Moving String with Transverse Wind Loadings by a Nonlinear Energy Sink

Ye-Wei Zhang, Jian Zang, Tian-Zhi Yang, Bo Fang, and Xin Wen

Department of Astronautics, Shenyang Aerospace University, Shenyang 110136, China

Correspondence should be addressed to Ye-Wei Zhang; zhanguwei1218@126.com

Received 23 July 2013; Accepted 29 August 2013

Academic Editor: Dumitru Baleanu

Copyright © 2013 Ye-Wei Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nonlinear targeted energy transfer (TET) is applied to suppress the excessive vibration of an axially moving string with transverse wind loads. The coupling dynamic equations used are modeled by a nonlinear energy sink (NES) attached to the string to absorb vibrational energy. By a two-term Galerkin procedure, the equations are discretized, and the effects of vibration suppression by numerical methods are demonstrated. Results show that the NES can effectively suppress the vibration of the axially moving string with transverse wind loadings, thereby protecting the string from excessive movement.

1. Introduction

The structures of axial speed-dependent behaviors have been analyzed in numerous studies [1–6]. Various engineering systems can be simplified into an axially moving string model such as aerial gondola cableways and conveyor belts. Specifically, the lateral vibrations of an axially moving string have been extensively examined. Both the aerial gondola cableway and the conveyor belt operate in rural environments and suffer from transverse wind loads. Excessive wind loads can destroy the string running security, causing catastrophic failure. The dynamic response of the axially moving string has been widely investigated [7]; however, studies on the vibration suppression of an axially moving string with transverse wind loads are rarely reported.

Traditional linear vibration absorbers have been used in various engineering fields. However, vibration suppression strongly responds only at the natural frequency of the vibration absorber. The nonlinear energy sink (NES) functions as an effective vibration absorber for a nonlinear system. The NES has recently been reported to engage in resonance over a very broad frequency range, has a small additional mass, and can perform targeted energy transfer (TET). Nonlinear TET has been used in numerous engineering structures for vibration suppression, such as drill-string [8], beam [9], rod

[10], and plate [11]. Specifically, Lee et al. [12] attached the NES to the fixed wing of the plane for vibration suppression of the limit cycle, which resulted in increased damping. Nucera et al. [13] applied the NES to a multistory frame structure to absorb the vibration. Savadkoohi et al. [14] examined the four-story frame structure by using two parallel NES.

The present study focuses on the vibration suppression of an axially moving string with certain and steady transverse wind loads by using NES. The coupling dynamic differential equations of an axially moving string and the NES with transverse wind loads are established. In addition, the governing equations are approximately discretized by the two-term Galerkin procedure. The effects of vibration suppression are finally demonstrated by numerical simulation.

2. Equation of Motion

Figure 1 shows the system under study, consisting of a simply supported axially moving string with an essentially nonlinear damped attachment. The attachment called NES is expected to irreversibly absorb the vibrational energy of the string.

The length of the axially moving string is represented by L . Let $U(X, T)$ and $\bar{U}(X, T)$ be the displacements of the string and the NES relative to the horizontal X -axis, respectively.

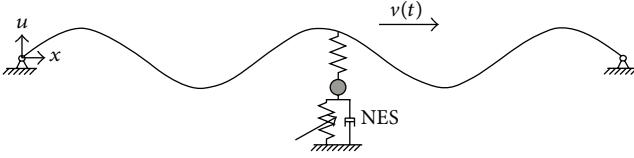


FIGURE 1: An axially moving string with a nonlinear energy sink (NES).

The governing equation of motion can be derived by Newton's second law:

$$\begin{aligned} & \rho A \left(\frac{\partial^2 U(X, T)}{\partial T^2} + 2V \frac{\partial^2 U(X, T)}{\partial X \partial T} + V^2 \frac{\partial^2 U(X, T)}{\partial X^2} \right) \\ & - P \frac{\partial^2 U(X, T)}{\partial X^2} + \eta \left(V \frac{\partial U(X, T)}{\partial X} + \frac{\partial U(X, T)}{\partial T} \right) \quad (1) \\ & = R(t) \delta(X - d) + F(X, T), \end{aligned}$$

where η is the viscosity coefficient of the string material, ρ is the linear density, A is the cross-sectional area, P is the initial tension, and V is the axial speed. $R(t)$ is the interaction force between the string and the NES.

In (1), $F(X, T)$ is expressed as follows [15]:

$$\begin{aligned} F(X, T) &= F_1 \frac{\partial U(X, T)}{\partial T} + F_3 \left(\frac{\partial U(X, T)}{\partial T} \right)^3, \quad (2) \\ F_1 &= 0.5a\rho_0 h v_0, \quad F_3 = 0.5b\rho_0 h v_0^{-1}. \end{aligned}$$

The NES equation of motion is given by

$$m_{\text{NES}} \frac{\partial^2 \bar{U}}{\partial T^2} + R(t) = 0. \quad (3)$$

The interaction force $R(t)$ can be written as

$$R(t) = K [\bar{U}(T) - Z(T)]^3 + D \left(\frac{\partial \bar{U}(T)}{\partial T} - \frac{\partial Z(T)}{\partial T} \right). \quad (4)$$

The attachment point displacement and velocity are expressed as follows [16]:

$$Z(T) = U(d, T) \quad \frac{\partial Z(T)}{\partial T} = \frac{\partial U(d, T)}{\partial T} + V \frac{\partial U(d, T)}{\partial X}, \quad (5)$$

where K is the nonlinear (cubic) spring stiffness, D is the NES dissipation, and d is the adding position to the NES on the string.

The nondimensional quantities are given as follows:

$$\begin{aligned} x &= \frac{X}{L}, & u &= \frac{U}{L}, & t &= \omega T, \\ v &= \frac{V}{\omega L}, & \lambda &= \frac{v_0}{\omega L}, & \alpha_0 &= \frac{\rho_0 L h}{\rho A}, \\ k &= \frac{K L^4}{P}, & \sigma &= \frac{D L}{\sqrt{\rho A P}}, & \varepsilon &= \frac{m_{\text{NES}}}{\rho A L^2}, \\ \eta_0 &= \frac{\eta}{\rho A \omega}, & \omega^2 &= \frac{P}{\rho A L^2}. \end{aligned} \quad (6)$$

Substituting (6) into (1) to (5) yields the following dimensionless form:

$$\begin{aligned} & \frac{\partial^2 u(x, t)}{\partial t^2} + 2v \frac{\partial^2 u(x, t)}{\partial x \partial t} + (v^2 - 1) \frac{\partial^2 u(x, t)}{\partial x^2} \\ & + \eta_0 \left(v \frac{\partial u}{\partial x} + \frac{\partial u}{\partial t} \right) \\ & = \left[k[u(d, t) - \bar{u}(t)]^3 + \sigma \left(\frac{\partial \bar{u}(t)}{\partial t} - \frac{\partial u}{\partial t}(d, t) - v \frac{\partial u}{\partial x}(d, t) \right) \right] \\ & \times \delta(x - d) + 0.5\alpha_0 \left(a\lambda \frac{\partial u(x, t)}{\partial t} + \frac{b}{\lambda} \left(\frac{\partial u(x, t)}{\partial t} \right)^3 \right), \\ & \varepsilon \frac{\partial^2 \bar{u}}{\partial t^2} + k[\bar{u}(t) - u(d, t)]^3 \\ & + \sigma \left(\frac{\partial \bar{u}(t)}{\partial t} - \frac{\partial u(d, t)}{\partial t} - \gamma \frac{\partial u(d, t)}{\partial x} \right) = 0. \end{aligned} \quad (7)$$

3. The Galerkin Method

By using a tractable finite-dimensional dynamical system, governing equations (7) can be approximated by the standard Galerkin-type projections as follows:

$$u(x, t) = \sum_{r=1}^N \phi_r(x) q_r(t), \quad (8)$$

where $\phi_r(x)$ denotes the eigenfunctions for the free undamped vibrations of a string satisfying the same boundary conditions and $q_r(t)$ represents the generalized coordinates of the discretized system.

Substituting (8) into (7) yields

$$\begin{aligned} & \sum_{r=1}^N \left[\phi_r(x) \ddot{q}_r(t) + 2\gamma \phi'_r(x) \dot{q}_r(t) + (v^2 - 1) \phi''_r(x) q_r(t) \right. \\ & \left. + \eta_0 v \phi'_r(x) q_r(t) + \eta_0 \phi_r(x) \dot{q}_r(t) \right] \\ & = \left[k \left(\sum_{r=1}^N \phi_r(d) q_r(t) - \bar{u}(t) \right)^3 \right. \\ & \left. + \sigma \left(\sum_{r=1}^N \phi_r(d) \dot{q}_r(t) + \gamma \sum_{r=1}^N \phi'_r(d) q_r(t) - \dot{\bar{u}}(t) \right) \right] \\ & \times \delta(x - d) + 0.5\alpha_0 a\lambda \sum_{r=1}^N \phi_r(x) \dot{q}_r(t) \\ & + 0.5\alpha_0 \frac{b}{\lambda} \left(\sum_{r=1}^N \phi_r(x) \dot{q}_r(t) \right)^3, \end{aligned} \quad (9)$$

$$\begin{aligned} & \varepsilon \ddot{\bar{u}}(t) + k \left(\bar{u}(t) - \sum_{r=1}^N \phi_r(d) q_r(t) \right)^3 \\ & + \sigma \left(\dot{\bar{u}}(t) - \sum_{r=1}^N \phi_r(d) \dot{q}_r(t) - \gamma \sum_{r=1}^N \phi'_r(d) q_r(t) \right) = 0. \end{aligned} \quad (10)$$

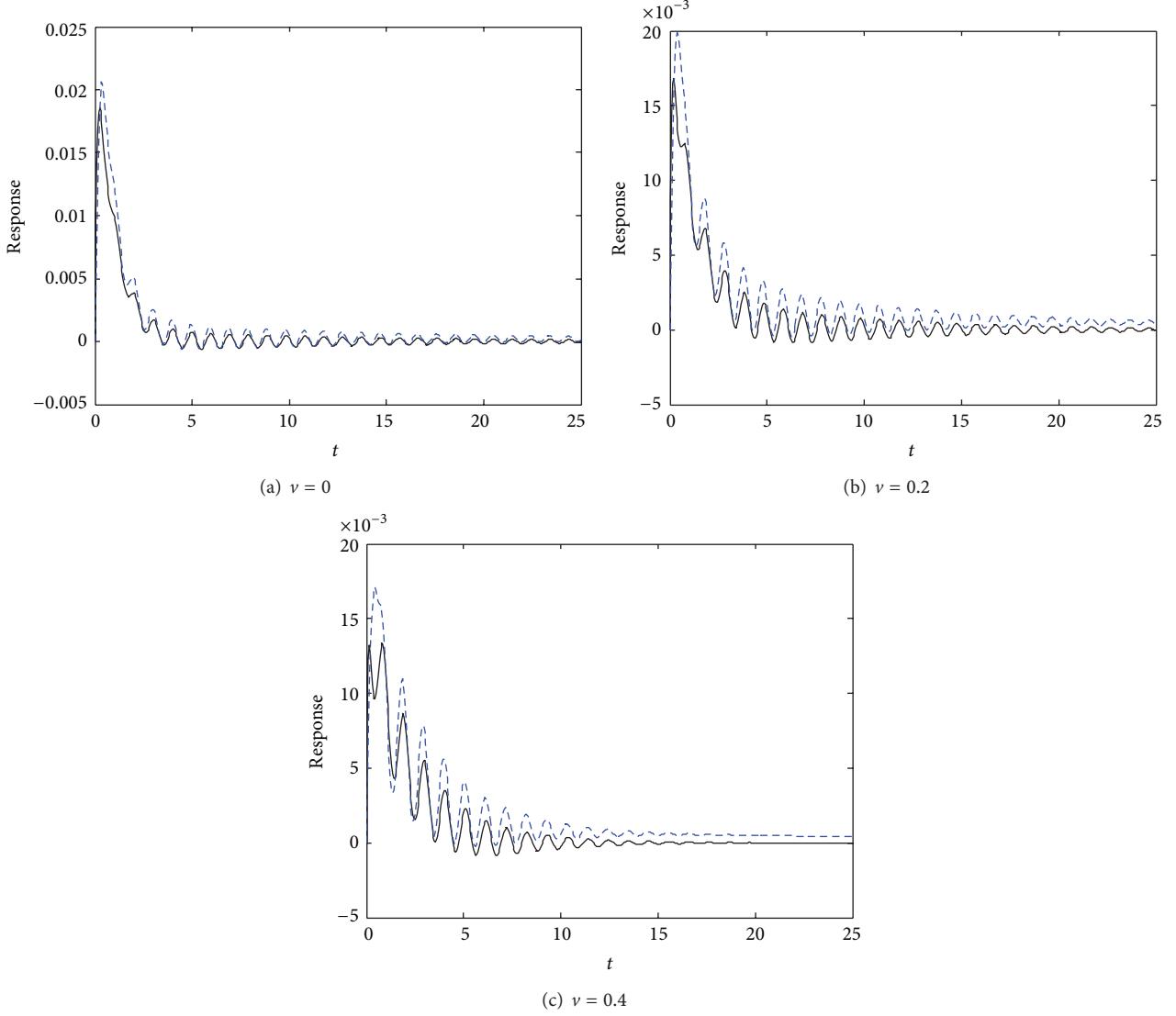


FIGURE 2: Response of the axially moving string and NES for different axial speeds (solid line: axially moving string's response, u ; dashed line: NES's response, \bar{u}).

In (10), for example, $\phi_r = \sqrt{2} \sin \lambda_r x$ and $\lambda_r = r\pi$ if a string is supported by pinned ends. The $\sqrt{2}$ factor ensures orthonormality. Multiplying both (9) and (10) by $\phi_s(x)$ and integrating over the domain $[0, 1]$ yields the following equations:

$$\begin{aligned} & \delta_{sr} \ddot{q}_r(t) + 2\nu b_{sr} \dot{q}_r(t) + c_{sr} (\nu^2 - 1) q_r(t) \\ & + \eta_0 \nu b_{sr} q_r(t) + \eta_0 \delta_{sr} \dot{q}_r(t) \\ & = \left[k \left[\sum_{r=1}^N \phi_r(d) q_r(t) - \bar{u}(t) \right]^3 \right. \\ & \quad \left. + \sigma \left(\sum_{r=1}^N \phi_r(d) \dot{q}_r(t) + \gamma \sum_{r=1}^N \phi'_r(d) q_r(t) - \dot{\bar{u}}(t) \right) \right] \\ & \times \phi_r(d) + 0.5\alpha_0 a \lambda \delta_{sr} \dot{q}_r(t) + 0.5\alpha_0 \frac{b}{\lambda} e_{s1} \dot{q}_1^3(t) \end{aligned}$$

$$\begin{aligned} & + 0.5\alpha_0 \frac{b}{\lambda} e_{s2} \dot{q}_2^3(t) + 1.5\alpha_0 \frac{b}{\lambda} f_{s1} \dot{q}_1^2(t) \dot{q}_2(t) \\ & + 1.5\alpha_0 \frac{b}{\lambda} f_{s2} \dot{q}_1(t) \dot{q}_2^2(t), \\ & \varepsilon \ddot{\bar{u}}(t) + k \left(\bar{u}(t) - \sum_{r=1}^N \phi_r(d) q_r(t) \right)^3 \\ & + \sigma \left(\dot{\bar{u}}(t) - \sum_{r=1}^N \phi_r(d) \dot{q}_r(t) - \gamma \sum_{r=1}^N \phi'_r(d) q_r(t) \right) = 0, \end{aligned} \tag{11}$$

where

$$\begin{aligned} \delta_{sr} &= \int_0^1 \phi_s(x) \phi_r(x) dx, & b_{sr} &= \int_0^1 \phi_s(x) \phi'_r(x) dx, \\ c_{sr} &= \int_0^1 \phi_s(x) \phi''_r(x) dx, & e_{sr} &= \int_0^1 \phi_s(x) \phi_r^3(x) dx, \end{aligned}$$

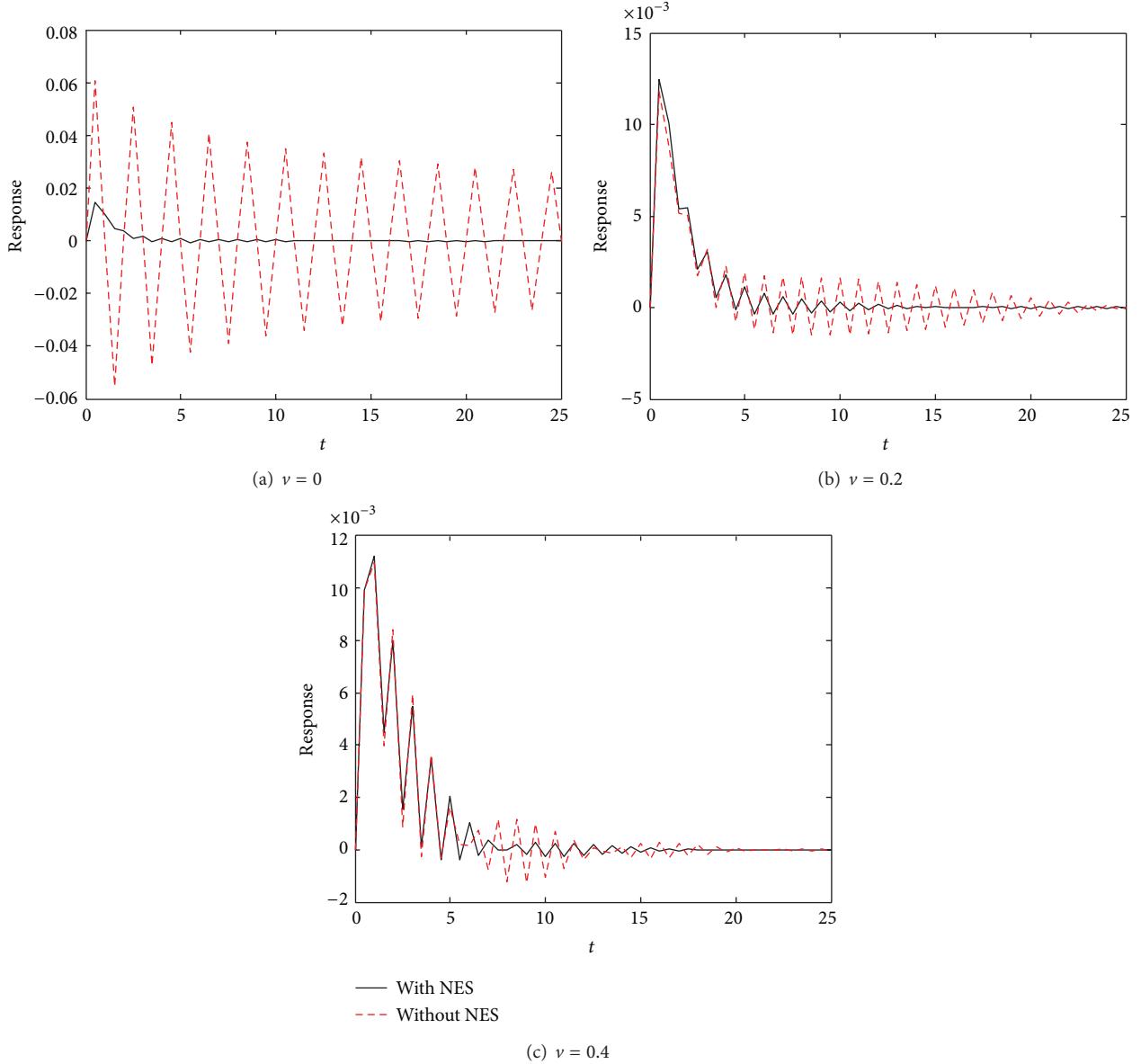


FIGURE 3: Comparison of the transient response of axially moving string with and without NES under varying speed (solid line: string coupled with NES; dotted line: string without NES).

$$\begin{aligned} f_{s2} &= \int_0^1 \phi_s(x) \phi_1(x) \phi_2^2(x) dx, \\ f_{s1} &= \int_0^1 \phi_s(x) \phi_1^2(x) \phi_2(x) dx, \end{aligned} \quad (12)$$

where δ_{sr} is the Kronecker delta and λ_r is the r th eigenvalue for the free undamped vibrations of a string with the same boundary conditions.

Equation (10) can be written as follows:

$$\begin{aligned} M\ddot{q}_r(t) + C\dot{q}_r(t) + Kq_r(t) \\ \times \left[k \left[\sum_{r=1}^N \phi_r(d) q_r(t) - \bar{u}(t) \right]^3 \right. \end{aligned}$$

$$\begin{aligned} &+ \sigma \left(\sum_{r=1}^N \phi_r(d) \dot{q}_r(t) + \gamma \sum_{r=1}^N \phi'_r(d) q_r(t) - \dot{\bar{u}}(t) \right) \Big] \\ &\times \phi_r(d) = 0, \end{aligned} \quad (13a)$$

$$\begin{aligned} \ddot{\bar{u}}(t) + k \left[\bar{u}(t) - \sum_{r=1}^N \phi_r(d) q_r(t) \right]^3 \\ + \sigma \left(\dot{\bar{u}}(t) - \sum_{r=1}^N \phi_r(d) \dot{q}_r(t) - \gamma \sum_{r=1}^N \phi'_r(d) q_r(t) \right) = 0, \end{aligned} \quad (13b)$$

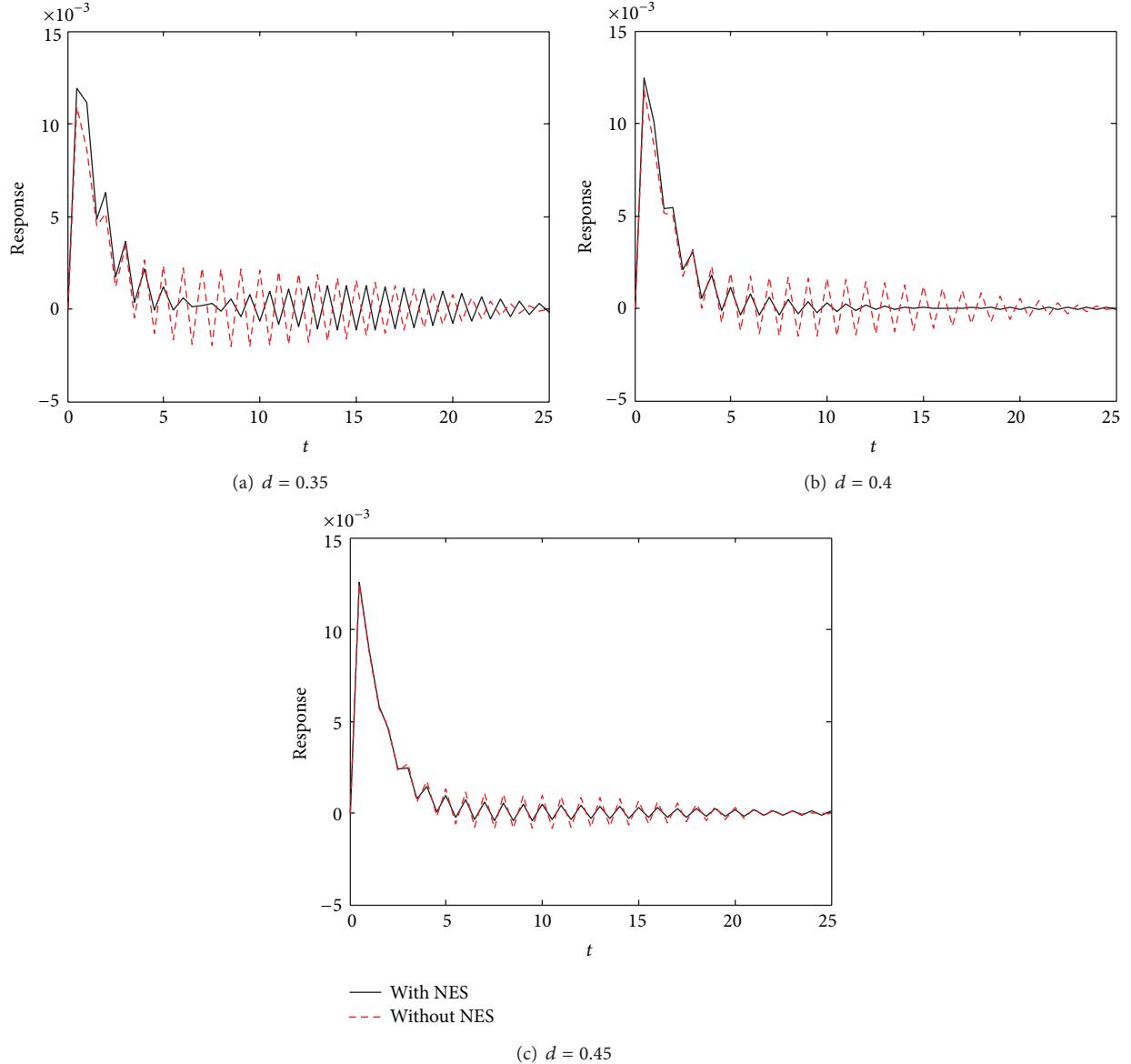


FIGURE 4: Comparison of the transient response of axially moving string with and without NES under varying attaching location (solid line: string coupled with NES; dotted line: string without NES).

where

$$\begin{aligned}
 M &= \delta_{sr}, \\
 C_1 &= 2vb_{sr} + \eta_0\delta_{sr} + 0.5\alpha_0a\lambda\delta_{sr}, & C_2 &= 0.5\alpha_0 \frac{b}{\lambda}e_{s1}, \\
 C_3 &= 0.5\alpha_0 \frac{b}{\lambda}e_{s2}, & C_4 &= 0.5\alpha_0 \frac{b}{\lambda}f_{s1}, & C_5 &= 0.5\alpha_0 \frac{b}{\lambda}f_{s2}, \\
 K &= (\nu^2 - 1)C_{sr} + \eta_0vb_{sr} = \omega_r^2,
 \end{aligned} \tag{14}$$

where M , C , and K are the mass, damping, and stiffness matrices, respectively. In addition, ω_r is the r th natural frequency of the axially moving string.

4. Effectiveness of the NES

The effectiveness of the NES coupled to an axially moving string at varying axial speeds is demonstrated. Equations (13a) and (13b) are a high-dimensional nonlinear dynamical system; thus, numerical methods should be used to truncate the expansion (13a), (13b) to a finite number of modes. In the Galerkin procedure for gyroscopic systems, at least two modes of the displacement amplitude can yield a good approximation [17–19]. Therefore, by choosing $N = 2$, the equations can achieve good numerical convergence. To initiate oscillations, an initial distributed velocity is imposed as follows:

$$\begin{aligned}
 \dot{q}_1(0) &= X, \\
 q_r(0) = \dot{q}_2(0) = \dots = \dot{q}_r &= \bar{u}(0) = \dot{\bar{u}}(0) = 0.
 \end{aligned} \tag{15}$$

Figure 2 shows the transient responses of the axially moving string and the NES for $\sigma = 0.65$, $k = 80000$, $d = 0.4$, $\alpha = 0.001$, $\varepsilon = 0.06$, $X = 0.15$, $\eta_0 = 0.005$, $a = 0.2$, $b = -0.2$, $p_0 = 1.293$, and $A = 2.7745 \times 10^{-4}$. As shown in the figure, the effectiveness of NES for the axial speed v varying from 0 to 0.4 is examined. Solid and dashed lines denote the responses of the beam and the NES, respectively. Figures 2(a)–2(c) show that the amplitude of the NES is markedly higher than that of the string, indicating the occurrence of energy transfer from the string to the NES. Both the string and the NES perform decaying vibration attributed to damping dissipation. During this process, the vibrational energy is irreversibly transferred and eventually damped by the NES. The NES can effectively absorb the vibrational energy and prevent the string from excessive vibrations at varying axial speeds. Energy absorption is achieved over a wide range of axial speeds.

To further demonstrate the effectiveness of the NES, the responses of the string with the NES and without the NES are presented in Figures 3(a)–3(c) for $\sigma = 0.65$, $k = 80000$, $d = 0.4$, $\alpha = 0.001$, $\varepsilon = 0.06$, $X = 0.15$, $\eta_0 = 0.005$, $a = 0.2$, $b = -0.2$, $p_0 = 1.293$, and $A = 2.7745 \times 10^{-4}$. The string attached with the NES exhibits a drastic reduction in transient response, as indicated by solid lines. By contrast, the string without the NES slowly decays as time increases, as indicated by dashed lines. Therefore, the NES can robustly absorb vibrational energy over a broad range of axial speeds.

In Figure 4, the effect of adding position to the NES is examined. Figures 4(a)–4(c) for $\sigma = 0.65$, $k = 80000$, $v = 0.2$, $\alpha = 0.001$, $\varepsilon = 0.06$, $X = 0.15$, $\eta_0 = 0.005$, $a = 0.2$, $b = -0.2$, $p_0 = 1.293$, and $A = 2.7745 \times 10^{-4}$ illustrate the different effects of varying the position of the NES on absorbed vibration. When the position d is 0.4, the maximum effect of vibration suppression is achieved.

5. Conclusions

The vibration of an axially moving string with transverse wind loads is effectively suppressed using the NES. The results of the simulation experiments indicate that at various flow speeds, the NES can irreversibly transfer and dissipate vibrational energy from the axially moving string. By considering the adding internal degrees of freedom to the NES, vibration suppression is most clearly demonstrated at $d = 0.4$. Therefore, the vibration of an axially moving string can be suppressed based on the additional internal degrees of freedom to NES and the speed of the axially moving string.

Acknowledgment

The authors acknowledge the funding support of the Natural Science Foundation of Liaoning Province (201102170).

References

- [1] H. R. Öz, M. Pakdemirli, and E. Özkaya, “Transition behaviour from string to beam for an axially accelerating material,” *Journal of Sound and Vibration*, vol. 215, no. 3, pp. 571–576, 1998.
- [2] J. A. Wickert and C. D. Mote Jr., “Classical vibration analysis of axially moving continua,” *Journal of Applied Mechanics*, vol. 57, no. 3, pp. 738–744, 1990.
- [3] T.-Z. Yang, B. Fang, Y. Chen, and Y.-X. Zhen, “Approximate solutions of axially moving viscoelastic beams subject to multi-frequency excitations,” *International Journal of Non-Linear Mechanics*, vol. 44, no. 2, pp. 240–248, 2009.
- [4] H. Ding and L.-Q. Chen, “Natural frequencies of nonlinear vibration of axially moving beams,” *Nonlinear Dynamics*, vol. 63, no. 1-2, pp. 125–134, 2011.
- [5] L.-Q. Chen and X.-D. Yang, “Steady-state response of axially moving viscoelastic beams with pulsating speed: comparison of two nonlinear models,” *International Journal of Solids and Structures*, vol. 42, no. 1, pp. 37–50, 2005.
- [6] L.-Q. Chen and Y.-Q. Tang, “Combination and principal parametric resonances of axially accelerating viscoelastic beams: recognition of longitudinally varying tensions,” *Journal of Sound and Vibration*, vol. 330, no. 23, pp. 5598–5614, 2011.
- [7] L.-F. Lu, Y.-F. Wang, and Y.-X. Liu, “Axially moving strings with transverse wind excitations,” *Gongcheng Lixue/Engineering Mechanics*, vol. 25, no. 2, pp. 40–45, 2008.
- [8] R. Viguié, G. Kerschen, J.-C. Golinval et al., “Using passive nonlinear targeted energy transfer to stabilize drill-string systems,” *Mechanical Systems and Signal Processing*, vol. 23, no. 1, pp. 148–169, 2009.
- [9] F. Georgiades and A. F. Vakakis, “Dynamics of a linear beam with an attached local nonlinear energy sink,” *Communications in Nonlinear Science and Numerical Simulation*, vol. 12, no. 5, pp. 643–651, 2007.
- [10] P. Panagopoulos, F. Georgiades, S. Tsakirtzis, A. F. Vakakis, and L. A. Bergman, “Multi-scaled analysis of the damped dynamics of an elastic rod with an essentially nonlinear end attachment,” *International Journal of Solids and Structures*, vol. 44, no. 18-19, pp. 6256–6278, 2007.
- [11] F. Georgiades and A. F. Vakakis, “Passive targeted energy transfers and strong modal interactions in the dynamics of a thin plate with strongly nonlinear attachments,” *International Journal of Solids and Structures*, vol. 46, no. 11-12, pp. 2330–2353, 2009.
- [12] Y. S. Lee, A. F. Vakakis, and D. M. McFarland, “Suppressing aerelastic instability using broadband passive targeted energy transfers—part I: theory,” *AIAA Journal*, vol. 45, no. 3, pp. 693–711, 2007.
- [13] F. Nucera, F. Lo Iacono, D. M. McFarland, L. A. Bergman, and A. F. Vakakis, “Application of broadband nonlinear targeted energy transfers for seismic mitigation of a shear frame: experimental results,” *Journal of Sound and Vibration*, vol. 313, no. 1-2, pp. 57–76, 2008.
- [14] A. Ture Savadkoohi, B. Vaurigaud, C.-H. Lamarque, and S. Pernot, “Targeted energy transfer with parallel nonlinear energy sinks—part II: theory and experiments,” *Nonlinear Dynamics*, vol. 67, no. 1, pp. 37–46, 2012.
- [15] Y. Wang, X. Liu, and L. Huang, “Stability analyses for axially moving strings in nonlinear free and aerodynamically excited vibrations,” *Chaos, Solitons and Fractals*, vol. 38, no. 2, pp. 421–429, 2008.
- [16] W. D. Zhu and C. D. Mote Jr., “Free and forced response of an axially moving string transporting a damped linear oscillator,” *Journal of Sound and Vibration*, vol. 177, no. 5, pp. 591–610, 1994.
- [17] M. P. Païdoussis and C. Semler, “Nonlinear and chaotic oscillations of a constrained cantilevered pipe conveying fluid: a full

- nonlinear analysis," *Nonlinear Dynamics*, vol. 4, no. 6, pp. 655–670, 1993.
- [18] R. J. McDonald and N. S. Namachchivaya, "Pipes conveying pulsating fluid near a 0:1 resonance: local bifurcations," *Journal of Fluids and Structures*, vol. 21, no. 5-7, pp. 629–664, 2005.
- [19] J. D. Jin and Z. Y. Song, "Parametric resonances of supported pipes conveying pulsating fluid," *Journal of Fluids and Structures*, vol. 20, no. 6, pp. 763–783, 2005.

Research Article

Modeling and Analyzing Operational Decision-Making Synchronization of C2 Organization in Complex Environment

Zou Zhigang, Liu Fuxian, Zhang Bo, and Li Xiang

Air and Missile Defense College, Air Force Engineering University, Xi'an 710051, China

Correspondence should be addressed to Zou Zhigang; usher12@126.com

Received 5 June 2013; Revised 24 July 2013; Accepted 24 July 2013

Academic Editor: J. A. Tenreiro Machado

Copyright © 2013 Zou Zhigang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to improve capability of operational decision-making synchronization (ODMS) in command and control (C2) organization, the paper puts forward that ODMS is the negotiation process of situation cognition with three phases about "situation cognition, situation interaction and decision-making synchronization" in complex environment, and then the model and strategies of ODMS are given in quantity. Firstly, measure indexes of three steps above are given in the paper based on the time consumed in negotiation, and three patterns are proposed for negotiating timely in high quality during situation interaction. Secondly, the ODMS model with two stages in continuous changing situation is put forward in the paper, and ODMS strategies are analyzed within environment influence and time restriction. Thirdly, simulation cases are given to validate the process of ODMS under different continuous changing situations the results of this model are better than the other previous models to fulfill the actual restrictions, and the process of ODMS can be adjusted more reasonable for improving the capability of ODMS. Then we discuss the case and summarize the influence factors of ODMS in the C2 organization as organization structure, shared information resources, negotiation patterns, and allocation of decision rights.

1. Introduction

As remarkable characteristics of joint operation under information condition, systems combat demands that operational entities should form the coincident combat intention within networked operational systems in cognition domain, and then the entities can be driven to take cooperative action in physical domain, which makes operation synchronization realized to improve combat effectiveness of weapons systems in the furthest extent [1]. So study in operation synchronization has great significance on combat effectiveness of networked operational systems.

In the concept of network centric warfare (NCW), operation synchronization is mainly referred to self-synchronization [2]. As the development of NCW, the concept of network centric operations (NCO) is paid more attention to decision-making synchronization and action synchronization. Based on situation awareness within complex environment in information domain, situation cognition difference can be formed within different operation entities, respectively. Then certain regulations or strategies should be followed by each operation

entity to realize situation interaction in high quality, which can achieve operational decision-making synchronization (ODMS) with unanimous situation cognition [3]. Then different operational entities are driven to take cooperative action within operational systems in physical domain. So in order to study operation synchronization in quantity, the primary question about how to model and improve ODMS should be solved firstly.

At present, the qualitative expatiation about ODMS is the essential requirement for networked systems combat in the future is put forward in [2, 3]. In order to study ODMS in quantity, ODMS is taken into account as complex activity in cognition domain within the command and control (C2) organization, and then some measure indexes and models of ODMS are given on the basis of C2 organization with mainly three methods as follows.

The first method is based on opinion emergence of group communication [4], ODMS is regarded as opinion exchange process on the basis of group communication in [5, 6]. There are two shortcomings as follows: (1) communication is the process of exchange opinion among individuals but not

always obtain the consensus in the whole organization [7]. The illogical strategies and ambiguous indexes of communication are difficult to be used from the unanimous opinions directly, but mostly in oppose opinions, multiple opinions, and so on. (2) In actual combat process, operation plans can be formed by operation entities with situation cognition, respectively, and then ODMS can be realized by operation plan-based negotiation among operation entities, which is named as situation interaction [8]. However, situation cognition is abstract and time variable in cognition domain. Presently as improvement on “three-layer” model of situation awareness [9, 10], the researches on situation cognition, such as individual situation awareness [11], and situation awareness velocity [12], are deficient in analyzing situation cognition formation, and the research about influence of complex changing environment on the ODMS is still in blank.

The second method is with the macroscopic viewpoint of difference in situation cognition, models about cognitive-entropy (CE) and cognitive self-synchronization (CSSync) are put forward by Manso and Moffat in the American department of defense (DoD) [13], and the compute method is given on the basis of the Experimental Laboratory for Investigating Collaboration, Information-sharing and Trust (ELICIT) [14]. The models are also not referred to situation cognition formation, especially cognition interaction process, and the relation between the CE and information entropy.

The third method is that ODMS model can be constructed as network dynamic as physics systems with the microscopic viewpoint of individual cognition. Dekker in Australian national defense research centre considers each operation entity as a coupled oscillator with different phase angles [15]. As an abstract model for synchronization in C2 organization, Kuramoto model is extended to describing the ODMS [16]. Information flow (including sensor information, operation scheme, and so on) in the organization is synthesized as a phase angle, and information difference is considered as feedback of phase angle; the synchronization process can be controlled by adjusting coupled coefficients. Although the model can be used to depict the synchronization process, it excessively depends on the network synchronization in physics, which is hardly applied to actual military operation.

The ODMS demands situation interaction among operation entities to achieve unanimous situation cognition, and situation interaction can be adjusted with different patterns agilely but only certain coefficients. Therefore, combining macroscopic view and microscopic view, this paper presents that ODMS is a negotiation process of situation cognition interaction in complex environment, which contains three phases about “situation cognition, situation interaction and decision-making synchronization.”

Taking time consumed into account, we construct the ODMS model with measure indexes of three phases in the paper. Then ODMS strategies are analyzed to adapt the environment influence and time restriction. Finally, the process of ODMS is verified in simulation cases in complex environment, including continuous changing situation, environment influence, and time restriction, which reflects that the model and strategies are feasible and available. On the basis of this, via analyzing and discussing the simulation

results, the ODMS inhibitors and enablers are summarized in the paper.

The paper is organized as follows: measure indexes of ODMS in complex environment are given in Section 2; ODMS model is constructed in Section 3, and ODMS strategies within the environment influence and time restriction are put forward in Section 4. And the simulation cases are designed to validate the model and strategies in Section 5. After comparing and discussing the simulation results, influence factors of ODMS are given in Section 6, and the conclusion and further work are given in Section 7.

2. Measure Indexes of ODMS in Complex Environment

The influence of complex environment on ODMS mainly contains two aspects: continuous changing situation and environment influence on communication. According to the process of ODMS with three phases in Figure 1, time-oriented measure indexes in each phase are given combining topological structure of C2 organization in the paper; especially different situation changing patterns and negotiation patterns are put forward in situation interaction stage.

2.1. Indexes of Continuous Changing Situation. Situation change can be reflected as the number change of operation entities with updating situation information rapidly. If situation changes quickly, it will arouse the action of updating situation information within sufficient limit time. So situation information elements apperceived item by item in operational entities can be regarded as situation information flow and depicted as continuous changing functions.

Definition 1. Situation change rate (SCR): it is the ratio of the number of operational entities with updating situation information, brought by the situation change directly in unit time, to the total number of operational entities in C2 organization, marked as h'_t in formula (1), which can be used to reflect of situation change:

$$h'_t = \frac{n_t}{N}, \quad (1)$$

where $n_t \in [0, N]$ denotes that number of operational entities with updating situation information in the moment of t , N is the total number of operational entities in the C2 organization, and $h'_t \in [0, 1]$ shows the SCR in the time of t . Combined [12], three patterns of continuous changing situation are put forward in Figure 2, whose general expressions are given in Table 1.

In Table 1, $a \in [0, 1]$, $b > 0$, and if $a = 0$, then the 1st pattern $h'_t = 0$ reflects that there is no operational entity updating the situation information in the time of t , which shows that the situation is stable (as one special case of continuous changing situation). Via compound calculation with the three patterns of continuous changing situation, more complex patterns of situations change can be generated in the paper.

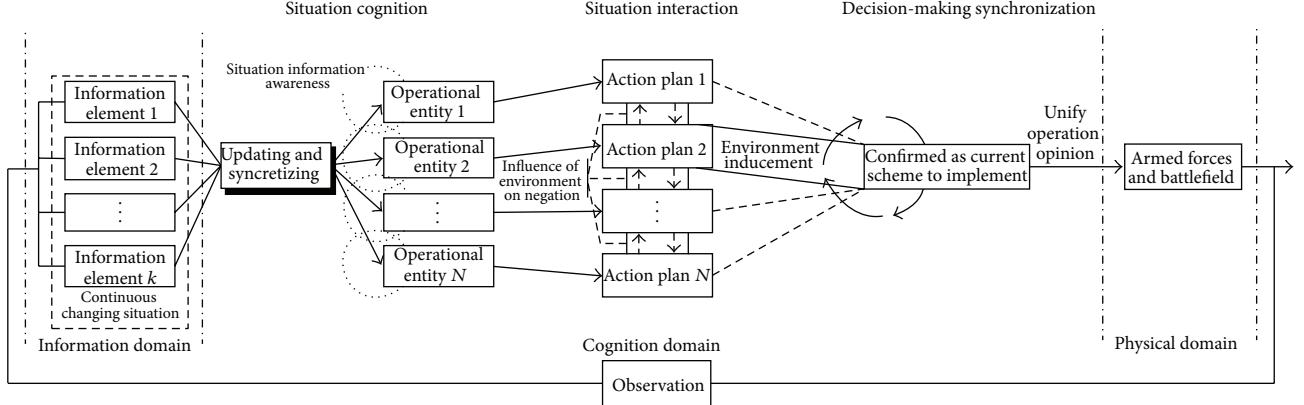


FIGURE 1: The process of ODMS in complex environment.

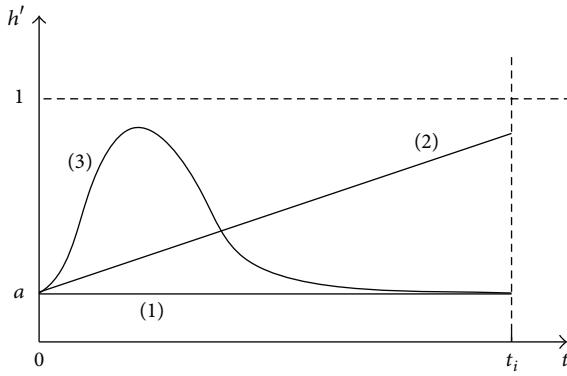


FIGURE 2: Three patterns in continuous changing situation.

2.2. Indexes of Situation Cognition. During the situation changing process, different situation cognition can be formed within different operational entities [10]. Due to the limits of time and resources, each operational entity can hardly negotiate with all the others. In the actual combat process, operational entity should negotiate with the surroundings. Taking the situation cognition level of certain operational entity (such as superior command center) as target, the others should adjust their own situation cognition via negotiating with the operational entity rapidly, and the ODMS in C2 organization can be realized when the situation cognition of others is consistent with target entity. Then indexes of situation cognition are defined as follows.

Definition 2. Situation cognition level (SCL): it is the maturity degree about operational entities in analyzing and comprehending the situation information of battlefield to form the action plans. As the improvement of discrete model in situation cognition in [17], SCL can be used to describe continuous process of individual situation cognition and marked as $x_i \in [0, 1]$. When $x_i = 0$ denotes that SCL of operational entity i is in the blank status, $x_i = 1$ shows that the SCL of operational entity i is in the complete status, and $x_i \in (0, 1)$ shows that the SCL of operational entity i is in the forming status.

TABLE 1: The SCR in three patterns.

No.	Situation changing patterns	SCR
1	Identical changing situation	$h'_t = a$
2	Gradual changing situation	$h'_t = b * t + a$
3	Sudden changing situation	$h'_t = f(t)$

Definition 3. Situation cognition target (SCT): in situation interaction process, taking the SCL of operational entity j as target, different patterns can be chosen by operational entity i to negotiate with j for improving its SCL. Thus operational entity i is called as compromise entity, and operational entity j is called as target entity, and x_j is named as SCT.

Definition 4. Information perception ability (IPA): in the process of situation cognition formation, it is used to reflect the ability of analyzing and comprehending situation information in operational entity i , marked as v_i . With the time of operational entity i analyzing the situation information increasing, the SCL can be continuous improved. The more network resources are possessed by the operational entity in the C2 organization, the higher value of its IPA [18]; and the higher network average degree of C2 organization, the stronger the degree of the whole IPA:

$$v_i = e^{\alpha} \cdot k_i^{\langle k - h'_t \rangle} t'_i, \quad (2)$$

where $k_i \in (0, 1)$ is entity importance degree (EID) of the i th object in C2 organization, t'_i is the time of situation cognition formation, and $\langle k \rangle$ denotes that network average degree of C2 organization, which is directly proportional with networked level. And $\alpha \in \{0, 1\}$ is environment inducement factor if $\alpha = 1$ shows that operational entity i is induced as target entity, and $\alpha = 0$ denotes that operational entity i can be regarded as compromise entity.

2.3. Indexes of Situation Interaction. According to measure indexes of centrality in social network [19], the situation interaction is described as action plan-based negotiation in the view of time consuming. Based on the topological

TABLE 2: The value of β in electromagnetic environment.

Electromagnetic environment classification	β
Level I (simple)	0~0.20
Level II (mild complexity)	0.20~0.55
Level III (moderate complexity)	0.55~0.80
Level IV (severe complexity)	0.80~1.0

structure of C2 organization, if entity i makes compromise with j , situation interaction route is marked as $r = \{(j, i) \mid j \rightarrow i\}$, where (j, i) denotes one segment in the situation interaction, and $j \rightarrow i$ shows that route of situation interaction is from j to i . Here we give measure indexes of situation interaction as follows.

Definition 5. Information share ability (ISA): it is information share speed among operational entities, which reflects difference of negotiation speed among the operational entities. And v_{ij} is marked as the ISA between operational entity i and j . The ISA of each entity is directly proportional to EID [20]:

$$v_{ij} = C_{ij} \cdot \frac{k_i + k_j}{2}, \quad (3)$$

where k_i and k_j are the EID of operational entity i and j separately, and $C_{ij} \in [0, 1]$ is the communication unblocked degree. And that $C_{ij} = 1 - \beta_{ij}$, $\beta_{ij} \in [0, 1]$ is environment influence factor. Then the value of β is given in Table 2 according to the classification of electromagnetic environment in [21].

Definition 6. Situation interaction time (SIT): it is the time cost in forming the unanimous opinion between operational entities i and j , which is marked as t_{ij} :

$$t_{ij} = \frac{\Delta x_{ij}}{v_{ij}} = \frac{x_i - x_j}{v_{ij}}. \quad (4)$$

Definition 7. Situation interaction quality (SIQ): as situation interaction is realized by action plan-based negotiation, compromise extent is used to measure negotiation quality in situation interaction between operational entities i and j , marked as $u_{ij} \in [0, 1]$ in formula (5). With the SIT increasing, the SIQ is improving correspondingly; when exceeding certain time, the SIQ keeps invariability in maximum, and the SCL is in the complete status:

$$\begin{aligned} u_{ij} &= 0 \quad t = 0, \\ u_{ij} &\in (0, 1) \quad 0 < t < t^*, \\ u_{ij} &= 1 \quad t^* \leq t. \end{aligned} \quad (5)$$

Combined EID and directions of situation interaction, three negotiation patterns are put forward in the paper. And the three functions of SIQ are defined as $u_{ij} = f(k_i, k_j, t)$ in Table 3.

In Figure 3, the 1st pattern shows the equality negotiation process between two operational entities with the same EID.

TABLE 3: Three SIQ functions in the direction of $j \rightarrow i$.

No.	Negotiation patterns	EID	SIQ functions
1	Equality pattern	$k_i = k_j$	$u_i = v'_{ij} t$
2	Persuading pattern	$k_i > k_j$	$u_i = (v'_{ij} t)^2$
3	Concession pattern	$k_i < k_j$	$u_i = (v'_{ij} t)^{1/2}$

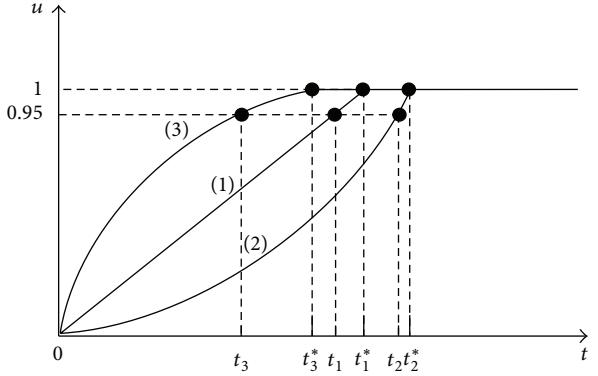


FIGURE 3: The SIQ of three negotiation patterns.

The 2nd pattern reflects that operational entity with high EID, as compromise entity, is persuaded to negotiate with target entity in low EID. The compromise entity is cautious to analyze and judge the SCL of target entity, as SCT, slowly in the beginning, and later gradually learn and understand from the target entity to keep the SCL consistent with SCT rapidly in a short time. The 3rd pattern denotes that operational entity with low EID as compromise entity makes a concession in negotiating with target entity in low EID. Then the compromise entity accepts the SCL of target entity as SCT directly without analysis and then learns SCT slowly after accepting in some extent to reach the SCL unanimous with target entity.

Suppose that ISA of three negotiation patterns in the direction of $j \rightarrow i$ is $v_{ij,1}$, $v_{ij,2}$, and $v_{ij,3}$ separately, and we can obtain that $v_{ij,2} < v_{ij,1} < v_{ij,3}$ according to formula (3). And suppose that SIT of three negotiation patterns is t_1^* , t_2^* , and t_3^* , we can gain that $t_3^* < t_1^* < t_2^*$ with formula (4). Then we can derivate relative ISA with $v'_{ij} = v_{ij}/\Delta x_{ij}$, then $t^* = 1/v'_{ij} = t_{ij}$, $0 < t \leq t^*$. And SIQ functions are general formulas.

Definition 8. Situation interaction function: it is used to describe situation interaction process between two adjacent entities. In the direction of $j \rightarrow i$, x_i can be increased by negotiating with x_j as SCT, and can be calculated in the time of t_i according to formula (6):

$$x_i(t_i) = x_i(t'_i) - u_i(t_i) \cdot [x_i(t'_i) - x_j^*], \quad (6)$$

where $x_j^* = x_j(t_j^*)$ denotes the SCL of operational entity j after situation interaction. When $u_{ij} = 0$ reflects that there is no effect in the negotiation, $u_{ij} \in (0, 1)$ shows that situation interaction is accomplished in certain extent, and $u_{ij} = 1$ shows that situation interaction is accomplished completely.

2.4. Indexes of Situation Interaction. Based on certain situation prediction, SCL of someone operational entity in advantageous environment is obtained to the complete status firstly, so the operational entity is induced to be the target entity in the C2 organization. In order to improve SCL, other entities should adopt corresponding negotiation patterns to make situation interaction in high quality quickly, which can make SCL of the whole C2 organization in complete status to reach ODMS. Relative definitions are given as follows.

Definition 9. Situation prediction level (SPL): before obtaining the situation information about opposite party, original SCL of commander can be formed with active thinking and estimation in forecast process [22], which is called SPL and marked as x_{io} . As network resources in C2 organization (such as the sensors and command systems) are possessed by operational entities with different EID separately, SPL of each operation entity can be generated in different extent before situation cognition formation, and $x_{io} \geq 0$.

Definition 10. Synchronization time (ST): for each operational entity in C2 organization, it is the maximum sum of SIT and time consumed in situation cognition formation, marked as T in the formula (7). After the time of T , the SCL of all the operational entities is in the complete status:

$$T = \max_{i \in N} \{t'_i + t_i\}. \quad (7)$$

Definition 11. Synchronization extent (SE): it is the ratio of SCL of compromise entity x_i after situation interaction and SCL of target entity x_o , and marked as γ_{io} in formula (8). Due to difference in interior condition and influence from exterior environment, ODMS is hardly accomplished completely:

$$\gamma_{io} = \frac{x_i}{x_o}, \quad (8)$$

where $\varepsilon_{io} = 1 - \gamma_{io}$ denotes accepted synchronization error. As $0 \leq x_i < x_o \leq 1$, $\gamma_{io} \in [0, 1]$, then $\varepsilon_{io} \in [0, 0.05]$ is defined in the paper to ensure basic ODMS and $\gamma_{io} \in [0.95, 1]$ in formula (8).

Definition 12. Synchronization time effectiveness (STE): it is the ratio of time consumed in situation cognition phase and situation interaction phase, and marked as τ in formula (9). STE reflects that time-consuming distribution of each phase in ODMS:

$$\tau = \frac{t_e}{t_r}, \quad (9)$$

where $\tau > 0$, $t_r = \max_{i \in N} \{t'_i\}$ and $t_e = T - t_r$ is the time consuming of situation cognition and situation interaction in the C2 organization separately.

Definition 13. Synchronization qualities (SQ): it is the ratio of SCL increment in situation interaction phase and SCL in situation cognition phase, and it is marked as χ in formula (10). So SQ reflects the process of increasing SCL in each phase of ODMS:

$$\chi = \frac{x_e}{x_r}, \quad (10)$$

where $\chi > 0$, $x_r = \sum_{i \in N} x_i(t'_i)/N$, and $x_e = \gamma - x_r$ is the average increment of SCL in phase of situation cognition and situation interaction separately.

3. ODMS Model in Complex Environment

3.1. Initialize C2 Organization. The triple set $\{G_N, K_N, X_N\}$ is used to represent C2 organization, N is the number of operational entities, G_N is the adjacency matrix of C2 organization, X_N is the SPL set of operational entities, and $K_N = \{k_i \mid i \in N\}$ is the EID set of operational entities. The EID reflects the ability of operational entity possessing the network resources, which lies in the betweenness centrality and degree centrality of the entity in the C2 organization [23], and the calculation method in [24] is adopted in the paper.

3.2. Construct Model of ODMS. Decision-making synchronization experienced two phases about situation cognition and situation prediction.

Phase 1. Situation cognition: with a certain SPL, the SCL of each operational entity is increased slowly during the process of analyzing and comprehending situation information under continuous changing situation. When SCL of certain operational entity in advantageous environment is improved to the complete status in a short time, then the entity is induced to be the target entity, and other entities are regarded as compromise entities. According to formula (2), situation cognition model is constructed as follows:

$$\begin{aligned} x_i(0) &= x_{io}, \\ \dot{x}_i(t'_i) &= v_i(t'_i) = e^\alpha k_i^{\langle k-h'(t'_i) \rangle} \cdot t'_i, \\ h'(t'_i) &= a \quad 0 \leq a \leq 1, \\ h'(t'_i) &= b \cdot t'_i + a \quad b > 0, \\ h'(t'_i) &= f(t'_i), \quad 0 \leq h'(t'_i) \leq 1, \end{aligned} \quad (11)$$

where $i \in N$, t'_i is the time of information perception in operational entity i before situation interaction, and $h'(t'_i)$ is the SCR in the time of t'_i . When $\alpha = 1$ in formula (2), the entity is regarded as target entity, and $\alpha = 0$ denotes that it is compromise entity.

Phase 2. Situation interaction: based on the SCL of target entity in the complete status firstly, compromise entities make situation interaction with the target entity by applying corresponding negotiation patterns to improve their SCL in complete status, which can achieve ODMS. According to

the formula (6), situation cognition model is constructed as follows:

$$\begin{aligned} S &= \frac{\sum_{i=1}^N \gamma_{io}}{N} \in [\gamma, 1] \\ x_i(t_i) &= x_i(t'_i) - u_1(t_i) \cdot [x_i(t'_i) - x_j^*] \quad k_{i+1} = k_i, \\ x_i(t_i) &= x_i(t'_i) - u_2(t_i) \cdot [x_i(t'_i) - x_j^*] \quad k_{i+1} = k_i, \quad (12) \\ x_i(t_i) &= x_i(t'_i) - u_3(t_i) \cdot [x_i(t'_i) - x_j^*] \quad k_{i+1} < k_i, \end{aligned}$$

$$\begin{aligned} 0.95 &\leq \gamma \leq 1, \\ (j, i) &\in \Lambda = \{(m, n) \mid m \rightarrow n, m \in P, n \in P\}, \end{aligned}$$

where $P = \{1, \dots, N\}$ is the set of operational entities, $u_k(t_i)$ ($k \in \{1, 2, 3\}$) is the SIQ of entity j applying the k th negotiation pattern, and $0 < t_i \leq t_i^*$. When $u_i \equiv 1$ shows the complete synchronization process and comprehensive synchronization extent as $S \equiv 1$ in formula (12) $\gamma \leq u_i \leq 1$ denotes the incomplete synchronization process, and $S \in [0.95, 1]$.

4. ODMS Strategies within Environment and Time Restriction

Due to influence of the environment in battlefield, such as complex environment and operation timerestrictions [21], information sharing among the operational entities in situation interaction is always incomplete, which leads to the incomplete ODMS. We suppose that situation interaction route in C2 organization is marked as $r = \{(p_i, p_{i+1}) \mid p_i \rightarrow p_{i+1}, p_i \in G_N, i \in [1, N-1]\}$. Taken situation interaction route $j \rightarrow k \rightarrow l \dots m$, contains n situation interaction segments, as an example, the properties of ODMS are analyzed as follows in detail.

4.1. ODMS Strategies within Environment Influence

Property 1 (environment adaptability in choosing situation interaction route). Suppose that there are s situation interaction routes existing from entity j to entity m : $r_{j \rightarrow m}^s = \{(p_i, p_{i+1}) \mid p_i \rightarrow p_{i+1}; p_i, p_{i+1} \in G_N \wedge (p_i^q = j, p_{i+1}^z = m)\}$. As differences in environment influence, it leads to difference in communication unblocked degree among different interaction routes. In order to realize unblocked interaction route by any possibility, environmental adaptability operator (EAO) is designed below to choose the best situation interaction route r_{s^*} with the minimum value, which is in favor of adapting the environment change:

$$\Delta_{r|j \rightarrow m}^{s^*} = \min \left[\left(\sum_{i \in n} \beta_{i,i+1} - \prod_{i \in n} \beta_{i,i+1} \right) \cdot \prod_{i \in n} (k_i + k_{i+1}) \right]. \quad (13)$$

And $\beta_{i,i+1}$ is the communication unblocked degree within the situation interaction segment $p_i \rightarrow p_{i+1}$.

Proof. According to recurrence method, the case of interaction situation route with two segments ($n_1 = n_2 = 2$) should be proved firstly. Suppose that there are two situation interaction routes:

$$\begin{aligned} (1) \quad r_{j \rightarrow m}^1 &= \{(p_i, p_{i+1}) \mid j \rightarrow l_1 \rightarrow m, i \in \{1, 2\}\}, \\ (2) \quad r_{j \rightarrow m}^2 &= \{(p_i, p_{i+1}) \mid j \rightarrow l_2 \rightarrow m, i \in \{1, 2\}\}. \end{aligned} \quad (14)$$

For the 1st situation interaction route, each interaction segment with formula (3):

$$\begin{aligned} v_{jl_1} &= C_{jl_1} \cdot \frac{k_j + k_{l_1}}{2} = (1 - \beta_{jl_1}) \cdot \frac{k_j + k_{l_1}}{2}, \\ v_{l_1 m} &= C_{l_1 m} \cdot \frac{k_{l_1} + k_m}{2} = (1 - \beta_{l_1 m}) \cdot \frac{k_{l_1} + k_m}{2}, \\ t_{jm}^1 &= \frac{x_j^* - x_{l_1}}{v_{jl_1}} + \frac{x_{l_1}^* - x_m}{v_{l_1 m}} \\ &= 2 \left\{ \left[(x_j^* - x_{l_1})(k_{l_1} + k_m) + (x_{l_1}^* - x_m)(k_j + k_{l_1}) \right] \right. \\ &\quad \left. - [\beta_{l_1 m} \cdot (k_{l_1} + k_m)(x_j^* - x_{l_1}) \right. \\ &\quad \left. + \beta_{jl_1} \cdot (k_j + k_{l_1})(x_{l_1}^* - x_m)] \right. \\ &\quad \times \left. ((k_j + k_{l_1}) \cdot (k_{l_1} + k_m) \right. \\ &\quad \left. - (\beta_{jl_1} + \beta_{l_1 m} - \beta_{jl_1} \beta_{l_1 m})) \right. \\ &\quad \left. \cdot (k_j + k_{l_1}) \cdot (k_{l_1} + k_m))^{-1} \right\}. \end{aligned} \quad (15)$$

In situation cognition stage, due to interaction position of entity l_1 and entity l_2 is in the same level, and then $x_{l_1} \approx x_{l_2}$. With the target entity in situation interaction process as $x_j^* = x_{l_1}^* = 1$, $(x_j^* - x_{l_1})(k_{l_1} + k_m)$ and $(x_{l_1}^* - x_m)(k_j + k_{l_1})$ are both fixed value. As $(x_j^* - x_{l_1}) \in [0, 1]$ and $(k_{l_1} + k_m) \in [0, 2]$, then $(k_{l_1} + k_m)(x_j^* - x_{l_1}) \in [0, 2]$, $(k_j + k_{l_1}) \cdot (k_{l_1} + k_m) \in [0, 4]$, and $(k_{l_1} + k_m)(x_j^* - x_{l_1}) \subset (k_j + k_{l_1}) \cdot (k_{l_1} + k_m)$. And $(\beta_{jl_1} + \beta_{l_1 m} - \beta_{jl_1} \beta_{l_1 m}) > \beta_{l_1 m}$, so the factor as below plays decisive role in denominator:

$$\Delta_{r|j \rightarrow m}^1 = (\beta_{jl_1} + \beta_{l_1 m} - \beta_{jl_1} \beta_{l_1 m}) \cdot (k_j + k_{l_1}) \cdot (k_{l_1} + k_m), \quad (16)$$

which is regarded as the environmental adaptability operator, and $\beta \rightarrow 0$, and then

$$\begin{aligned} t_{jm}^1|_{\beta \rightarrow 0} &= 2 \left[((x_j^* - x_{l_1})(k_{l_1} + k_m) \right. \\ &\quad \left. + (x_{l_1}^* - x_m)(k_j + k_{l_1})) \right. \\ &\quad \times \left. ((k_j + k_{l_1}) \cdot (k_{l_1} + k_m))^{-1} \right]. \end{aligned} \quad (17)$$

With the same method, for the 2nd situation interaction route, the EAO is

$$\Delta_{r|j \rightarrow m}^2 = (\beta_{jl_2} + \beta_{l_2m} - \beta_{jl_2}\beta_{l_2m}) \cdot (k_j + k_{l_2}) \cdot (k_{l_2} + k_m). \quad (18)$$

Due to $\beta, k \in [0, 1]$, we obtain that $\Delta_r^1, \Delta_r^2 > 0$. If $\Delta_r^1 < \Delta_r^2$, the 1st situation interaction route is chosen; otherwise, we should choose the 2nd situation interaction route.

When there are several situation interaction routes ($n_1 \geq 2, n_2 \geq 2, \dots, n_s \geq 2$), the EAO of each situation interaction route is $\Delta_{r|j \rightarrow m}^{s_h} = (\sum_{i \in n_j} \beta_{i,i+1} - \prod_{i \in n_j} \beta_{i,i+1}) \cdot \prod_{i \in n_j} (k_i + k_{i+1}), s_h \in \{1, 2, \dots, s\}$.

The best situation interaction route should be chosen with the EAO with minimum value:

$$\begin{aligned} \Delta_{r|j \rightarrow m}^{s^*} &= \min \Delta_{r|j \rightarrow m}^{s_h} \\ &= \min \left[\left(\sum_{i \in n_j} \beta_{i,i+1} - \prod_{i \in n_j} \beta_{i,i+1} \right) \cdot \prod_{i \in n_j} (k_i + k_{i+1}) \right]. \end{aligned} \quad (19)$$

□

Inference 1. Suppose that there are situation interaction routes in two types: merely-single interaction segment ($n_1 = 1$) and multiple interaction segments ($n_2 \geq 2$), and we should take the merely-single interaction segment as the situation interaction route.

Proof. Suppose that the situation interaction route with merely-single interaction segment ($n_1 = 1$) is $r_{j \rightarrow m}^1 = \{(p_1, p_2) \mid j \rightarrow m\}$, and then $\Delta_{r|j \rightarrow m}^1 = (\beta_{1,2} - 0) \cdot (k_1 + k_2) = \beta_{1,2} \cdot (k_1 + k_2)$.

And situation interaction route with multiple interaction segments:

$$\begin{aligned} r_{j \rightarrow m}^2 &= \{(p_i, p_{i+1}) \mid j \rightarrow l_2^1 \rightarrow \dots \rightarrow l_2^q \rightarrow m\}, \\ \Delta_{r|j \rightarrow m}^2 &= \left(\sum_{i \in n_2} \beta_{i,i+1} - \prod_{i \in n_2} \beta_{i,i+1} \right) \cdot \prod_{i \in n_2} (k_i + k_{i+1}). \end{aligned} \quad (20)$$

Due to $\sum_{i \in n_2} \beta_{i,i+1} - \prod_{i \in n_2} \beta_{i,i+1} > \beta_{1,2}$, $\prod_{i \in n_2} (k_i + k_{i+1}) > (k_1 + k_2)$, and then $\Delta_{r|j \rightarrow m}^1 < \Delta_{r|j \rightarrow m}^2$. According to Property 1, we can obtain that $\Delta_{r|j \rightarrow m}^{s^*} = \min(\Delta_{r|j \rightarrow m}^1, \Delta_{r|j \rightarrow m}^2) = \Delta_{r|j \rightarrow m}^1$. So it is appropriate to take the merely-single interaction segment as the situation interaction route. □

Property 2 (multitarget entities adjustability in situation interaction routes). Suppose that there are several situation interaction routes existing from s target entities to entity m , and it is described as $r_{J \rightarrow m}^s = \{(p_i, p_{i+1}) \mid p_i \rightarrow p_{i+1}; p_i, p_{i+1} \in G_N \wedge (p_i^q \in J = \{j_1, \dots, j_s\}, p_{i+1}^z = m)\}$. As difference in environment influence, environmental adaptability

operator (EAO) is designed below to adjust, and we should choose the best situation interaction route with the minimum value of EAO:

$$\Delta_{r|j^* \rightarrow m}^{s^*} = \min \Delta_{r|j_h \rightarrow m}^{s_h} (s_h \in \{1, 2, \dots, s\}). \quad (21)$$

Then r_{s^*} is chosen as situation interaction route, which reflects that plan-based negotiation is from the target entity j_h to comprise entity m , and other target entities have no effect on the negotiation. So multitarget entities can be adjustable in situation interaction routes.

Proof. According to recurrence method, the case of interaction situation route with two target entities ($s = 2$) should be proved firstly. Suppose that situation interaction routes with two interaction segments ($n_1 = n_2 = 2$) are given by

$$\begin{aligned} (1) \quad r_{j_1 \rightarrow m}^1 &= \{(p_i, p_{i+1}) \mid j_1 \rightarrow l_1 \rightarrow m, i \in \{1, 2\}\}, \\ (2) \quad r_{j_2 \rightarrow m}^2 &= \{(p_i, p_{i+1}) \mid j_2 \rightarrow l_2 \rightarrow m, i \in \{1, 2\}\}. \end{aligned} \quad (22)$$

And each EAO is

$$\begin{aligned} \Delta_{r|j_1 \rightarrow m}^1 &= (\beta_{j_1 l_1} + \beta_{l_1 m} - \beta_{j_1 l_1}\beta_{l_1 m})(k_{j_1} + k_{l_1})(k_{l_1} + k_m), \\ \Delta_{r|j_2 \rightarrow m}^2 &= (\beta_{j_2 l_2} + \beta_{l_2 m} - \beta_{j_2 l_2}\beta_{l_2 m})(k_{j_2} + k_{l_2})(k_{l_2} + k_m). \end{aligned} \quad (23)$$

When $\Delta_{r|j_1 \rightarrow m}^1 < \Delta_{r|j_2 \rightarrow m}^2$, we should choose the 1st situation interaction route from target entity j_1 to compromise entity m , and target entity j_2 is free to search other compromise entities, which reflects the cooperation among the different target entities to improve situation interaction in high efficiency.

Similarly, for the multitarget entities ($s > 2$) and each route with multiinteraction segments ($n_1 > 2, \dots, n_s > 2$), we can obtain that

$$\begin{aligned} \Delta_{r|j^* \rightarrow m}^{s^*} &= \min \Delta_{r|j_h \rightarrow m}^{s_h} \\ &= \min \left[\left(\sum_{i \in n_j} \beta_{i,i+1} - \prod_{i \in n_j} \beta_{i,i+1} \right) \cdot \prod_{i \in n_j} (k_i + k_{i+1}) \right]. \end{aligned} \quad (24)$$

Then we can choose the target entity j^* in optimization to interact with compromise entity m , and other target entities J/j^* are used to negotiate with other compromise entities, which are using resources efficiently to realize ODMS. □

Inference 2. Suppose that there are situation interaction routes from s target entities to comprise entity m with two types: merely-single interaction segment ($n_1 = 1$) and multiple interaction segments ($n_2 \geq 2$), and we also should take the merely-single interaction segment as the situation interaction route.

Proof. Suppose that situation interaction route with merely-single interaction segment ($n_1 = 1$) is $r_{j_1 \rightarrow m}^1 = \{(p_1, p_2) \mid j_1 \rightarrow m\}$, and then $\Delta_{r|j_1 \rightarrow m}^1 = (\beta_{1,2} - 0) \cdot (k_1 + k_2) = \beta_{1,2} \cdot (k_1 + k_2)$.

From target entities j_k ($k \in \{2, 3, \dots, s\}$) to entity m , situation interaction routes with several interaction segments ($n_k \geq 2$) are given by $r_{j_k \rightarrow m}^k = \{(p_i, p_{i+1}) \mid j_k \rightarrow l_{k,2}^1 \rightarrow \dots \rightarrow l_{k,2}^q \rightarrow m\}$. Then we can gain that $\Delta_{r|j_k \rightarrow m}^k = \min[(\sum_{i \in n_k} \beta_{i,i+1} - \prod_{i \in n_k} \beta_{i,i+1}) \cdot \sum_{i \in n_k} (k_i + k_{i+1})]$.

Within the same environment influence as $\beta_{1,2} = \beta_{j_1,m} \approx \beta_{i,i+1}$ (for all $i \in n_k$), then we can obtain $\sum_{i \in n_k} \beta_{i,i+1} - \prod_{i \in n_k} \beta_{i,i+1} > \beta_{1,2}$. So $\prod_{i \in n_k} (k_i + k_{i+1}) > (k_1 + k_2)$, $\Delta_{r|j_1 \rightarrow m}^1 < \Delta_{r|j_k \rightarrow m}^k$. According to Property 2, we can deduce that $\Delta_{r|j^* \rightarrow m}^{s^*} = \min \Delta_{r|j_h \rightarrow m}^{s_h} = \Delta_{r|j_1 \rightarrow m}^1$. And the situation interaction route with merely-single interaction segment is suitable to be adopted. \square

In order to adapt the influence of environment change, combining Properties 1 and 2, strategies below should be taken to find the situation interaction route quickly, which can improve ODMS capability.

- (1) In the case target entity and compromise entity are both fixed, according to Property 1 and Inference 1, and when β is low, we should choose k_j , namely, EID, with high value in the situation interaction route, which can reduce the interaction limit time $t_{jm}^{s^*}|_{\beta \rightarrow 0}$. Otherwise, when β is in high value, choosing the entity with k_j in high value will cause communication blocked severely, so we should choose entity with k_j in low value, and then $\Delta_{r|j \rightarrow m}^{s^*}$ is also reduced, which can make $t_{jm}^{s^*}|_{\beta > 0}$ depress.
- (2) In the case about target entities are uncertainty but compromise entities are fixed, according to Property 2 and Inference 2, we should choose situation interaction route with merely-single interaction segment in possible, which can be used to confirm the suitable target entities and form the situation interaction route in the whole C2 organization.

Thus ODMS strategies set of choosing situation interaction routes under environment influence is:

$$R = \left\{ r_{st_i \rightarrow en_i}^{s_i^*} \mid st_i \vee en_i \in G_N, i \in N \right\}, \quad (25)$$

where s_i^* is the situation interaction route in optimization within $st_i \rightarrow en_i$.

4.2. ODMS Strategies within Time Restriction. In the actual operation, in order to ensure the cooperation action among operational entities, ODMS is usually incomplete within the time restriction, and there are certain accepted synchronization errors in ODMS. We can summarize properties as follows.

Property 3 (error accumulation in incomplete ODMS). Due to that $\gamma \leq u_i \leq 1$ and the SIQ of each situation interaction segment is in little difference, so $u(t_k) \approx u(t_l) \cdots u(t_m)$, which are marked as u in unified. Accepted synchronization error in the interaction segment $m \rightarrow j$, in which entity j is taken as target entity, is approximate described as

$$\begin{aligned} \Delta x_{jm} &\approx (1 - u^n) \cdot (x_j^* - x_k(t_k')) \\ &\approx (1 - u(t_k)) \cdot u(t_l) \cdots u(t_m)) \\ &\quad \cdot (x_j^* - x_k(t_k')), \end{aligned} \quad (26)$$

where $u(t_k)$ is the SIQ of entity k .

Proof. According to recurrence method, the process of the interaction segment of $j \rightarrow k$ can be described as $x_k^* = x_k(t_k') - u(t_k) \cdot [x_k(t_k') - x_j^*]$. Then the local synchronization error is

$$\Delta x_{jk} = x_j^* - x_k^* = (1 - u(t_k)) \cdot (x_j^* - x_k(t_k')). \quad (27)$$

In the situation interaction segment $k \rightarrow l$, we can obtain that

$$\begin{aligned} x_l^* &= x_l(t_l') - u(t_l) \cdot [x_l(t_l') - x_k^*] \\ &= (1 - u(t_l)) x_l(t_l') \\ &\quad + u(t_l) (1 - u(t_k)) \cdot x_k(t_k') + u(t_l) u(t_k) \cdot x_j^*. \end{aligned} \quad (28)$$

As $u(t_l) \approx u(t_k)$ and both of them are marked as u , then

$$\begin{aligned} \Delta x_{jl} &= x_j^* - x_l^* \approx (1 - u) \cdot (x_j^* - x_l(t_l')) \\ &\quad + u (1 - u) \cdot (x_j^* - x_k(t_k')). \end{aligned} \quad (29)$$

Due to $x_l(t_l') \approx x_k(t_k')$, we can gain that $\Delta x_{jl} = (1 - u^2) \cdot (x_j^* - x_k(t_k')) \approx (1 - u(t_k)u(t_l)) \cdot (x_j^* - x_k(t_k'))$. If $u = 1$, then $\Delta x_{jl} = 0$. Else if $0 < u < 1$, we can obtain that $0 < \Delta x_{jk} < \Delta x_{jl}$, and synchronization accumulation is enlarged through continuous situation interaction. So according to analogy analysis in the situation interaction route, we can gain Property 3. \square

Property 4 (time adjustability in incomplete ODMS). In incomplete ODMS, time adjustability of negotiation with concession pattern is the best of the three negotiation patterns. The time consumed in concession pattern is less than other two negotiation patterns with the same SIQ, which provides larger time range for subsequent interaction segments to adjust and fulfill the time restriction.

Proof. In the situation interaction segment $j \rightarrow k$, time consumed of complete ODMS with the i th negotiation pattern is t_i^* , and time consumed of incomplete ODMS ($u = b < 1$) is t_i correspondingly, then $t_i < t_i^*$ ($i \in \{1, 2, 3\}$), and $\Delta t_1 = t_1^* - t_1 = (a-b)/v'_{kl}$, $\Delta t_2 = t_2^* - t_2 = (\sqrt{a}-\sqrt{b})/v'_{kl}$, $\Delta t_3 = t_3^* - t_3 = (a^2 - b^2)/v'_{kl}$.

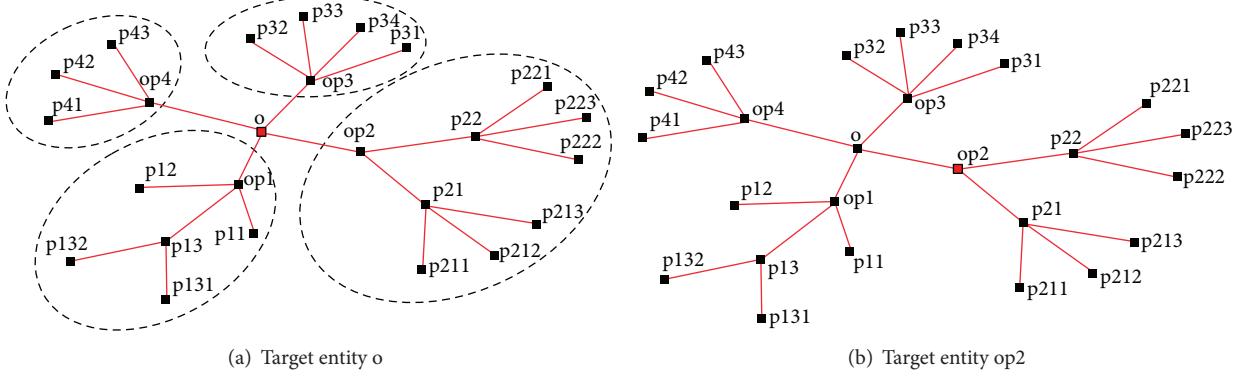


FIGURE 4: C2 organization with tree structure: NET1.

Due to $a = 1$ and $0 < b < 1$, then $a + b > 1$, $\sqrt{a} + \sqrt{b} > 1$, so $\sqrt{a} - \sqrt{b} < a - b < a^2 - b^2$. Then we can deduce that $\Delta t_3 > \Delta t_1 > \Delta t_2$. As known in Figure 3, the time t_3 consumed in concession negotiation process is much less than t_3^* in the incomplete ODMS with $u = 0.95$, which reflects the best time adjustability in three negotiation patterns. With the decrease of u , the difference of gap is enlarged in the different negotiation process. \square

Combined Properties 3 and 4, when the performance of ODMS cannot fulfill the time restriction, we should take following strategies to adjust the performance of ODMS.

- (1) Adjusting time consumed in situation interaction segment with low EID operational entities in any possible way: according to Property 3, synchronization error of operation entities with high EID will cause error accumulation in following situation interaction.
- (2) Adopting the concession pattern in negotiation in any possible way: according to Property 4, when target entity is changed, structure of C2 organization should be adjusted for controlling more resource quickly, which can transfer the target entity to be the most important entity for concession negotiation.

Based on the analysis above, ODMS strategies set within time restriction can be constructed as $U = \{u_i \mid \gamma \leq u_i \leq 1, i \in N\}$, where u_i is the SIQ of each interaction segment.

5. Simulation Cases

5.1. Parameters Setting. Taking the operational organization in [25], for example, the organization is abstracted as network of C2 organization with 25 nodes and 4 command corporations shown in Figure 4, whose average degree is $\langle k \rangle = 1.92$, and parameters are initialized as seen in Table 4. Suppose that SPL is uniform distribution within $[0, 0.1]$, and entity o with the highest EID is induced as target entity with facility.

5.2. ODMS Performance in Continuous Changing Situation. In the ideal condition ($\beta = 0$), according to parameters setting above, stable situation (as special continuous changing situation) and three continuous changing situation patterns

TABLE 4: Initialization parameters of C2 organization.

No.	ID	x_{i0}	NET1 NID ₁	NET2 NID ₂	NET3 NID ₃
1	o	0.0203	0.7100	0.5300	0.2866
2	op1	0.0199	0.5242	0.4426	0.4150
3	op2	0.0604	0.6608	0.6892	0.6550
4	op3	0.0272	0.4742	0.3792	0.4644
5	op4	0.0199	0.4042	0.4152	0.3916
6	p11	0.0015	0.0850	0.0700	0.0634
7	p12	0.0747	0.0850	0.0700	0.0634
8	p13	0.0445	0.2416	0.2266	0.2200
9	p131	0.0932	0.0500	0.0500	0.0500
10	p132	0.0466	0.0500	0.0500	0.0500
11	p21	0.0419	0.3900	0.3360	0.3326
12	p211	0.0846	0.0550	0.0550	0.0550
13	p212	0.0525	0.0550	0.0550	0.0550
14	p213	0.0203	0.0550	0.0550	0.0550
15	p22	0.0672	0.3900	0.3360	0.3326
16	p221	0.0838	0.0550	0.0550	0.0550
17	p222	0.0020	0.0550	0.0550	0.0550
18	p223	0.0681	0.0550	0.0550	0.0550
19	p31	0.0379	0.0574	0.0574	0.0466
20	p32	0.0832	0.0574	0.0574	0.0466
21	p33	0.0503	0.0574	0.0574	0.0466
22	p34	0.0709	0.0574	0.0574	0.0466
23	p41	0.0429	0.0550	0.0440	0.0422
24	p42	0.0305	0.0550	0.0440	0.0422
25	p43	0.0190	0.0550	0.0440	0.0422

are given together in Table 5. As limit of paper, the sudden changing situation within $t \in [0, 5]$ and the IPA of each entity are analyzed in Figure 5.

The axis of abscissa shows time consumed and axis of ordinate express SCL, and then complete ODMS under four changing situations, shown in Figure 6, can be described as curve cluster about the SCL changing process of 25 operational entities.

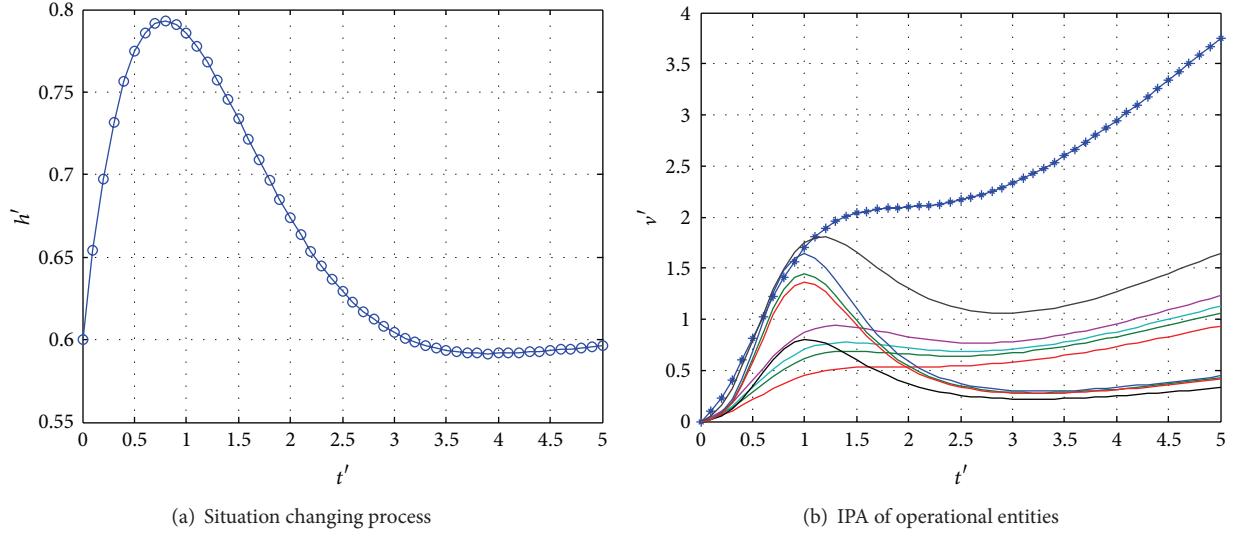


FIGURE 5: Situation in sudden changing pattern.

TABLE 5: The SCR in continuous changing situation.

No.	Situation change patterns	SCR
1	Stable situation	$h'_t = 0$
2	Identical changing situation	$h'_t = 0.8$
3	Gradual changing situation	$h'_t = 0.17 * t + 0.2$
4	Sudden changing situation	$h'_t = 0.6 * (1 - \sin(t - \pi)) * e^{-t}$

With two phases in ODMS process, STE can be obtained on the basis of time consuming within two stages correspondingly and shown in Figure 7(a). Meanwhile, SQ can be gained with cognition change in two stages and shown in Figure 7(b). Other measure indexes of ODMS are listed in Table 6. From the statistics in Table 6, with the increase of SCR, the ST is shortened correspondingly.

As in process P1 in Figure 6(a), due to little change in the situation, there are fewer nodes updating situation information. Then long time is consumed in situation cognition stage, but the SCL is not improved highly, which leads to more time used to negotiate in situation interaction stage, and ST is extended in the ODMS, as P1 process in Figure 7.

As in process P2 and process P3 in Figures 6(b) and 6(c) respectively, with SCR increasing, the number of nodes updating information raises correspondingly, and then IPA and SCL in situation cognition stage are improved in a short time. Compared with process P1, time consumed in situation interaction stage is shortened, and then ST can be reduced, which is shown as process P2 and process P3 in Figure 7.

As process P4 in Figure 6(d), the SCL of each operational entity improves with SCR increasing. When situation suddenly has a great change, it leads to situation information explosion apperceived by C2 organization in a short time. The communication routes are easy to be blocked, which can make the IPA decrease among the operational nodes and the SCL reduced (as subgraph shown in Figure 10 within time interval [1, 2]). After the decrease of SCR, the C2

TABLE 6: Indexes list of ODMS.

No.	ID	T	γ_{io}	τ	χ
1	P1	24.3425	1	1.4420	6.7622
2	P2	15.2562	1	1.8034	2.1064
3	P3	14.2024	1	1.8306	1.5813
4	P4	11.9370	1	1.7543	1.2449

organization's ability of efficiently processing and interpreting information recovers gradually, and SCL of each entity is still improved. Due to overload processing of situation information, the operational nodes have the highest SCL while consuming the minimum time, which is benefit for realizing ODMS in shortest time, shown in Figure 7.

According to above analysis, under continuous changing situation, the situation cognition stage is the foundation of situation interaction stage, which directly influences the STE and SQ of ODMS. When the target entity is transferred from o to op2, the ODMS performance of C2 organization with tree structure drops obviously under the stable situation (shown in Figure 8), which denotes that it is weak for C2 organization with tree structure to adapt situation change.

So in order to adapt the complex environment change, C2 organization should make entities obtain IPA as many as possible. The main ways are changing organization structure and selecting appropriate strategy to complete the situation interaction process. The following sections about two aspects strategies are used to improve the ODMS performance within environment influence and time restriction separately.

5.3. ODMS Strategies under Environment Influence. Taken stable situation as an example, environment influence factor $\beta = [0, 0.2]$ is introduced to analyze adaptive ability of C2 organization with different structures. Then the cross-structure ($\langle k \rangle = 2.02$) is constructed on the basis of tree structure shown in Figure 9.

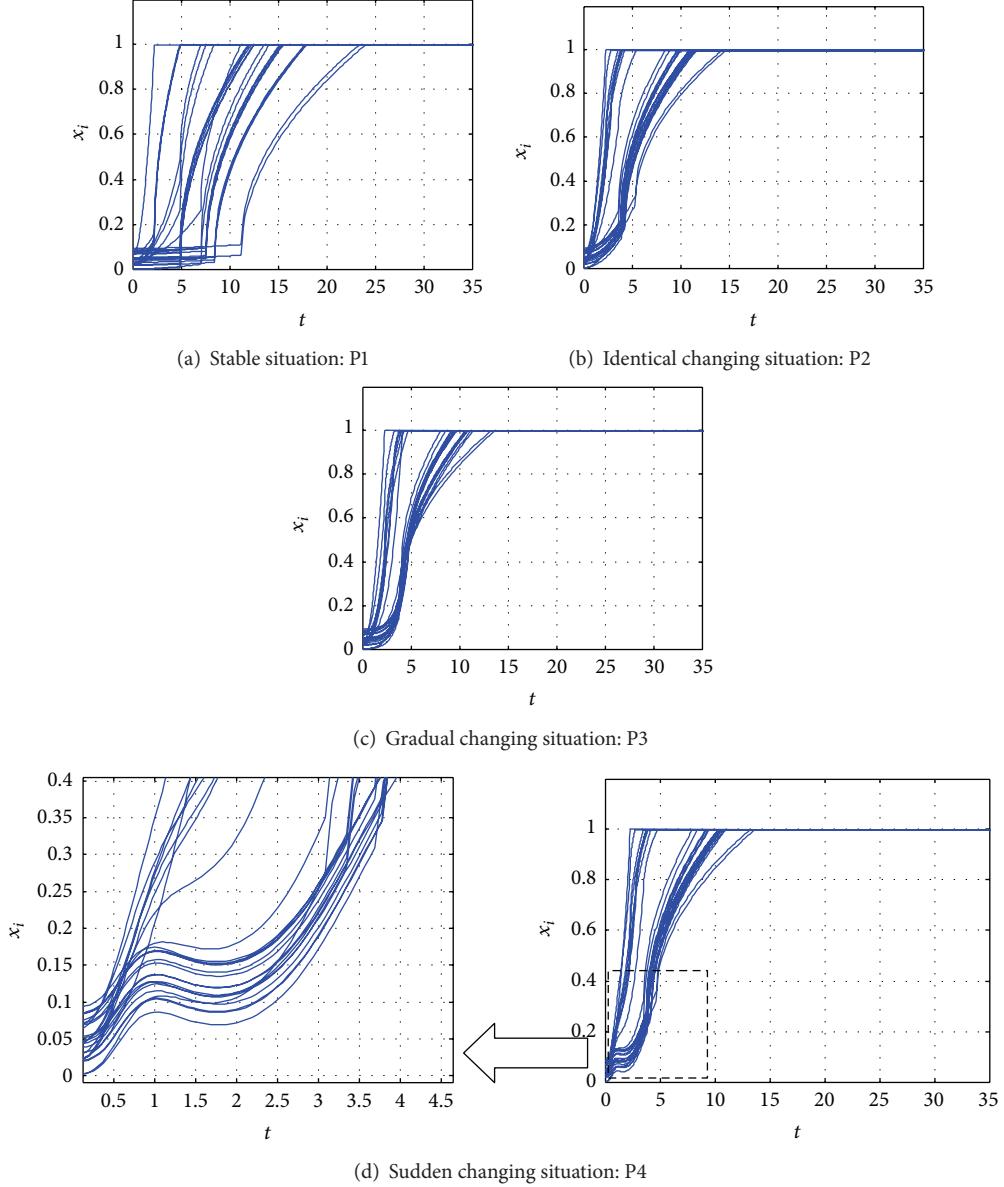


FIGURE 6: Complete ODMS under four changing situation.

When $\beta = 0$, the situation interaction route in NET1 is still adopted in NET2, and the ODMS performance is shown in Figure 15. When $\beta = [0, 0.2]$, it can reduce the communication unblocked degree C . Taking op2 as target entity, ODMS strategies set of choosing situation interaction route is analyzed in Table 7 according to Properties 1 and 2, and (*) denotes the chosen situation interaction routes.

In Table 7, $r_{\text{op}2 \rightarrow p31}^{1*}$, $r_{\text{op}2 \rightarrow p41}^{1*}$, and $r_{\text{op}2 \rightarrow p12}^{2*}$ are the best situation interaction route separately with Property 1. Moreover, for the target entities op2 and p21, $r_{\text{op}2 \rightarrow p211}^{1*}$ and $r_{\text{op}2 \rightarrow p221}^{2*}$ are the best situation interaction route to p22, respectively, according to Property 2. So the situation interaction routes in NET2 are given as $R_{\text{NET2}} = \{r_{\text{op}2 \rightarrow p31}^{1*}, r_{\text{op}2 \rightarrow p41}^{1*}, r_{\text{op}2 \rightarrow p12}^{2*}, r_{\text{op}2 \rightarrow p211}^{1*}, r_{\text{op}2 \rightarrow p221}^{2*}\}$.

Comparing Figure 8 with Figure 10(a), networked level of NET2 is higher than NET1, then and IPA of NET2 improves but shortens its ST under stable situation with $\beta = 0$. Comparing Figure 9 with Figure 10(b), although environment influence factor is increased ($\beta = [0, 0.2]$), situation interaction routes chosen in reason can improve the quality of situation interaction, which also improves the ODMS performance. So the adaptive ability of NET2 with cross-structure is stronger than NET1 with tree structure. Based on this, C2 organization with network structure is generated by further extending networked level in Figure 11, then ODMS performance is shown in Figure 12, and strategies of choosing situation interaction route are listed in Table 8.

Contrastive analysis in both NET3 and NET2, network degree of NET3 is higher than NET2, namely, $\langle k \rangle = 2.24$. It

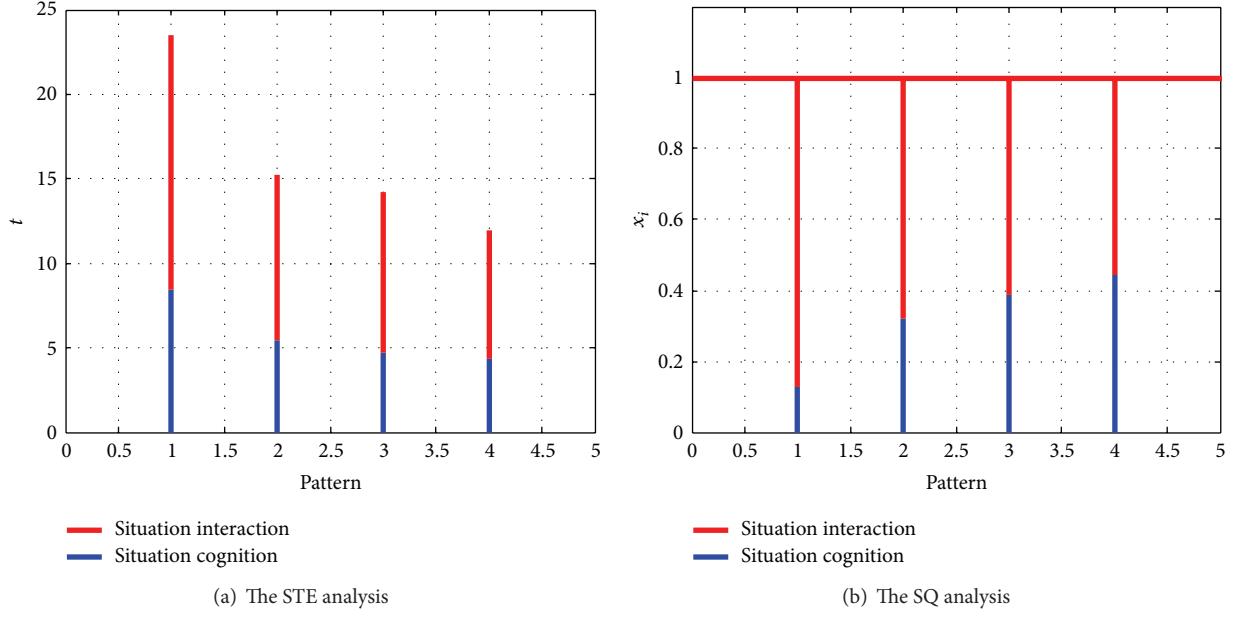


FIGURE 7: The STE and SQ under different changing situations.

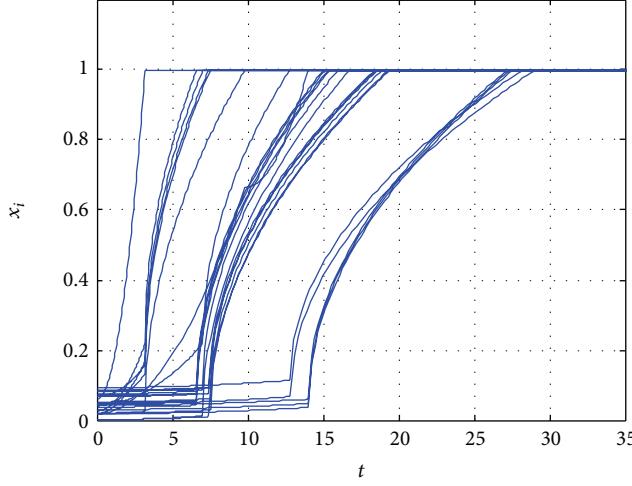


FIGURE 8: Complete ODMS in NET1 under stable situation with target entity op2: P5.

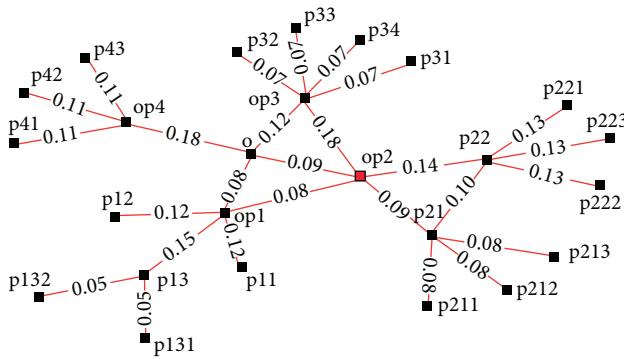


FIGURE 9: C2 organization with cross structure: NET2.

TABLE 7: Situation interaction routes within NET1.

$R_{st_f \rightarrow en_f}$	No.	Situation interaction routes	Δ_r
$r_{op2 \rightarrow p^{31}}$	1*	$op2 \rightarrow op3 \rightarrow p^{31}$	0.1107
	2	$op2 \rightarrow o \rightarrow op3 \rightarrow p^{31}$	0.1351
$r_{op2 \rightarrow p^{41}}$	1*	$op2 \rightarrow o \rightarrow op4 \rightarrow p^{43}$	0.2001
	2	$op2 \rightarrow op1 \rightarrow o \rightarrow op4 \rightarrow p^{43}$	0.2149
$r_{op2 \rightarrow p^{211}}$	3	$op2 \rightarrow op3 \rightarrow o \rightarrow op4 \rightarrow p^{43}$	0.2486
	1	$op2 \rightarrow o \rightarrow op1 \rightarrow p^{12}$	0.1757
$r_{op2 \rightarrow p^{12}}$	2*	$op2 \rightarrow op1 \rightarrow p^{12}$	0.1105
	1*	$op2 \rightarrow p^{21} \rightarrow p^{211}$	0.0653
$r_{op2 \rightarrow p^{221}}$	2	$op2 \rightarrow p^{22} \rightarrow p^{211}$	0.0859
	1	$op2 \rightarrow p^{22} \rightarrow p^{221}$	0.1009
	2*	$op2 \rightarrow p^{21} \rightarrow p^{22} \rightarrow p^{221}$	0.0859

* means that route with the symbol is the best situation interaction route from the fixed target entity to the fixed compromise entity and can be selected to negotiate.

has benefit to allocating decision rights, which is good for adjusting situation interaction routes to adapt environment influence. From entity $op2$ to entity p^{31} , $r_{op2 \rightarrow p^{31}}^{2*}$: $op2 \rightarrow o \rightarrow op3 \rightarrow p^{31}$ in Table 8 and $r_{op2 \rightarrow p^{31}}^{1*}$: $op2 \rightarrow op3 \rightarrow p^{31}$ in Table 7 are the results of allocating decision rights agilely. So situation interaction routes in NET3 are given as $R_{NET3} = \{r_{op2 \rightarrow p^{31}}^{2*}, r_{op2 \rightarrow p^{41}}^{4*}, r_{op2 \rightarrow p^{12}}^{1*}, r_{op2 \rightarrow p^{211}}^{1*}, r_{op2 \rightarrow p^{221}}^{2*}\}$.

5.4. ODMS Strategies within Time Restriction. Suppose that, in order to realize cooperation among operational entities, the time restriction is within 20 unit time (remarked as red vertical real line in Figure 12), and the accepted synchronization error within 5%, and then SE should reach 95% at least (remarked as red horizontal dotted line in Figure 12). Then

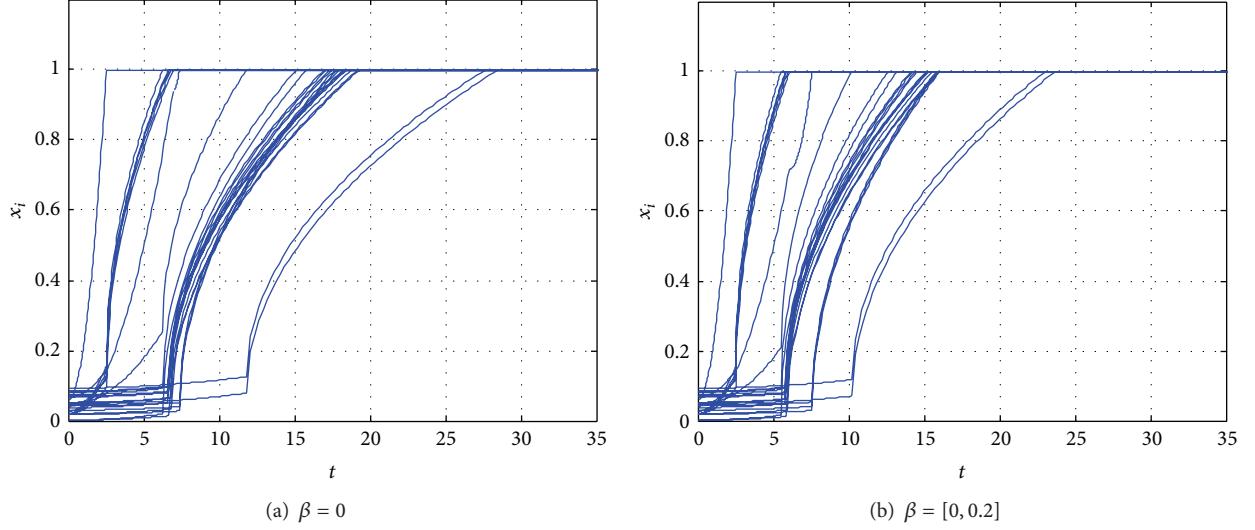


FIGURE 10: Complete ODMS in NET2 with different β .

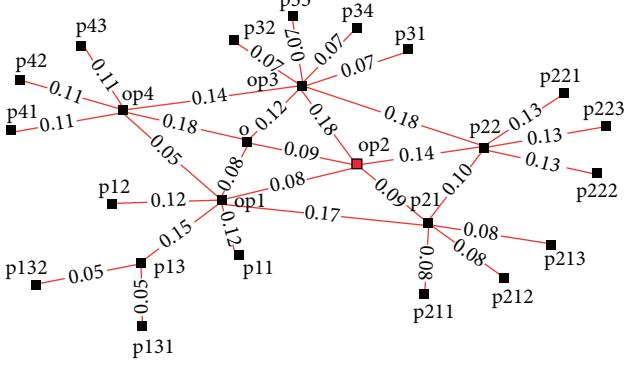


FIGURE 11: C2 organization with network structure: NET3.

ODMS performance is acceptable under other three situations changing patterns with NET1 but the stable situation with NET3. As communication is blocked in some extent within NET3, it takes long time for situation interaction, which makes ST beyond the time restriction.

As shown in sub graph of Figure 12, entities p132 and p131 do not fulfill the time restriction. The strategies set of incomplete ODMS is given as UP_1^1 and UP_2^1 in Table 9, and the processes of incomplete ODMS are shown in Figure 13.

Because that entity op2 is the most important entity in NET3 (shown in Table 4), other entities should adopt concession negotiation pattern to realize situation interaction in high quality, so the ODMS performance is new, and the ST of C2 organization is in low adjustability. Following the strategies in Table 9, the ST of incomplete ODMS is shortened with 0.772 and 0.765, respectively, and comprehensive synchronization extent S is 0.9946 and 0.9967, respectively, which can fulfill the operation demand.

6. Discussion

As known in simulation cases, ODMS with three phases is studied in detail in the paper. Then compared with models

TABLE 8: Situation interaction routes within NET2.

$R_{st_i \rightarrow en_i}$	No.	Situation interaction routes	Δ_r
$r_{op2 \rightarrow p31}$	1	op2 → op3 → p31	0.1255
	2*	op2 → o → op3 → p31	0.0583
$r_{op2 \rightarrow p41}$	1	op2 → op1 → op4 → p41	0.0017
	2	op2 → o → op4 → p41	0.0020
	3	op2 → op3 → op4 → p41	0.0034
	4*	op2 → o → op1 → op4 → p41	0.0015
$r_{op2 \rightarrow p12}$	1*	op2 → o → op1 → p12	0.0916
	2	op2 → op1 → p12	0.0975
$r_{op2 \rightarrow p211}$	1*	op2 → p21 → p211	0.0623
	2	op2 → p22 → p21 → p211	0.0812
$r_{op2 \rightarrow p221}$	1	op2 → p22 → p221	0.0964
	2*	op2 → p21 → p22 → p221	0.0812

* means that route with the symbol is the best situation interaction route from the fixed target entity to the fixed compromise entity and can be selected to negotiate.

referred to in Section 1, the advantages and shortcomings of the ODMS model are analyzed and discussed in this paper.

For the group communication model in [5, 6], it does not refer to the opinion formation. Taking the incomplete ODMS process P_1^0 in Figure 12 as example, the SCL in the time of $t = 13$ is regarded as original opinions, and communication parameters in [5] are introduced as opinion accepted degree $u = 0.2$ and communication allow degree $d = 0.2$. Suppose that 1 unit time is consumed in the each communication process in the whole C2 organization, and then the results of two methods can be expressed in the same picture. After 25 units time, the synchronization process in [5] is shown with red real line in Figure 14. As communication processes have no effect on the SCL improvement of certain entity in the end, it can hardly get to the unanimous status, which also makes ODMS realization in hard.

For the models of CE and CSSync in [13], they do not refer to cognitive formation but only the difference of cognition.

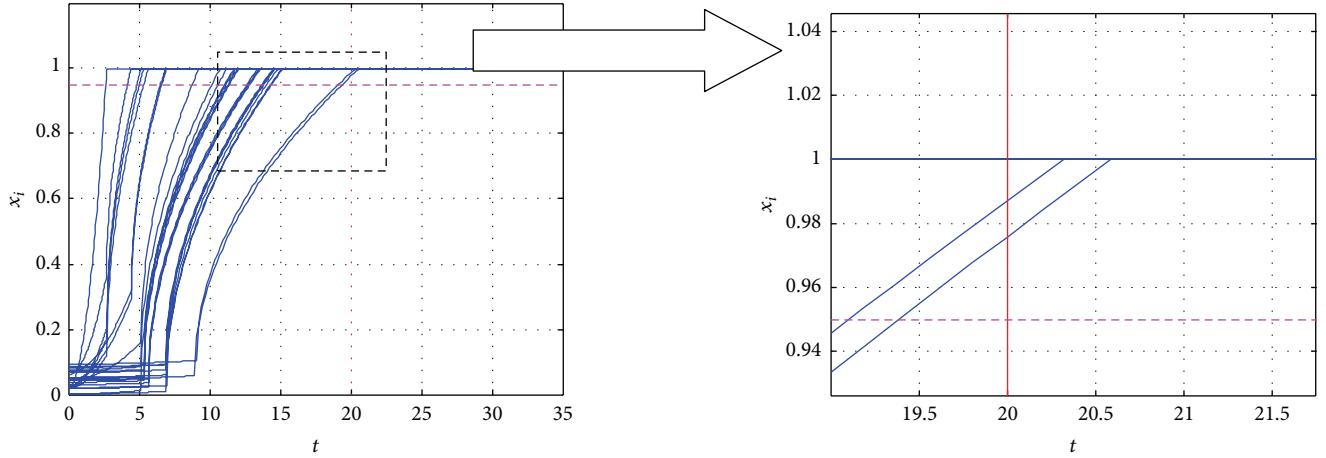
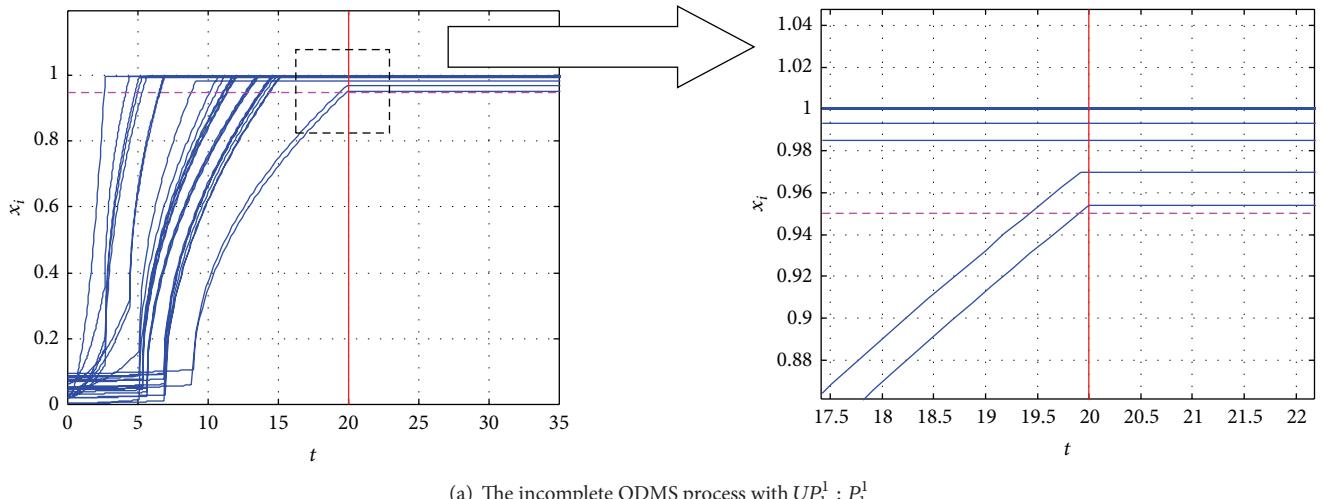
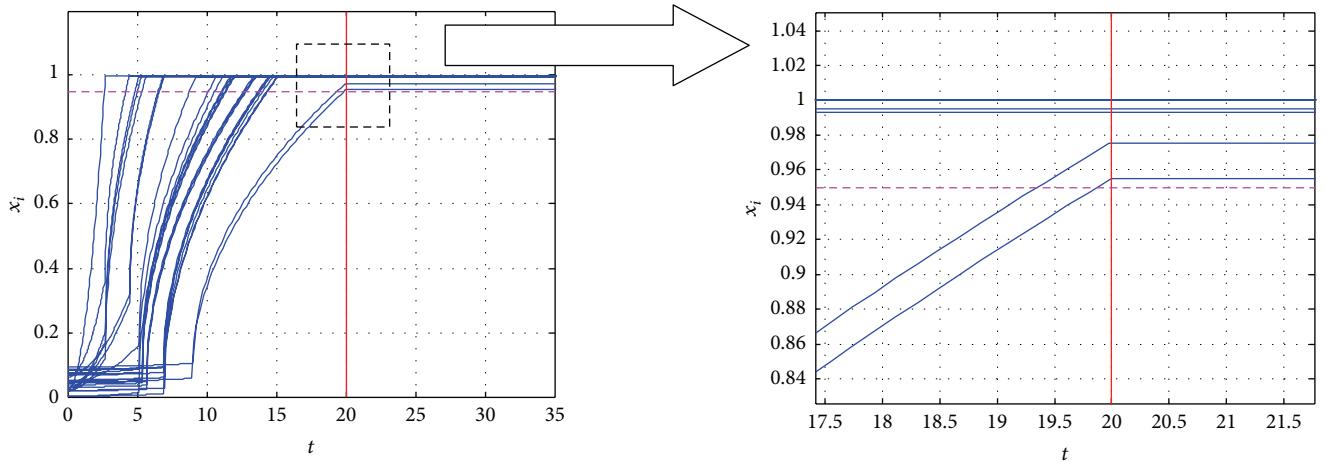
FIGURE 12: Complete ODMS in NET3 with $\beta = [0, 0.2]$ as P_1^0 .(a) The incomplete ODMS process with $UP_1^1 : P_1^1$ (b) The incomplete ODMS process with $UP_1^2 : P_1^2$

FIGURE 13: Incomplete ODMS under stable situation.

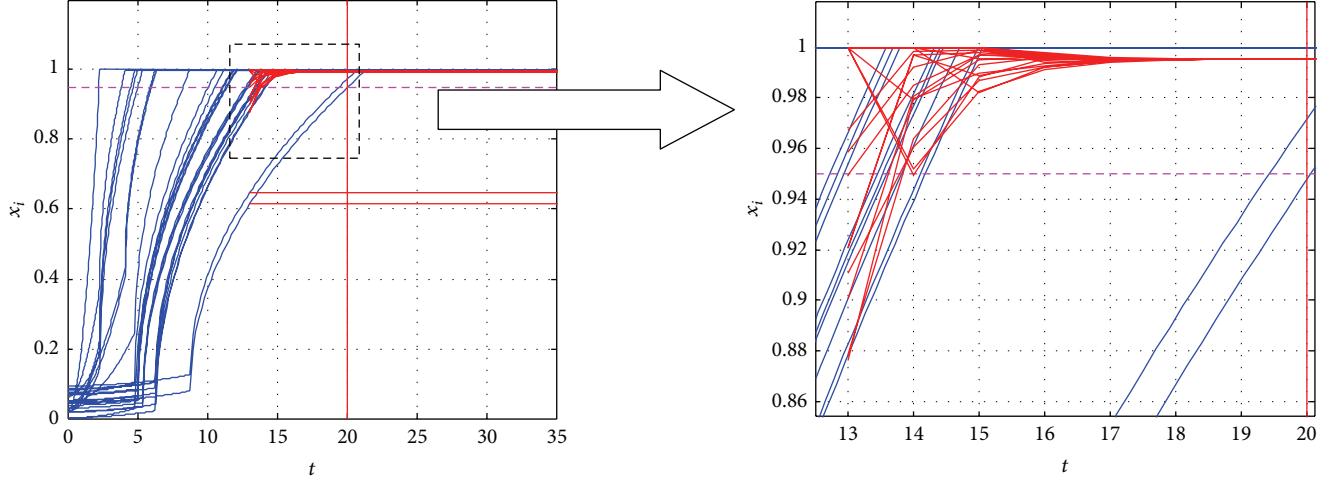


FIGURE 14: Comparing with the models in [5].

TABLE 9: Strategies sets of incomplete ODMS.

No.	EID	UP_1^1	UP_1^2
1	o	—	—
2	op1	0.985	0.993
3	op2	1	1
4	op3	1	1
5	op4	1	1
6	p11	1	1
7	p12	1	1
8	p13	0.992	0.994
9	p131	0.996	0.990
10	p132	0.983	0.987
11	p21	1	1
12	p211	1	1
13	p212	1	1
14	p213	1	1
15	p22	1	1
16	p221	1	1
17	p222	1	1
18	p223	1	1
19	p31	1	1
20	p32	1	1
21	p33	1	1
22	p34	1	1
23	p41	1	1
24	p42	1	1
25	p43	1	1

After the SCL of the entities is all in the complete status, then $CSSync = 1$, which reflects that complete ODMS. The process of $CSSync$ is shown in Figure 15.

For Kuramoto model in [16], synchronization process is only controlled by a coupled coefficient (as feedback of synthesized information difference), and it does not take the

strategies of ODMS into account, which makes it hard to adapt to the environment change.

According to the ODMS processes in NET1, NET2 and NET3, we can summarize four factors for influencing the ODMS capability of different organizations, including organization structure, shared information resources, interactions patterns, and allocation of decision rights. Then inhibitors and enablers for different ODMS factors are given for compared analyzing in Table 10.

For the 1st factor about organization structure, with the networked level of C2 organization increasing from NET1 to NET3, IPA of operational entities can be improved remarkably for promoting the SCL, which is benefit for enhancing the ODMS performance. Then enabler of networked structure is networked structure, and the inhibitor of organization structure is traditional tree structure. For the 2nd factor about shared information resources, as P1 in Figure 6(a) and ODMS process in Figure 10, the situation information resources are shared in sufficiency, which is good for shortening the ST. Thus enabler for the factor is shared across members and all information accessible across entities, and inhibitor for the factor is none or a few shared (mainly kept within own entities). For the 3rd factor about interactions patterns, comparing P5 with P1, P2, P3, and P4, as the entity op2 is not the most important entity in NET1, there are fixed restrictions of allocating decision rights, which is not good for forming the suitable situation interaction route for adapting to the environment influence. So interactions patterns with highly fixed and constrained in superior-subordinate will inhibit ODMS capability improvement. However, compared with the complete ODMS processes in NET2 (within Figure 10) and NET3 (within Figure 12), there are several interaction patterns to be selected, and we can change situation interaction pattern in agility, which is good for improving the SIQ; thus interactions patterns with unconstrained and agile across entities can be regarded as the enabler of the factor. For the 4th factor about allocation of decision rights, the fixed task-role-based decision rights can inhibit the situation cognition

TABLE 10: ODMS inhibitors and enablers.

ODMS factors	Inhibitors	Enablers
Organization structure	Traditional tree structure	Networked structure
Shared information resources	None or a few shared (mainly kept within own entities)	Shared across members and all information accessible across entities.
Interactions patterns	Highly fixed and constrained in up-down grade	Unconstrained and agile across entities
Allocation of decision rights	Fixed task-role based	Distributed to all subjects

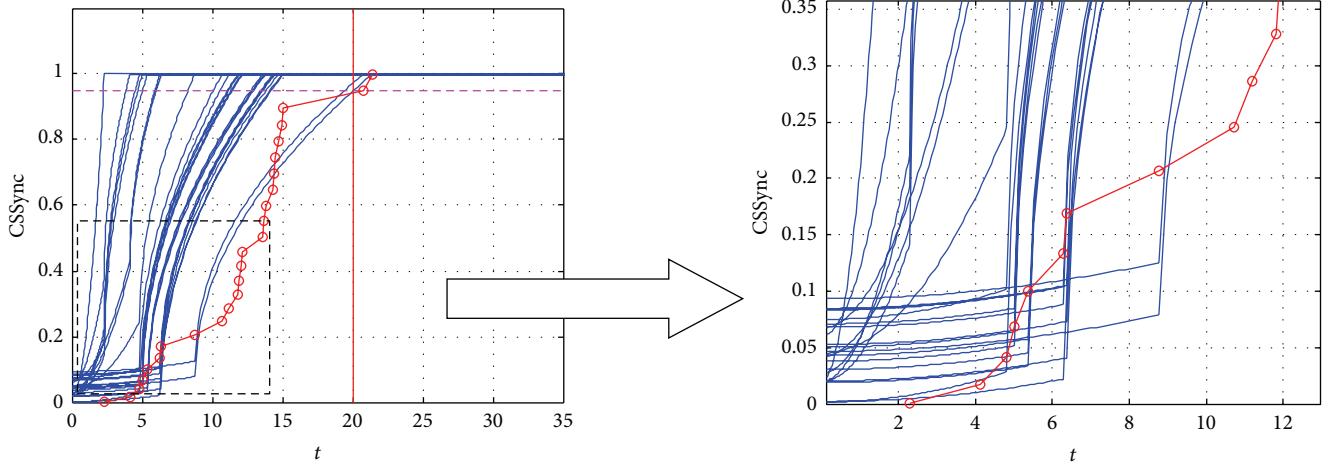


FIGURE 15: Comparing with the models in [13].

level formation and interaction pattern optimizing selection, which will depress the ODMS capability of the organization. However, with the networked level of organizations increasing, the decision rights can be distributed to different subjects according to situation change, which is good for adjusting situation interaction routes to adapt to environment influence and can be regarded as enabler for the factor.

7. Conclusion and Further Work

With the viewpoint of time consumed, the model of ODMS is constructed with three phases under complex environment, and the strategies of ODMS are put forward in the paper within environment influence and time restriction to improve the ODMS performance for adapting to environment. Concerning the several factors in ODMS, such as situation interaction patterns, situation interaction route, and situation in different changing patterns, the process of ODMS can be controlled more reasonable, and the results of this model are better than the other previous models to fulfill the actual demand.

As the environment of battlefield is complex and uncertain, the influence on ODMS is changing instantly. The C2 organization is the critical factor. So the future work might involve the dynamic self-adaptation of C2 organization within different environment. If several operational entities cannot percept situation information, the organization structure should be adjusted and optimized to make these entities form IPA. Then how to adjust C2 organization in optimization, such as adding and cutting down entities, or

changing the structure on the basis of fixed entities, is the difficult problems to enhance the ODMS performance.

Acknowledgment

Research works in this paper are supported by the National Natural Science Foundation of China (61272011).

References

- [1] D. Alberts and J. Garstka, *Network Centric Warfare: Developing and Leveraging Information Superiority* (2nd Edition Revised), The Command and Control Research Program (CCRP) Publication Series, Washington, DC, USA, 2002.
- [2] J. R. Butler, *Network Centric Warfare (NCW): The Mechanism for Change*, ADA520230, Marine Corps University, 2003.
- [3] J. Garstka and D. Alberts, *Network Centric Operations Conceptual Framework (Version 2.0)*, Evidence Based Research, 2007.
- [4] H. Liu and L. Zhang, "Analysis on opinion emergence of group communication," *Complex Systems and Complexity Science*, vol. 4, no. 1, pp. 45–52, 2004.
- [5] B. Wang, *Research on Operation Synchronization Modeling Based on Complex Networks*, National University of Defense Technology, 2007.
- [6] X. K. Bao, *Modeling and Analyzing of Networked Operation Synchronization*, National University of Defense Technology, 2009.
- [7] S. F. Fatima, M. Wooldridge, and N. R. Jennings, "An agenda-based framework for multi-issue negotiation," *Artificial Intelligence*, vol. 152, no. 1, pp. 1–45, 2004.

- [8] J. Scrocca, M. Molz, and A. Kott, "Collaborative awareness: experiments with tools for battle command," in *Proceedings of the 11th International Command and Control Research and Technology Symposium*, 2006.
- [9] A. Kirlik and R. Strauss, "Situation awareness as judgment I: statistical modeling and quantitative measurement," *International Journal of Industrial Ergonomics*, vol. 36, no. 5, pp. 463–474, 2006.
- [10] R. Strauss and A. Kirlik, "Situation awareness as judgment II: experimental demonstration," *International Journal of Industrial Ergonomics*, vol. 36, no. 5, pp. 475–484, 2006.
- [11] A. Kirlik and R. Strauss, "Situation awareness as judgment I: statistical modeling and quantitative measurement," *International Journal of Industrial Ergonomics*, vol. 36, no. 5, pp. 463–474, 2006.
- [12] H. Meng and D. G. Zhang, "Model of situation awareness velocity under continuous changing situation," *Systems Engineering and Electronics*, vol. 31, no. 12, pp. 2909–2913, 2009.
- [13] M. Manso and J. Moffat, "Defining and measuring cognitive-entropy and cognitive self-synchronization," in *Proceedings of the International Command and Control Research and Technology Symposium*, 2011.
- [14] M. Ruddy, "ELICIT—the experimental laboratory for investigating collaboration, information sharing and trust," in *Proceedings of the International Command and Control Research and Technology Symposium*, 2007.
- [15] A. H. Dekker, "Average distance as predictor of synchronisability in networks of coupled oscillators," in *Proceedings of the 33rd Australasian Computer Science Conference*, vol. 102, pp. 127–131, 2010.
- [16] A. H. Dekker, "Analyzing C2 structures and self-synchronization with simple computational models," in *Proceedings of the International Command and Control Research and Technology Symposium*, 2011.
- [17] L. J. Pan, X. J. Li, and D. S. Hao, "Research on modeling of battle collaboration effect," *Journal of System Simulation*, vol. 23, no. 12, pp. 2561–2565, 2011.
- [18] P. Holme, "Congestion and centrality in traffic flow on complex networks," *Advances in Complex Systems*, vol. 32, no. 6, pp. 163–176, 2003.
- [19] D. Gómez, E. González-Arangüena, C. Manuel, G. Owen, M. del Pozo, and J. Tejada, "Centrality and power in social networks: a game theoretic approach," *Mathematical Social Sciences*, vol. 46, no. 1, pp. 27–54, 2003.
- [20] S. A. David and M. James, *Network Centric Warfare and Complexity Theory*, Publishing House of Electronics Industry, Beijing, China, 2004.
- [21] Y. Liu, J. Zhao, J. N. Liu, D. J. Feng, and G. Y. Wang, "Novel battlefield electromagnetic environment classification method based on combat efficacy of weapon systems," *Systems Engineering and Electronics*, vol. 33, no. 5, pp. 1059–1062, 2011.
- [22] X. C. Gong, Y. C. Jiang, and H. Chen, "Command and control in the network-centric warfare," *Fire Control & Command*, vol. 35, no. 3, pp. 1–8, 2010.
- [23] J. Wu and Y. J. Tan, "Finding the most vital node by node contraction in communication networks," in *Proceedings of the International Conference on Communications, Circuits and Systems*, vol. 2, pp. 1283–1286, May 2005.
- [24] Y. Zhao, Z. Wang, J. Zheng, and X. Guo, "Finding most vital node by node importance contribution matrix in communication networks," *Journal of Beijing University of Aeronautics and Astronautics*, vol. 35, no. 9, pp. 1076–1079, 2009.
- [25] J. Jiang, *Research on Executable Architecture and the Executable Method of DoDAF*, National University of Defense Technology, 2008.

Research Article

Windowed Least Square Algorithm Based PMSM Parameters Estimation

Song Wang

School of Mechanical, Electrical & Information Engineering, Shandong University, Weihai, China

Correspondence should be addressed to Song Wang; wangsong_sdu@163.com

Received 4 June 2013; Revised 26 July 2013; Accepted 26 July 2013

Academic Editor: Juan J. Nieto

Copyright © 2013 Song Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Stator resistance and inductances in d -axis and q -axis of permanent magnet synchronous motors (PMSMs) are important parameters. Acquiring these accurate parameters is usually the fundamental part in driving and controlling system design, to guarantee the performance of driver and controller. In this paper, we adopt a novel windowed least algorithm (WLS) to estimate the parameters with fixed value or the parameter with time varying characteristic. The simulation results indicate that the WLS algorithm has a better performance in fixed parameters estimation and parameters with time varying characteristic identification than the recursive least square (RLS) and extended Kalman filter (EKF). It is suitable for engineering realization in embedded system due to its rapidity, less system resource possession, less computation, and flexibility to adjust the window size according to the practical applications.

1. Introduction

The high-field-strength neodymium-iron-boron (NdFeB) magnets have become commercially available with affordable prices, so the permanent magnet synchronous motor (PMSM) is receiving increasing attention due to its high speed, high power density, and high efficiency. It is very suitable for some high-performance requirement applications, for example, robotics, aerospace, electric ship propulsion systems, and wind power generation systems [1–3]. It has been shown that PMSM can provide significant performance improvement in many variable speed applications [4]. The commonly used control method in motor control is vector control. The method has a requirement of obtaining related parameters of the motor. Therefore, acquiring accurate parameters of the motor is usually the fundamental part in driving system design. We cannot measure the motor parameters with normal no-load test and locked rotor test in the work site. Moreover, with the increasing working time of the motor and the surrounding environment changes, some parameters of the motor will be changed. Therefore, servo drivers usually have the function of parameters identification and self-tuning [3].

Stator resistance and inductances in d -axis and q -axis are important parameters of motor model, which are considered as constants usually. However, these parameters vary with different operation conditions when motor is running [5]. The study object of this paper was permanent magnet servomotor produced by Huada Company in Wuhan of China. The experiment data showed that stator resistance value ranged from R_s to $1.3R_s$ and that inductance value ranged from L_d to $1.004L_d$ when the temperature ranged from 20 degrees to 80 degrees. Thus, the temperature of motor had a great influence on stator resistance and inductance in d -axis and q -axis. When these parameters are treated as constants, the stability and control performance of the system will be affected. Therefore, the realization of parameters (R_s , L_d , and L_q) identification is essential for motion control of PMSM.

Model identification and parameter estimation techniques have become mature after years of development. From the least square estimation theory [6–14] and its various improved algorithm [15–23] to Kalman filter algorithm [24–27], neural network [28, 29], genetic algorithm [30–32], and so forth, they can serve as parameter estimation tools. However, these methods have their own characteristics and applicability.

The least square estimation is one of the most simple and most mature parameter estimation methods. However the amount of calculation of the traditional least square method will increase with time sequence increase. It is hard to realize in embedded chip due to the large amount of calculation, and there is a problem of data saturation. Kalman filtering algorithm is put forward for system identification by Kalman in 1960, and there is a wide range of use. However, it is sensitive to the initial conditions [26], and its performance is poor for time varying parameter identification [33, 34]. With the development of artificial intelligence technology, neural network [28, 29] and genetic algorithm [30–32] are used to the parameter estimation. These intelligent methods can get the identification results with high accuracy. However, it is hard to apply them in practical parameter estimation due to the large amount of calculation and complexity of the algorithm. Therefore, least square algorithm is also a commonly parameter estimation method. Kinds of improved algorithms are proposed to promote the identification performance of traditional least square algorithm [18–22, 35–41]. For example, the forgetting factor is introduced in the recursive least square estimation, and the past time of data will be forgotten by index rate [42, 43]. However, it still cannot discard the past time of data [16] but just weakens the impact of the past time of data for the current parameter estimation. Another method is to use window method for the time series data [16, 23, 44]. This method can discard the past time of data flexibly and eliminate the impact of the past time of data for future parameter estimation. The window size can be set flexible according to practical application. In this paper, we adopt windowed least square algorithm for stator resistance R_s , L_d , and L_q inductance estimation and make a comparison with recursive least square and extended kalman Filter (EKF). From the simulation result, we can see that windowed least square algorithm has a better performance in convergence speed and identification precision for fixed parameters and parameters with time varying characteristics. From the view of algorithm complexity, the windowed least square algorithm is suitable for engineering realization in embedded chip, such as DSP and ARM.

This paper is consisted of the five sections. Section 2 describes the principle of least square theory and the recursive least square algorithm. Section 3 illustrates the windowed least square algorithm. Section 4 does some simulations for PMSM parameter estimation. Section 5 analyses the simulation results and shows some conclusions.

2. Least Square Estimation and Recursive Least Square Estimation

2.1. The Principle of Least Square Estimation. The earliest stimulus for the development of the least square estimation theory was apparently provided by astronomical studies in which planet and comet motions were studied using telescopic measurement data. The principle of the parameter estimation is simple and does not need any statistical characteristics of the variables. It is used in system identification and parameter estimation widely. The least square estimation still

can provide an accurate solution when other identification methods lose efficacy.

Supposing $y(i)$ and $x_1(i), x_2(i) \dots x_n(i)$ are the observation sequences of y and x at $t_1, t_2 \dots t_m$. The relationship of y and x is expressed

$$\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(m) \end{bmatrix} = \begin{bmatrix} x_1(1) \dots x_n(1) \\ x_1(2) \dots x_n(2) \\ \vdots \\ x_1(m) \dots x_n(m) \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{bmatrix}, \quad (1)$$

where $\theta = (\theta_1, \theta_2, \dots, \theta_n)$ is the measured parameter set and n is the number of parameters. We hope to estimate their values by the observation value of y and x at different time sequences. m is the time sequences; to estimate the n parameters $\theta_i, m \geq n$ is required, and if $m = n$, we can get the single solution from (1) as (2):

$$\hat{\theta} = X^{-1}y, \quad (2)$$

where $\hat{\theta}$ is the estimation value of θ and inverse matrix X^{-1} of X is required:

$$\varepsilon = y - X\theta, \quad (3)$$

where $\varepsilon = (\varepsilon_1, \varepsilon_2 \dots \varepsilon_m)^T$ is the error vector.

The target function is shown in the following:

$$J = \sum_{i=1}^m \varepsilon_i^2 = \varepsilon^T \varepsilon. \quad (4)$$

Obtaining $\hat{\theta}$ to make J minimum,

$$\frac{\partial J}{\partial \theta} \Big|_{\theta=\hat{\theta}} = -2X^T y + 2X^T X \hat{\theta} = 0. \quad (5)$$

The result is

$$\hat{\theta} = (X^T X)^{-1} X^T y, \quad (6)$$

$\hat{\theta}$ is the least square estimate, LS of θ .

2.2. Recursive Least Square Estimation. In practical parameter estimation, the data is always constantly to be refreshed. Therefore, we can further deduce (2) to a recursion algorithm. This algorithm does not need to compute the inverse matrix calculation repeatedly and reduces the time-consuming and system resources occupation.

The least square estimation using m groups of data is shown as follows:

$$\hat{\theta}(m) = (X_m^T X_m)^{-1} X_m^T Y_m. \quad (7)$$

The $m+1$ moment data is $(x(m+1), y(m+1))$, and then

$$Y_{m+1} = X_{m+1} \theta(m+1), \quad (8)$$

where

$$Y_{m+1} = \begin{bmatrix} y(1) \\ \vdots \\ y(m) \\ \dots \\ y(m+1) \end{bmatrix} = \begin{bmatrix} Y_m \\ y(m+1) \end{bmatrix}, \quad (9)$$

$$X_{m+1} = \begin{bmatrix} x_1(1) & \dots & x_n(1) \\ \vdots & \dots & \dots \\ x_1(m) & \dots & x_n(m) \\ \dots & \dots & \dots \\ x_1(m+1) & \dots & x_n(m) \end{bmatrix} = \begin{bmatrix} X_m \\ x^T(m+1) \end{bmatrix}.$$

The new least estimation equation is shown as follows:

$$\hat{\theta}(m+1) = (X_{m+1}^T X_{m+1})^{-1} X_{m+1}^T Y_{m+1}. \quad (10)$$

In order to get out of inverse matrix calculation of $X_{m+1}^T \cdot X_{m+1}$, $\hat{\theta}(m+1)$ is deduced as follows:

$$\begin{aligned} \hat{\theta}(m+1) &= \hat{\theta}(m) + P(m)x(m+1) \\ &\times [1 + x^T(m+1)P(m)x(m+1)]^{-1} \quad (11) \\ &\cdot [y(m+1) - x^T(m+1)\hat{\theta}(m)]. \end{aligned}$$

Equation (11) is the recursive least square (RLS) estimation. Recursive least square algorithm is called the generalization Kalman filter algorithm [45, 46]. It is the engineering realization method of the least square estimation theory [17].

From the calculation of least square estimation and recursive least square estimation, we can see that the past time of data has a big effect to future parameter estimation, and a large number of data calculations have occupied the system resources seriously. Therefore, it is difficult to be realized in embedded systems.

3. Windowed Least Square Estimation

Recursive least square algorithm can be used to real-time parameter estimation. However, the algorithm uses the past time of data, and the past time of data has the same importance as the present data in the algorithm. It weakened the importance of current data, caused a lot of system resources possession, and affected the estimation speed and precision [44]. In order to guarantee the instantaneity of parameter estimation, the paper adopts windowed least squares (WLS) to estimate the parameters of PMSM. The algorithm simulates the window processing function of communication signal. The time series data used for parameter estimation are added window handle, to reduce the calculation of estimation and system resources possession, making the algorithm easy for engineering realization.

Suppose that $P(m) = (X_m^T X_m)^{-1}$, the parameter estimation by $(x(k), y(k))$, $k = 1, 2, \dots, m$, is as follows:

$$\hat{\theta}(m) = P(m) X_m^T Y_m, \quad (12)$$

where

$$X_m = \begin{bmatrix} x^T(1) \\ x^T(2) \\ \vdots \\ x^T(m) \end{bmatrix}, \quad Y_m = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(m) \end{bmatrix}. \quad (13)$$

Consider increasing a group of data $(x(m+1), y(m+1))$, and then the parameter estimation is shown as follows [16, 23, 44]:

$$\begin{aligned} \bar{\theta}(m+1) &= (X_{m+1}^T X_{m+1})^{-1} X_{m+1}^T Y_{m+1} \\ &= \hat{\theta}(m) + P(m)x(m+1) \\ &\times [1 + x^T(m+1)P(m)x(m+1)]^{-1} \\ &\times [y(m+1) - x^T(m+1)\hat{\theta}(m)], \quad (14) \end{aligned}$$

$$\begin{aligned} \bar{P}(m+1) &= (X_{m+1}^T X_{m+1})^{-1} \\ &= P(m) - P(m)x(m+1) \\ &\times [1 + x^T(m+1)P(m)x(m+1)]^{-1} \\ &\times x^T(m+1)P(m), \end{aligned}$$

where

$$X_{m+1} = \begin{bmatrix} x^T(1) \\ x^T(2) \\ \vdots \\ x^T(m) \\ x^T(m+1) \end{bmatrix}, \quad Y_{m+1} = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(m) \\ y(m+1) \end{bmatrix}. \quad (15)$$

Therefore, $X_{m+1} = [x^T(1) \ X]$.

Consider eliminating a group of data $(x(1), y(1))$, and then the parameter estimation is shown as follows, where

$$P(m+1) = (X^T X)^{-1}, \quad X = \begin{bmatrix} x^T(2) \\ \vdots \\ x^T(m) \\ x^T(m+1) \end{bmatrix},$$

$$\begin{aligned}
& X_{m+1}^T X_{m+1} = X^T X + x(1) x^T(1), \\
P(m+1) &= \left(X_{m+1}^T X_{m+1} - x(1) x^T(1) \right)^{-1} \\
&= \left((\bar{P}(m+1))^{-1} + (-x(1)) x^T(1) \right)^{-1} \\
&= \bar{P}(m+1) - \bar{P}(m+1)(-x(1)) \\
&\quad \times [1 + x^T(1) \bar{P}(m+1)(-x(1))]^{-1} x^T(1) \bar{P}(m+1) \\
&= \bar{P}(m+1) + \bar{P}(m+1)x(1) \\
&\quad \times [1 - x^T(1) \bar{P}(m+1)x(1)]^{-1} x^T(1) \bar{P}(m+1). \tag{16}
\end{aligned}$$

Therefore, the estimation result is shown as follows:

$$\hat{\theta}(m+1) = (X^T X)^{-1} X^T Y = P(m+1) X^T Y, \tag{17}$$

where

$$\begin{aligned}
(x(k), y(k)), \quad k = 2, \dots, m+1, \quad Y = \begin{bmatrix} y(2) \\ \vdots \\ y(m) \\ y(m+1) \end{bmatrix}, \\
X_{m+1}^T Y_{m+1} = X^T Y + y(1)x(1), \tag{18}
\end{aligned}$$

$$\hat{\theta}(m+1)$$

$$\begin{aligned}
&= P(m+1) (X_{m+1}^T Y_{m+1} - y(1)x(1)) \\
&= [\bar{P}(m+1) + \bar{P}(m+1)x(1) \\
&\quad \times [1 - x^T(1) \bar{P}(m+1)x(1)]^{-1} \\
&\quad \times x^T(1) \bar{P}(m+1)] \\
&\quad \times [X_{m+1}^T Y_{m+1} - y(1)x(1)] \\
&= \bar{P}(m+1) X_{m+1}^T Y_{m+1} - y(1) \bar{P}(m+1)x(1) \\
&\quad \times [1 - x^T(1) \bar{P}(m+1)x(1)]^{-1} \\
&\quad \times x^T(1) \bar{P}(m+1)x(1) \\
&\quad - y(1) \bar{P}(m+1)x(1) + \bar{P}(m+1)x(1) \\
&\quad \times [1 - x^T(1) \bar{P}(m+1)x(1)]^{-1} \\
&\quad \times x^T(1) \bar{P}(m+1) X_{m+1}^T Y_{m+1} \\
&= \bar{\theta}(m+1) - \frac{y(1) \bar{P}(m+1)x(1)}{1 - x^T(1) \bar{P}(m+1)x(1)}
\end{aligned}$$

$$\begin{aligned}
&\times (x^T(1) \bar{P}(m+1)x(1) + 1 \\
&\quad - x^T(1) \bar{P}(m+1)x(1)) \\
&+ \bar{P}(m+1)x(1) [1 - x^T(1) \bar{P}(m+1)x(1)]^{-1} \\
&\times x^T(1) \bar{\theta}(m+1) \\
&= \bar{\theta}(m+1) - \frac{y(1) \bar{P}(m+1)x(1)}{1 - x^T(1) \bar{P}(m+1)x(1)} \\
&\quad + \frac{\bar{P}(m+1)x(1)x^T(1)\bar{\theta}(m+1)}{1 - x^T(1) \bar{P}(m+1)x(1)} \\
&= \bar{\theta}(m+1) + \frac{\bar{P}(m+1)x(1)}{1 - x^T(1) \bar{P}(m+1)x(1)} \\
&\quad \times (x^T(1)\bar{\theta}(m+1) - y(1)). \tag{19}
\end{aligned}$$

Therefore, for any moment

$$\begin{aligned}
\hat{\theta}(m+k) &= \bar{\theta}(m+k) \\
&\quad + \frac{\bar{P}(m+k)x(k)}{1 - x^T(k)\bar{P}(m+k)x(k)} \\
&\quad \times (x^T(k)\bar{\theta}(m+k) - y(k)),
\end{aligned} \tag{20}$$

where m is the window size, and

$$\begin{aligned}
P(m+k) &= \bar{P}(m+k) X_{m+k} \\
&\quad \times [1 - X_{m+k}^T \bar{P}(m+k) X_{m+k}]^{-1} X_{m+k}^T \bar{P}(m+k).
\end{aligned} \tag{21}$$

The window size is adjustable according to actual needs, based on the data length of regulation. This can guarantee the speed of calculation and can reduce the system resources possession too. At this time, the parameter estimation is related to the current m data sample, the past time of data has no effect on parameter estimation, and this can ensure the instantaneity and accuracy of the parameter estimation.

4. Simulations

4.1. PMSM Model. The voltage equations, flux linkage equations, and electromagnetic torque equations of PMSM in d, q frames are as follows [47, 48]:

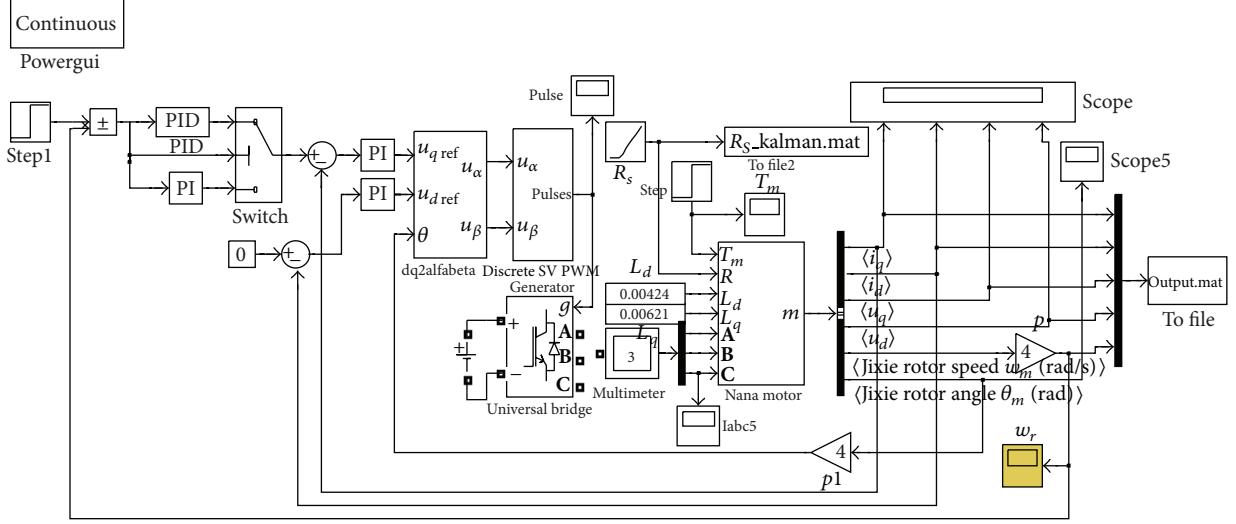


FIGURE 1: MATLAB motor simulation model.

$$\begin{aligned} u_q &= R_s i_q + L_q p i_q + \omega_r L_d i_d + \omega_r \psi_f, \\ u_d &= R_s i_d + L_d p i_d - \omega_r L_q i_q. \end{aligned} \quad (22)$$

In the steady state,

$$\begin{aligned} u_q &= R_s i_q + \omega_r L_d i_d + \omega_r \psi_f, \\ u_d &= R_s i_d - \omega_r L_q i_q. \end{aligned} \quad (23)$$

Flux linkage equations are as follows:

$$\begin{aligned} \psi_d &= L_d i_d + \psi_f \\ \psi_q &= L_q i_q, \end{aligned} \quad (24)$$

where u_q and u_d are voltages in q -axis and d -axis, respectively, i_q and i_d are currents in q -axis and d -axis, R_s is phase resistance of stator, L_d and L_q are inductances in d -axis and q -axis, ω_r is rotor velocity, ψ_f is flux linkage established by magnets, and p is the differential operator.

The mathematical model of PMSM is discretized to estimate parameters (R_s , L_d , and L_q). The discrete model of PMSM is as follows:

$$\begin{aligned} u_q(k) &= R_s i_q(k) + p \psi_q(k) + \omega_r \psi_d(k) \\ u_d(k) &= R_s i_d(k) + p \psi_d(k) + \omega_r \psi_q(k), \end{aligned} \quad (25)$$

where ψ_d and ψ_q are as follows:

$$\begin{aligned} \psi_d(k) &= L_d i_d(k) + \psi_f, \\ \psi_q(k) &= L_q i_q(k). \end{aligned} \quad (26)$$

The PMSM simulation model is established by MATLAB/SIMULINK, and the PMSM running data is obtained by the model. The simulation model is shown in Figure 1.

In MATLAB/SIMULINK, the R_s , L_d , and L_q parameters of PMSM are fixed in simulation. We cannot simulate the time varying characteristic of R_s , L_d , and L_q . Therefore, we design a motor simulation model according to the requirement. The R_s , L_d , and L_q can be changed flexibly in the simulation.

4.2. PMSM Parameters Estimation of Windowed Least Square Algorithm. When R_s , L_d , and L_q are fixed value, the PMSM simulation data is obtained by MATLAB. The windowed least square algorithm is used to identify the parameters. The algorithm with different window sizes is used for R_s , L_d , and L_q identification. From the identification result (Figures 2, 3 and 4), we can see that bigger window size has a better identification result. However, the window size does not obviously have an effect on the promotion of parameter identification precision when R_s , L_d , and L_q are fixed.

Different window sizes have big effect on the results of parameter estimation when the estimated parameters have time varying characteristic. The motor parameters R_s , L_d , and L_q are measured at 23.4°C, 30°C, 40°C, 50°C, 60°C, 70°C, and 80°C. Using piecewise linear method to simulate the time varying of the three parameters, the motor running data is obtained by MATLAB.

The windowed least square is used to identify the R_s , L_d , and L_q . The identification result is shown in Figures 5, 6, and 7 when the R_s , L_d and L_q are changed at the same time. From the Figures 5–7, we can see that shorter window size has lower effect on the promotion of identification precision. However, the window size is too big to improve the identification precision. The identification result is better when the window size is 300–400. In the motor model, ψ_f is bigger than

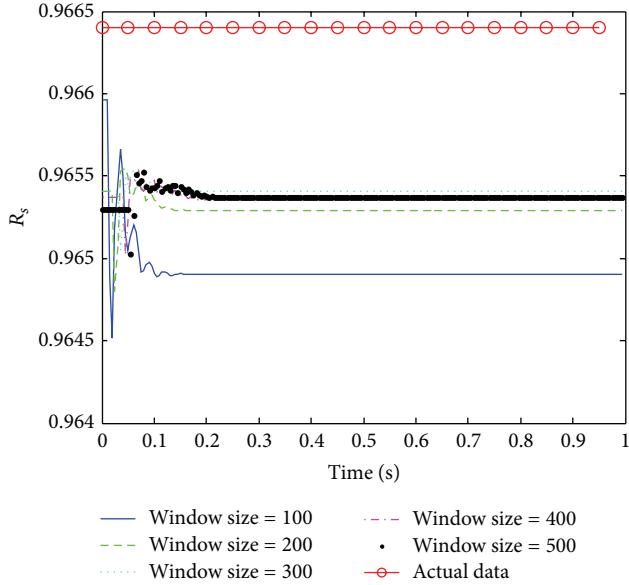


FIGURE 2: Estimation result of R_s of WLS with different window sizes when the parameters are fixed.

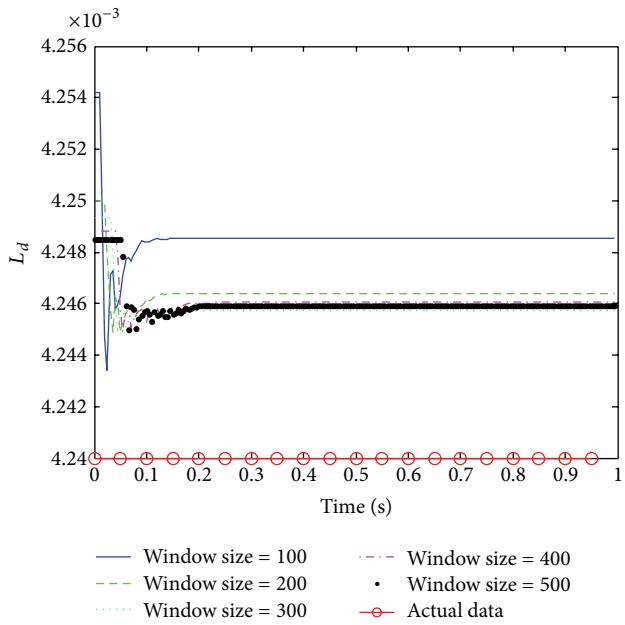


FIGURE 3: Estimation result of L_d of WLS with different window sizes when the parameters are fixed.

$L_d i_d$, so the change of L_d has little effect on the model output. Therefore, the identification result of L_d is not very well. However, the algorithm can also identify the parameter correctly.

4.3. PMSM Parameters Estimation of Extended Kalman Filter

The Kalman filter is a common parameter identification

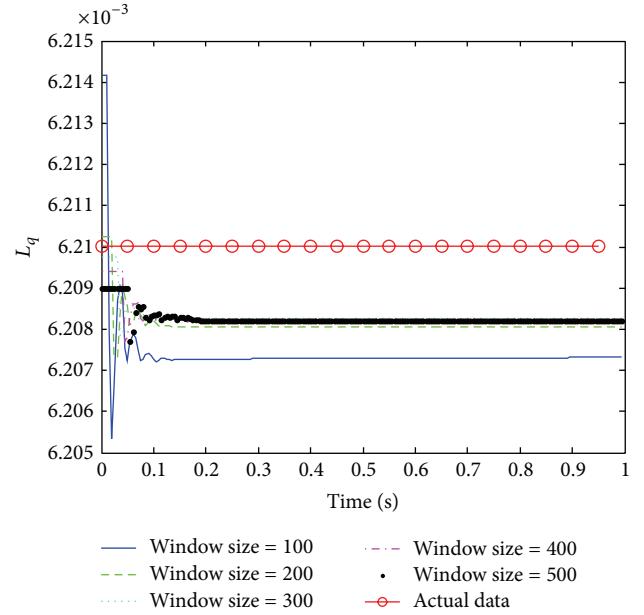
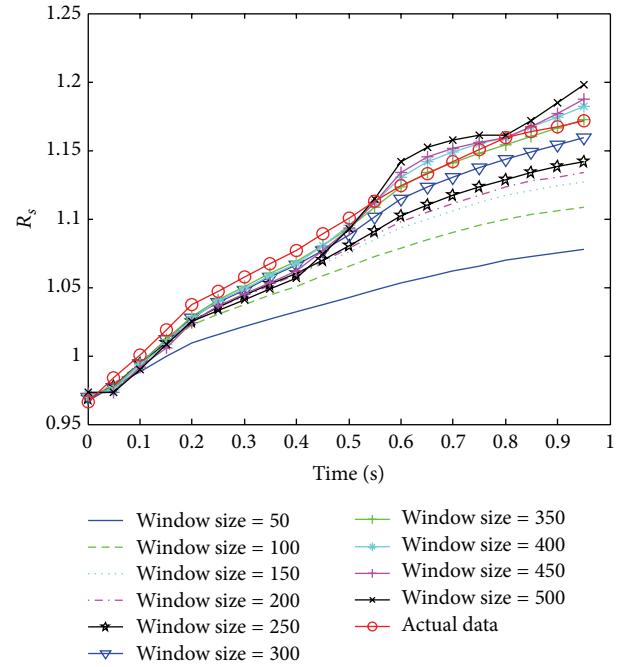


FIGURE 4: Estimation result of L_q of WLS with different window sizes when the parameters are fixed.



method. It is proposed in 1960 by Kalman [49]. The theory is applied to practical engineering immediately when it is put forward. The Apollo program and C-5 plane navigation system design are the most successful application examples. Extended Kalman filter (EKF) is an improved model of the Kalman filter which is one of the most widely applied in nonlinear system filter.

Discrete system state equation of EKF is

$$\begin{aligned}\mathbf{X}(k) &= \mathbf{A}(k-1)\mathbf{X}(k-1) + \mathbf{B}(k-1)\mathbf{U}(k-1) \\ &\quad + \mathbf{C}(k-1) + \mathbf{w}(k-1), \\ \mathbf{Z}(k) &= \mathbf{H}(k-1)\mathbf{X}(k) + \mathbf{v}(k),\end{aligned}\tag{27}$$

where

$$\begin{aligned}\mathbf{A}(k-1) &= \frac{\partial f(\mathbf{X}, \mathbf{U})}{\partial \mathbf{X}} T_s \Big|_{\mathbf{X}=\widehat{\mathbf{X}}(k-1)}, \\ \mathbf{B}(k-1) &= \frac{\partial f(\mathbf{X}, \mathbf{U})}{\partial \mathbf{U}} T_s \Big|_{\mathbf{U}=\mathbf{U}(k-1)},\end{aligned}$$

$$\begin{aligned}C(k-1) &= \left[f(\mathbf{X}, \mathbf{U}) T_s - \frac{\partial f(\mathbf{X}, \mathbf{U})}{\partial \mathbf{X}} \mathbf{X} T_s \right] \Big|_{\mathbf{X}=\widehat{\mathbf{X}}(k-1)}, \\ \mathbf{H}(k) &= \frac{\partial h(\mathbf{X})}{\partial \mathbf{X}} \Big|_{\mathbf{X}=\widehat{\mathbf{X}}(k-1)},\end{aligned}\tag{28}$$

where $\mathbf{X}(k)$ is the system state vector, $\mathbf{U}(k)$ is the system input vector, $\mathbf{Z}(k)$ is the system observation vector, $\mathbf{w}(k)$ is the system random noise vector, $\mathbf{v}(k)$ is the system random observation noise vector. $\mathbf{w}(k)$ and $\mathbf{v}(k)$ are noise sequences with zero mean, and the covariance matrices are $\mathbf{Q}(k)$ and $\mathbf{R}(k)$.

T_s is the sampling period. the discrete linear state space equation (29) of PMSM is established by discretization and linearization of the model (30):

$$\frac{d}{dt} = \begin{cases} i_d = \frac{u_d}{L_d} - \frac{R_s}{L_d} i_d + \frac{\psi_q}{L_d} \omega_e \\ i_q = \frac{u_q}{L_q} - \frac{R_s}{L_q} i_q - \frac{\psi_d}{L_q} \omega_e \\ \psi_d = 0 \\ \psi_q = 0 \\ R_s = 0, \end{cases}\tag{29}$$

$$\begin{bmatrix} i_d(k) \\ i_q(k) \\ \psi_d(k) \\ \psi_q(k) \\ R_s(k) \end{bmatrix} = \begin{bmatrix} 1 - \frac{\widehat{R}_s(k-1)}{L_d} T_s & 0 & 0 & \frac{\widehat{\omega}_e(k-1)}{L_d} T_s & -\frac{\widehat{i}_d(k-1)}{L_d} T_s \\ 0 & 1 - \frac{\widehat{R}_s(k-1)}{L_q} T_s & -\frac{\widehat{\omega}_e(k-1)}{L_q} T_s & 0 & -\frac{\widehat{i}_q(k-1)}{L_q} T_s \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}\tag{30}$$

$$\times \begin{bmatrix} i_d(k-1) \\ i_q(k-1) \\ \psi_d(k-1) \\ \psi_q(k-1) \\ R_s(k-1) \end{bmatrix} + \begin{bmatrix} \frac{T_s}{L_d} & 0 \\ 0 & \frac{T_s}{L_q} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_d(k-1) \\ u_q(k-1) \end{bmatrix} + T_s \begin{bmatrix} \frac{\widehat{i}_d(k-1) \widehat{R}_s(k-1)}{L_d} \\ \frac{\widehat{i}_q(k-1) \widehat{R}_s(k-1)}{L_q} \\ 0 \\ 0 \\ 0 \end{bmatrix} + \mathbf{w}(k-1),$$

$$\begin{bmatrix} i_d(k) \\ i_q(k) \\ \psi_d(k) \\ \psi_q(k) \\ R_s(k) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} i_d(k) \\ i_q(k) \\ \psi_d(k) \\ \psi_q(k) \\ R_s(k) \end{bmatrix} + \mathbf{v}(k).\tag{31}$$

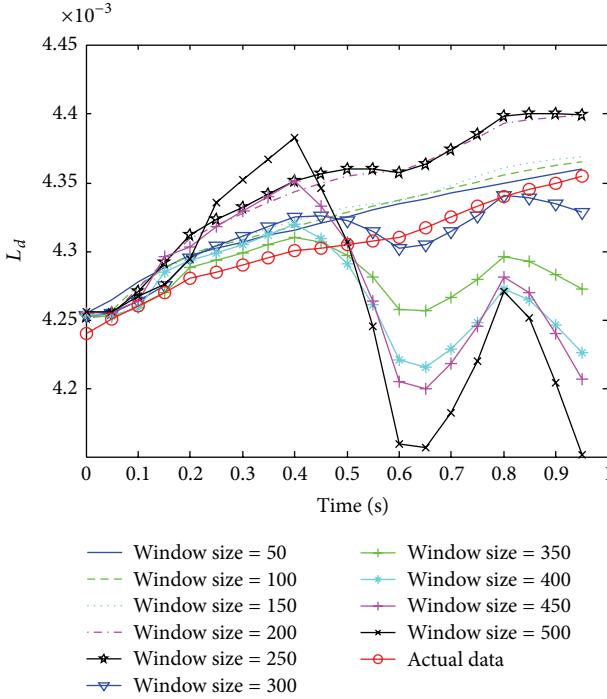


FIGURE 6: Estimation result of L_d of WLS with different window sizes when the parameters have time varying characteristic.

Using the simulation model and getting the motor output i_d , i_q , u_d , u_q , ω_e ; the initial values of P , Q , and R in EKF algorithm are

$$\begin{aligned} P &= \text{diag}([0.1 \ 0.1 \ 0.0004 \ 0.002 \ 0.02]), \\ Q &= \text{diag}([30 \ 15 \ 0.005 \ 0.03 \ 0.03]), \\ R &= \text{diag}([0.1 \ 0.02]). \end{aligned} \quad (32)$$

When R_s , L_d , and L_q of motor are fixed, the identification result table of EKF, recursive least square and windowed least square, algorithm is shown in Table 1.

From Table 1 we can see that identification result of EKF algorithm is as good as the windowed least square algorithm when R_s , L_d , and L_q are fixed. The comparison diagrams of identification result are shown in Figure 8 when the parameters are fixed.

EKF and recursive least square algorithms cannot achieve reasonable result when the parameters have time varying characteristic or have a drastic change. However, windowed least square algorithm can achieve good identification result when R_s , L_d , and L_q have time varying characteristic at the same time (Figures 5–7).

5. Analysis and Conclusion

Through the previous different PMSM parameters identification experiments, we can see the following.

- (1) When the parameters of PMSM have no time varying characteristic, three methods can achieve better

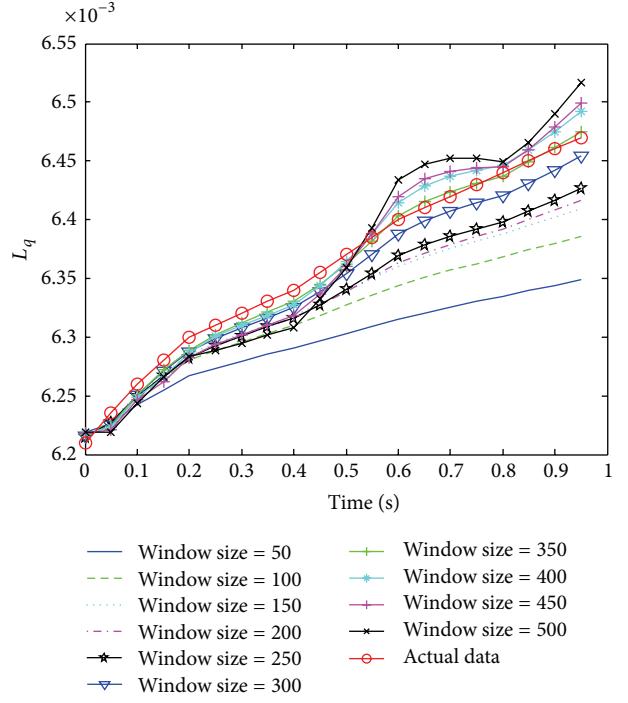


FIGURE 7: Estimation result of L_q of WLS with different window sizes when the parameters have time varying characteristic.

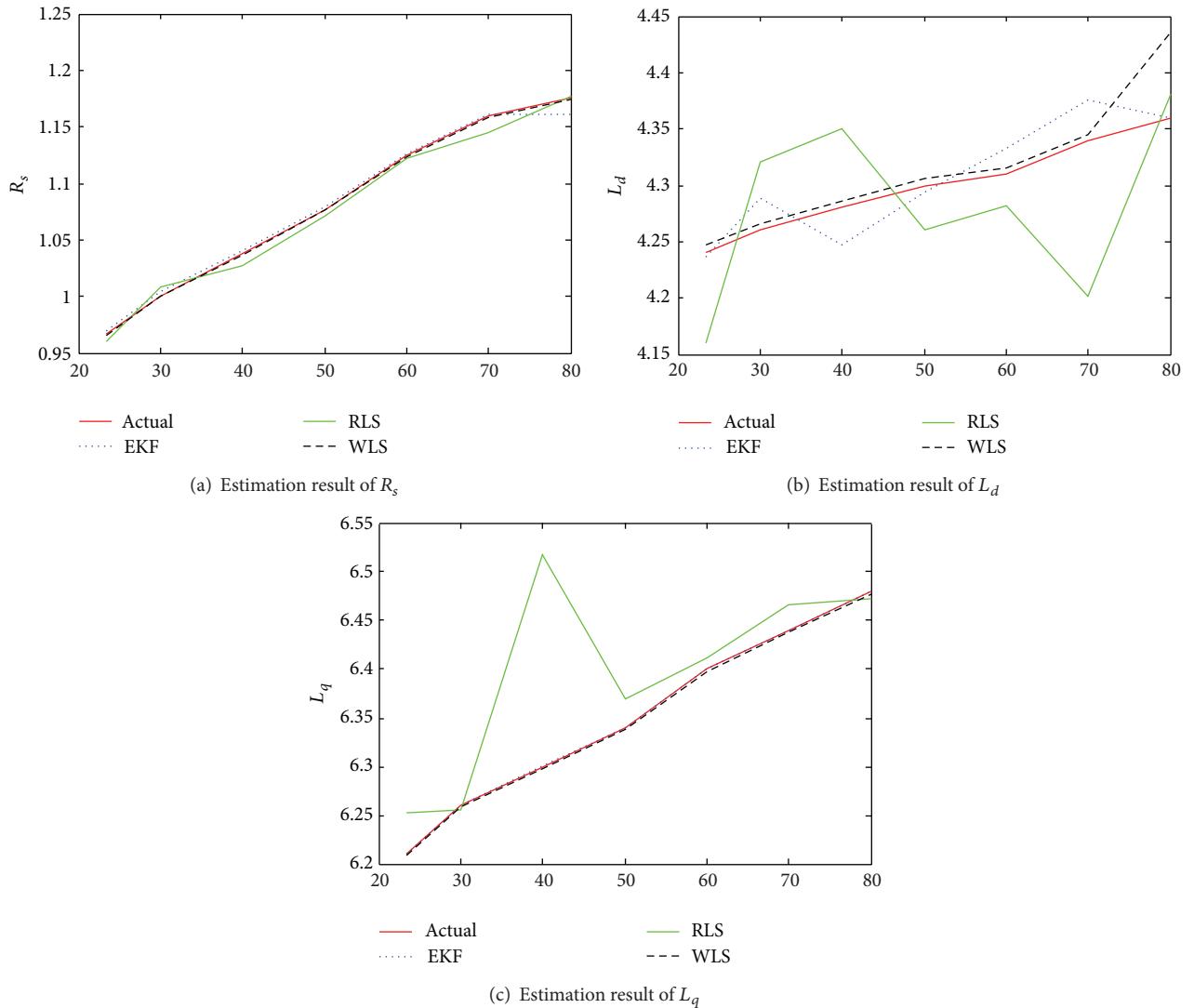
identification result in precision and accuracy. In the calculation and instantaneity of identification, recursive least square algorithm has a fatal flaw of data saturation, so the precision and accuracy of the algorithm are hard to guarantee. It is difficult to realize in embedded system the real-time parameters identification due to the amount of calculations and system resources possession of EKF [28]. The window size of windowed least square algorithm is flexible, so we can choose the collected data according to the changes of the parameters. It will reduce the influence of the past time of data to the current parameter identification, guarantee the accuracy and instantaneity of identification, and reduce the system resources possession at the same time.

- (2) When the parameters of PMSM have strong time varying characteristic, the EKF and recursive least square algorithms cannot guarantee the precision and accuracy of identification. However, windowed least square algorithm can get better identification result. Therefore, EKF and recursive least square algorithms are suitable for fixed parameters estimation or parameters with weak time varying characteristic identification. Windowed least square algorithm can get a good result both for fixed parameters and for time varying parameters identification.

Embedded technology is widely used in the motor driver and controller at present. However, the embedded chip (MCU, DSP, ARM, etc.) has certain restriction in computing speed and storage space. Therefore, windowed least square

TABLE 1: Estimation result comparison of EKF, RLS, and WLS when R_s , L_d , and L_q are fixed value.

Temperature (°C)		23.4	30	40	50	60	70	80
R_s	Actual data (Ω)	0.9664	1.0008	1.0373	1.0770	1.1245	1.1592	1.1751
	EKF	0.9694	1.0036	1.0400	1.0795	1.1266	1.1612	1.1611
	RLS	0.9607	1.0077	1.0271	1.0712	1.1216	1.1451	1.1771
$L_d \times 10^{-3}$	Actual data (mH)	4.24	4.26	4.28	4.30	4.31	4.34	4.36
	EKF	4.2368	4.2882	4.2475	4.2941	4.3331	4.3755	4.3595
	RLS	4.1602	4.3210	4.3506	4.2611	4.2814	4.2021	4.3807
$L_q \times 10^{-3}$	Actual data (mH)	6.21	6.26	6.30	6.34	6.40	6.44	6.48
	EKF	6.2095	6.2604	6.3001	6.3400	6.4012	6.4395	6.4800
	RLS	6.2531	6.2551	6.5172	6.3703	6.4118	6.4658	6.4727
	WLS	6.2089	6.2581	6.2982	6.3387	6.3981	6.4384	6.4780

FIGURE 8: Comparison diagram of EKF, RLS, and WLS when R_s , L_d , and L_q are fixed value.

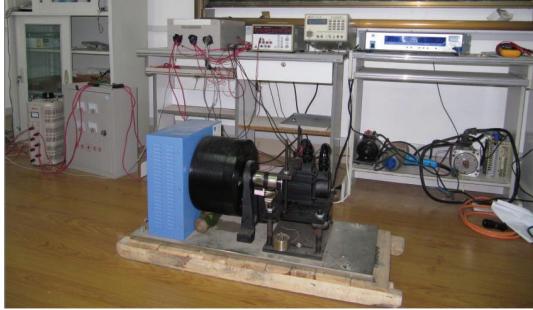


FIGURE 9: The PMSM experiment system.



FIGURE 10: The prototype DSP-based PMSM driver.

algorithm is a better choice for PMSM parameters identification of motor driver and controller. This paper is the beginning of work. There are a lot of work to do, such as transplant the algorithm to practical controller and control system (in Figures 9 and 10) which is designed to control the PMSM in practical application.

Acknowledgment

This paper is supported by the Shandong Province Science and Technology Development Plan of China (Grant no. 2011GGE27053).

References

- [1] M. A. Rahman and P. Zhou, "Analysis of brushless permanent magnet synchronous motors," *IEEE Transactions on Industrial Electronics*, vol. 43, no. 2, pp. 256–267, 1996.
- [2] M. Ooshima, A. Chiba, A. Rahman, and T. Fukao, "An improved control method of buried-type IPM bearingless motors considering magnetic saturation and magnetic pull variation," *IEEE Transactions on Energy Conversion*, vol. 19, no. 3, pp. 569–575, 2004.
- [3] K. Liu, Z. Q. Zhu, Q. Zhang, and J. Zhang, "Influence of nonideal voltage measurement on parameter estimation in permanent-magnet synchronous machines," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 6, pp. 2438–2447, 2012.
- [4] F. Caricchi, F. Crescimbini, and O. Honorati, "Low-cost compact permanent magnet machine for adjustable-speed pump application," *IEEE Transactions on Industry Applications*, vol. 34, no. 1, pp. 109–116, 1998.
- [5] P. Milanfar and J. H. Lang, "Monitoring the thermal condition of permanent-magnet synchronous motors," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 4, pp. 1421–1429, 1996.
- [6] T. Kailath, "An innovations approach to least-squares estimation—part I: linear filtering in additive white noise," *IEEE Transactions on Automatic Control*, vol. 13, pp. 646–655, 1968.
- [7] D. G. Robertson and J. H. Lee, "A least squares formulation for state estimation," *Journal of Process Control*, vol. 5, no. 4, pp. 291–299, 1995.
- [8] J. S. Gibson, G. H. Lee, and C. F. Wu, "Least-squares estimation of input/output models for distributed linear systems in the presence of noise," *Automatica*, vol. 36, no. 10, pp. 1427–1442, 2000.
- [9] S. Tunali and I. Batmaz, "Dealing with the least squares regression assumptions in simulation metamodeling," *Computers & Industrial Engineering*, vol. 38, no. 2, pp. 307–320, 2000.
- [10] R. M. Fernandez-Alcalá, J. Navarro-Moreno, and J. C. Ruiz-Molina, "Linear least-square estimation algorithms involving correlated signal and noise," *IEEE Transactions on Signal Processing*, vol. 53, no. 11, pp. 4227–4235, 2005.
- [11] V. Krätschmer, "Least-squares estimation in linear regression models with vague concepts," *Fuzzy Sets and Systems*, vol. 157, no. 19, pp. 2579–2592, 2006.
- [12] M. J. García-Ligero, A. Hermoso-Carazo, and J. Linares-Pérez, "Least-squares linear estimation of signals from observations with Markovian delays," *Journal of Computational and Applied Mathematics*, vol. 236, no. 2, pp. 234–242, 2011.
- [13] S. Ma, C. Quan, R. Zhu, C. J. Tay, L. Chen, and Z. Gao, "Application of least-square estimation in white-light scanning interferometry," *Optics and Lasers in Engineering*, vol. 49, no. 7, pp. 1012–1018, 2011.
- [14] Q. Wang and L. Zhang, "Least squares online linear discriminant analysis," *Expert Systems with Applications*, vol. 39, no. 1, pp. 1510–1517, 2012.
- [15] C. J. Demeure and L. L. Scharf, "Sliding windows and lattice algorithms for computing QR factors in the least squares theory of linear prediction," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 4, pp. 721–725, 1990.
- [16] K. Zhao, L. Fuyun, H. Lev-Ari, and J. G. Proakis, "Sliding window order-recursive least-squares algorithms," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 1961–1972, 1994.
- [17] H. Liu and Z. He, "A sliding-exponential window RLS adaptive filtering algorithm: properties and applications," *Signal Processing*, vol. 45, no. 3, pp. 357–368, 1995.
- [18] K. Yoo and H. Park, "Fast residual computation for sliding window recursive least squares methods," *Signal Processing*, vol. 45, no. 1, pp. 85–95, 1995.
- [19] Y. Xia, M. S. Kamel, and H. Leung, "A fast algorithm for AR parameter estimation using a novel noise-constrained least-squares method," *Neural Networks*, vol. 23, no. 3, pp. 396–405, 2010.
- [20] A. Aknouche, E. M. Al-Eid, and A. M. Hmeid, "Offline and online weighted least squares estimation of nonstationary power ARCH processes," *Statistics & Probability Letters*, vol. 81, no. 10, pp. 1535–1540, 2011.
- [21] L. Xie, H. Yang, and F. Ding, "Recursive least squares parameter estimation for non-uniformly sampled systems based on the data filtering," *Mathematical and Computer Modelling*, vol. 54, no. 1–2, pp. 315–324, 2011.
- [22] J. Oliver, R. Aravind, and K. M. M. Prabhu, "Improved least squares channel estimation for orthogonal frequency division multiplexing," *IET Signal Processing*, vol. 6, no. 1, pp. 45–53, 2012.

- [23] T. Sadiki, M. Triki, and D. T. M. Slock, "Window optimization issues in recursive least-squares adaptive filtering and tracking," in *Proceedings of the 38th IEEE Annual Asilomar Conference on Signals, Systems and Computers*, pp. 940–944, Pacific Grove, Calif, USA, November 2004.
- [24] G. Welch and G. Bishop, "An introduction to the Kalman filter," 1997.
- [25] P. J. Hargrave, "A tutorial introduction to Kalman filtering," in *Proceedings of the IEE Colloquium on Kalman Filters: Introduction, Applications and Future Developments*, pp. 1/1–1/6, 1989.
- [26] M. Gautier and P. Poignet, "Extended Kalman filtering and weighted least squares dynamic identification of robot," *Control Engineering Practice*, vol. 9, no. 12, pp. 1361–1372, 2001.
- [27] H. M. Al-Hamadi and S. A. Soliman, "Kalman filter for identification of power system fuzzy harmonic components," *Electric Power Systems Research*, vol. 62, no. 3, pp. 241–248, 2002.
- [28] T. Boileau, N. Leboeuf, B. Nahid-Mobarakeh, and F. Meibody-Tabar, "Online identification of PMSM parameters: parameter identifiability and estimator comparative study," *IEEE Transactions on Industry Applications*, vol. 47, no. 4, pp. 1944–1957, 2011.
- [29] A. Bechouche, H. Sediki, D. O. Abdeslam, and S. Haddad, "Identification of induction motor at standstill using artificial neural network," in *Proceedings of the 36th Annual Conference on IEEE Industrial Electronics Society (IECON '10)*, pp. 2908–2913, Glendale, Ariz, USA, 2010.
- [30] F. Alonge, F. D'Ippolito, and F. M. Raimondi, "Least squares and genetic algorithms for parameter identification of induction motors," *Control Engineering Practice*, vol. 9, no. 6, pp. 647–657, 2001.
- [31] S. Mishra, "A hybrid least square-fuzzy bacterial foraging strategy for harmonic estimation," *IEEE Transactions on Evolutionary Computation*, vol. 9, no. 1, pp. 61–73, 2005.
- [32] R. Liao, H. Zheng, S. Grzybowski, and L. Yang, "Particle swarm optimization-least squares support vector regression based forecasting model on dissolved gases in oil-filled power transformers," *Electric Power Systems Research*, vol. 81, no. 12, pp. 2074–2080, 2011.
- [33] R. A. Zadeh, A. Ghosh, and G. Ledwich, "Combination of Kalman filter and least-error square techniques in power system," *IEEE Transactions on Power Delivery*, vol. 25, no. 4, pp. 2868–2880, 2010.
- [34] S. Bolognani, R. Oboe, and M. Zigliotto, "Sensorless full-digital PMSM drive with EKF estimation of speed and rotor position," *IEEE Transactions on Industrial Electronics*, vol. 46, no. 1, pp. 184–191, 1999.
- [35] M. Haardt, "Structured least squares to improve the performance of ESPRIT-Type algorithms," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 792–799, 1997.
- [36] M. Ghogho, A. Swami, and A. K. Nandi, "Non-linear least squares estimation for harmonics in multiplicative and additive noise," *Signal Processing*, vol. 78, no. 1, pp. 43–60, 1999.
- [37] J. Angeby, "Estimating signal parameters using the nonlinear instantaneous least squares approach," *IEEE Transactions on Signal Processing*, vol. 48, no. 10, pp. 2721–2732, 2000.
- [38] J. F. Weng and S. H. Leung, "Nonlinear RLS algorithm for amplitude estimation in class a noise," *IEE Proceedings—Communications*, vol. 147, no. 2, pp. 81–86, 2000.
- [39] D. Zachariah, M. Sundin, M. Jansson, and S. Chatterjee, "Alternating least-squares for low-rank matrix reconstruction," *IEEE Signal Processing Letters*, vol. 19, no. 4, pp. 231–234, 2012.
- [40] R. Montoliu and F. Pla, "Generalized least squares-based parametric motion estimation," *Computer Vision and Image Understanding*, vol. 113, no. 7, pp. 790–801, 2009.
- [41] Z. Yingjie and G. Liling, "Improved moving least squares algorithm for directed projecting onto point clouds," *Measurement*, vol. 44, no. 10, pp. 2008–2019, 2011.
- [42] S. Seongwook, J.-S. Lim, S. J. Baek, and K.-M. Sung, "Variable forgetting factor linear least squares algorithm for frequency selective fading channel estimation," *IEEE Transactions on Vehicular Technology*, vol. 51, no. 3, pp. 613–616, 2002.
- [43] S. Morimoto, M. Sanada, and Y. Takeda, "Mechanical sensorless drives of IPMSM with online parameter identification," *IEEE Transactions on Industry Applications*, vol. 42, no. 5, pp. 1241–1248, 2006.
- [44] H. Sakai and H. Nakaoka, "Fast sliding window QRD-RLS algorithm," *Signal Processing*, vol. 78, no. 3, pp. 309–319, 1999.
- [45] S. Reece and S. Roberts, "An introduction to Gaussian processes for the Kalman filter expert," in *Proceedings of the 13th Conference on Information Fusion (FUSION '10)*, pp. 1–9, 2010.
- [46] A. Giorgano, F. M. Hsu, and J. Wiley, "Book review: least-square estimation with applications to digital signal processing," *IEE Proceedings F—Communications, Radar and Signal Processing*, vol. 132, no. 7, 1985.
- [47] S. Wang, R. Zhao, W. Chen, G. Li, and C. Liu, "Parameter identification of PMSM based on windowed least square algorithm," in *Proceedings of the Manufacturing Science and Technology (ICMST '11)*, pp. 5940–5944, Singapore, 2012.
- [48] S. Wang, S. Shi, C. Chen, G. Yang, and Z. Qu, "Identification of PMSM based on EKF and elman neural network," in *Proceedings of the IEEE International Conference on Automation and Logistics (ICAL '09)*, pp. 1459–1463, Shenyang, China, 2009.
- [49] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME—Journal of Basic Engineering D*, vol. 82, pp. 35–45, 1960.

Research Article

θ -Metric Space: A Generalization

Farshid Khojasteh,¹ Erdal Karapınar,² and Stojan Radenovic³

¹ Department of Mathematics, Arak Branch, Islamic Azad University, Arak, Iran

² Department of Mathematics, Atilim University, İncek, 06836 Ankara, Turkey

³ Faculty of Mechanical Engineering, University of Belgrade, Kraljice Marije 16, 11120 Belgrade, Serbia

Correspondence should be addressed to Erdal Karapınar; erdalkarapınar@yahoo.com

Received 27 May 2013; Accepted 25 July 2013

Academic Editor: Bashir Ahmad

Copyright © 2013 Farshid Khojasteh et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We introduce the notion of θ -metric as a generalization of a metric by replacing the triangle inequality with a more generalized inequality. We investigate the topology of the spaces induced by a θ -metric and present some essential properties of it. Further, we give characterization of well-known fixed point theorems, such as the Banach and Caristi types in the context of such spaces.

1. Introduction

Celebrated Banach contraction mapping principle [1] can be considered as a revolution in fixed point theory and hence in nonlinear functional analysis. The statement of this well-known principle is simple, but the consequences are so strong: every contraction mapping in a complete metric space has a unique fixed point. Since fixed point theory has been applied to different sciences and also distinct branches of mathematics, this pioneer result of Banach has been generalized, extended, and improved in various ways in several abstract spaces. After that, many authors stated many types, generalizations, and applications of fixed point theory until now (see also [2, 3]).

In 1976, Caristi [4] defined an order relation in a metric space by using a functional under certain conditions and proved a fixed point theorem for such an ordered metric space. Denote \mathbb{R} and \mathbb{N} as the sets of all real and natural numbers, respectively.

The order relation is defined as follows.

Lemma 1. Let (X, d) be a metric space and $\phi : X \rightarrow \mathbb{R}$ a functional. Define the relation " \leq " on X by

$$x \leq y \iff d(x, y) \leq \phi(x) - \phi(y). \quad (1)$$

Then, " \leq " is a partial order relation on X introduced by ϕ , and (X, d) is called an ordered metric space introduced by ϕ . Apparently, if $x \leq y$, then $\phi(x) \geq \phi(y)$.

Caristi's fixed point theorem states that a mapping $T : X \rightarrow X$ has a fixed point provided that (X, d) is a complete metric space and there exists a lower semicontinuous map $\phi : X \rightarrow \mathbb{R}$ such that

$$d(x, Tx) \leq \phi(x) - \phi(Tx), \quad \text{for every } x \in X. \quad (2)$$

This general fixed point theorem has found many applications in nonlinear analysis.

Many authors generalized Caristi's fixed point theorem and stated many types of it in complete metric spaces (see [5–8]). In particular, in 2010, Amini-Harandi [6] extended Caristi's fixed point and Takahashi's minimization theorems in complete metric space via the extension of partial ordered relation which is introduced in Lemma 1 and introduced some applications of such results.

One of the interesting generalizations of the notion of a metric is the concept of a fuzzy metric, given by Kramosil and Michálek [9], and Grabiec [10], independently. Later, George and Veeramani [11] investigated fuzzy metric structure and observed some important topological properties of such spaces. Furthermore, the authors [9–11] announced existence

and uniqueness of a fixed point of certain mappings in the framework of such spaces.

Definition 2 (see [11]). A binary operation $* : [0, 1] \times [0, 1] \rightarrow [0, 1]$ is a continuous t -norm if it is a topological monoid with unit 1 such that $a * b \leq c * d$ whenever $a \leq c$ and $b \leq d$ ($a, b, c, d \in [0, 1]$).

Definition 3 (see [9, 10]). The 3-tuple $(X, M, *)$ is said to be a fuzzy metric space if X is an arbitrary set, $*$ is a continuous t -norm, and $M : X \times X \times [0, +\infty) \rightarrow [0, 1]$ is a fuzzy set satisfying the following conditions:

- (1) $M(x, y, 0) = 0$,
- (2) $M(x, y, t) = 1$ for all $t > 0$ if and only if $x = y$,
- (3) $M(x, y, t) = M(y, x, t)$,
- (4) $M(x, y, t) * M(y, z, t) \leq M(x, z, t + s)$,
- (5) $M(x, y, \cdot) : [0, +\infty) \rightarrow [0, 1]$ is left continuous,

where $x, y, z \in X$ and $t, s > 0$.

In this paper, inspired from the definition of fuzzy metric spaces, we will introduce θ -metric as an extension of metric spaces which is obtained by replacing the triangle inequality with a more generalized inequality. We also investigate the topology of the θ -metric space and observe some fundamental properties of it. Furthermore, we give the characterization of the Banach and Caristi type fixed point theorems in the context of θ -metric space.

First, we give the following definition.

Definition 4. Let $\theta : [0, +\infty) \times [0, +\infty) \rightarrow [0, +\infty)$ be a continuous mapping with respect to each variable. Let $\text{Im}(\theta) = \{\theta(s, t) : s \geq 0, t \geq 0\}$. A mapping θ is called an B -action if and only if it satisfies the following conditions:

$$(I) \quad \theta(0, 0) = 0 \text{ and } \theta(t, s) = \theta(s, t) \text{ for all } t, s \geq 0,$$

(II)

$$\theta(s, t) < \theta(u, v) \quad \text{if } \begin{cases} \text{either } s < u, t \leq v, \\ \text{or } s \leq u, t < v, \end{cases} \quad (3)$$

$$(III) \quad \text{for each } r \in \text{Im}(\theta) \text{ and for each } s \in [0, r], \text{ there exists } t \in [0, r] \text{ such that } \theta(t, s) = r,$$

$$(IV) \quad \theta(s, 0) \leq s, \text{ for all } s > 0.$$

We denote by Υ the set of all B -actions.

Example 5. The following functions are examples of B -action:

$$(\theta_1) \quad \theta(t, s) = k(t + s), \text{ where } k \in (0, 1],$$

$$(\theta_2) \quad \theta(t, s) = k(t + s + ts), \text{ where } k \in (0, 1],$$

$$(\theta_3) \quad \theta(t, s) = ts/(1 + ts),$$

$$(\theta_4) \quad \theta(s, t) = \sqrt{s^2 + t^2},$$

$$(\theta_5) \quad \theta(t, s) = t + s + ts,$$

$$(\theta_6) \quad \theta(t, s) = t + s + \sqrt{ts},$$

$$(\theta_7) \quad \theta(t, s) = (t + s)(1 + ts).$$

Example 5 shows that the category of B -actions is uncountable. Next, we derive some lemmas which play a crucial role in our main results.

Lemma 6. Let

$$\Psi = \left\{ f : [0, +\infty) \longrightarrow [0, +\infty) : \begin{array}{l} f \text{ is continuous,} \\ f \text{ strictly increasing} \\ f(0) = 0 \\ f(t) < t \quad \forall t > 0. \end{array} \right\}. \quad (4)$$

Then, there exists a correspondence between Υ and Ψ . In other words, Υ is an infinite set.

Proof. For each $(t, s) \in [0, +\infty) \times [0, +\infty)$, define $\theta_f(t, s) = \lambda f(t + s)$, where $\lambda \in [0, 1)$ and $f \in \Psi$. It is previous that $\theta_f \in \Upsilon$. Now, define

$$\begin{aligned} H : \Psi &\longrightarrow \Upsilon \\ H(f) &= \theta_f. \end{aligned} \quad (5)$$

H is well-defined and injective function which completes the proof. \square

Lemma 7. Let θ be a B -action. For each $r \in \text{Im}(\theta)$ and $s \in B = [0, r]$, there exist $t \in [0, r]$ and a function $\eta : [0, +\infty) \times [0, +\infty) \rightarrow [0, +\infty)$ such that $\eta(r, s) = t$. Then, one derives the following.

$$(a_1) \quad \eta(0, 0) = 0.$$

$$(a_2) \quad \theta(\eta(r, s), s) = r \text{ and } \theta(r, \eta(s, r)) = s.$$

$$(a_3) \quad \eta \text{ is continuous with respect to the first variable.}$$

$$(a_4) \quad \text{If } \eta(r, s) \geq 0, \text{ then } 0 \leq s \leq r.$$

Proof. By (III) of Definition 4, for each $r \in \text{Im}(\theta)$ and for each $s \in [0, r]$, there exists $t \geq 0$ such that $\theta(t, s) = r$. Now, define $\eta(r, s) = t$. Let t and t' be two values such that $\eta(r, s) = t$ and $\eta(r, s) = t'$. If $t \neq t'$, then $t < t'$ or $t > t'$. If $t < t'$, then $r = \theta(t, s) < \theta(t', s) = r$ and this is a contradiction. For $t > t'$, we have the same argument. Thus, η is well defined.

On the other hand, (a_1) , (a_2) , and (a_4) are straightforward from (III) of Definition 4. Also, (a_3) holds since if $r \in \text{Im}(\theta)$ and $\{r_n\}$ is a sequence in $\text{Im}(\theta)$ such that $r_n \rightarrow r$, then $\theta(\eta(r_n, \cdot), \cdot) = r_n$. Thus,

$$\begin{aligned} \theta\left(\lim_{n \rightarrow \infty} \eta(r_n, \cdot), \cdot\right) &= \lim_{n \rightarrow \infty} \theta(\eta(r_n, \cdot), \cdot) \\ &= r = \theta(\eta(r, \cdot), \cdot). \end{aligned} \quad (6)$$

If $\lim_{n \rightarrow \infty} \eta(r_n, \cdot) > \eta(r, \cdot)$ or $\lim_{n \rightarrow \infty} \eta(r_n, \cdot) < \eta(r, \cdot)$, then by (II) of Definition 4 and (6), we conclude a contradiction. So, η is continuous with respect to the first variable. \square

The following definition arises from Lemma 7.

Definition 8. The function η , mentioned in Lemma 7, is called B -inverse action of θ . One says that θ is regular if η satisfies $\eta(r, r) = 0$, for each $r > 0$.

The set of all regular B -inverse actions will be denoted by Υ_r .

Example 9. Let $\theta_1(t, s) = t + s$ and $\theta_2(t, s) = \sqrt[n]{t^n + s^n}$. It is evident that $\eta_1(t, s) = t - s$ and $\eta_2(t, s) = \sqrt[n]{t^n - s^n}$. Furthermore, η_1, η_2 satisfy all the conditions of Lemma 7. Note that θ_1, θ_2 are regular.

Lemma 10. If $a \in X, c \in \text{Im}(\theta)$ and $b \in [0, c]$, then $\theta(a, b) \leq c$ implies that $a \leq \eta(c, b)$.

Proof. Suppose, on the contrary, that $\eta(c, b) < a$. Then, we have

$$c = \theta(\eta(c, b), b) < \theta(a, b), \quad (7)$$

a contradiction. \square

2. Main Results

In this section, we introduce the notion θ -metric and discuss the induced topology generated by d_θ on a nonempty subset X . In particular, we will show that τ_{d_θ} is Hausdorff and first countable. Further, we derive that (X, d_θ) is a metrizable topological space.

Definition 11. Let X be a nonempty set. A mapping $d_\theta : X \times X \rightarrow [0, +\infty)$ is called a θ -metric on X with respect to B -action $\theta \in Y$ if d_θ satisfies the following:

- (A1) $d_\theta(x, y) = 0$ if and only if $x = y$,
- (A2) $d_\theta(x, y) = d_\theta(y, x)$, for all $x, y \in X$,
- (A3) $d_\theta(x, y) \leq \theta(d_\theta(x, z), d_\theta(z, y))$, for all $x, y, z \in X$.

A pair (X, d_θ) is called a θ -metric space.

Remark 12. If (X, d_θ) is a θ -metric space, $\theta(s, t) = k(s + t)$, and $k \in (0, 1]$, then (X, d_θ) is a metric space.

Conversely, for $\theta(s, t) = k(s + t)$, $k \in (0, 1)$, we have that there exists metric space (X, d) which is not θ -metric space. For example, if $X = \{1, 2, 3\}$ and $d : X \times X \rightarrow [0, +\infty)$ defined by

$$\begin{aligned} d(1, 2) &= 1, & d(1, 3) &= 1, & d(2, 3) &= 2, \\ d(1, 1) &= 0, & d(2, 2) &= 0, & d(3, 3) &= 0, & k &= \frac{1}{2}, \end{aligned} \quad (8)$$

d is a metric, but d is not a θ metric.

Example 13. Let $X = \{x, y, z\}$ and $d_\theta : X \times X \rightarrow [0, +\infty)$ defined by

$$\begin{aligned} d_\theta(x, y) &= 2, & d_\theta(x, z) &= 6, & d_\theta(z, y) &= 10, \\ d_\theta(x, x) &= 0, & d_\theta(y, y) &= 0, & d_\theta(z, z) &= 0, \\ d_\theta(x, y) &= d_\theta(y, x), & d_\theta(x, z) &= d_\theta(z, x), \\ d_\theta(z, y) &= d_\theta(z, x). \end{aligned} \quad (9)$$

For $\theta(s, t) = s + t + st$, the function d_θ forms a metric, and hence, the pair (X, d_θ) is a θ -metric space.

Remark 14. Notice that d_θ , in Example 13, is not a metric on X , since we have $d_\theta(z, y) > d_\theta(x, y) + d_\theta(x, z)$.

Definition 15. Let (X, d_θ) be a θ -metric space. An open ball $B_{d_\theta}(x, r)$ at a center $x \in X$ with a radius $r \in \text{Im}(\theta)$ is defined as follows:

$$B_{d_\theta}(x, r) = \{y \in X : d_\theta(x, y) < r\}. \quad (10)$$

Lemma 16. Every open ball is an open set.

Proof. We show that, for each $x \in X$ and $r > 0$ and for each $y \in B_{d_\theta}(x, r)$, there exists $\delta > 0$ such that

$$B_{d_\theta}(y, \delta) \subset B_{d_\theta}(x, r). \quad (11)$$

By (III) of Definition 4, we can choose $\delta > 0$ such that $\theta(\delta, d_\theta(x, y)) = r$. Now, if $z \in B_{d_\theta}(y, \delta)$, then we have

$$d_\theta(z, x) \leq \theta(d_\theta(z, y), d_\theta(y, x)) < \theta(\delta, d_\theta(y, x)) = r. \quad (12)$$

It means that $z \in B_{d_\theta}(x, r)$ and (11) is proved. \square

Lemma 17. If (X, d_θ) is a θ -metric space, then the collection of open sets forms a topology, denoted by τ_{d_θ} . A pair (X, τ_θ) is called topological space induced by a θ -metric.

Lemma 18. The set $\{B_{d_\theta}(x, 1/n) : n \in \mathbb{N}\}$ is a local base at x , and the above topology is first countable.

Proof. For each $x \in X$ and $r > 0$, we can find $n_0 \in \mathbb{N}$ such that $1/n_0 < r$. Thus, $B_{d_\theta}(x, 1/n_0) \subset B_{d_\theta}(x, r)$. This means that $\{B_{d_\theta}(x, 1/n) : n \in \mathbb{N}\}$ is a local base at x and the above topology is first countable. \square

Theorem 19. A topological space (X, τ_{d_θ}) is Hausdorff.

Proof. Let x, y be two distinct points of X . Suppose that $0 < \alpha < d_\theta(x, y)$ is arbitrary. By Definition 4, we conclude that $\alpha \in \text{Im}(\theta)$. Therefore, there exist $r, s > 0$ such that $\theta(r, s) = \alpha$. It is clear that $B_{d_\theta}(x, r) \cap B_{d_\theta}(y, s) = \emptyset$. For if there exists $z \in B_{d_\theta}(x, r) \cap B_{d_\theta}(y, s)$, then

$$\begin{aligned} d_\theta(x, y) &\leq \theta(d_\theta(x, z), d_\theta(z, y)) \\ &< \theta(r, s) = \alpha < d_\theta(x, y), \end{aligned} \quad (13)$$

a contradiction. \square

Theorem 20. Let (X, d_θ) be a θ -metric space and τ_{d_θ} the topology induced by the θ -metric. Then, for a sequence $\{x_n\}$ in X , $x_n \rightarrow x$ if and only if $d_\theta(x_n, x) \rightarrow 0$ as $n \rightarrow \infty$.

Proof. Suppose that $x_n \rightarrow x$. Then, for each $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that $x_n \in B_{d_\theta}(x, \epsilon)$, for all $n \geq n_0$. Thus, $d_\theta(x_n, x) < \epsilon$; that is, $d_\theta(x_n, x) \rightarrow 0$ as $n \rightarrow \infty$. The converse is verified easily. \square

Theorem 21. Let (X, d_θ) be a θ -metric space and $x_n \rightarrow x$, $y_n \rightarrow y$. Then,

$$d_\theta(x_n, y_n) \rightarrow d_\theta(x, y). \quad (14)$$

Proof. For each $n \in \mathbb{N}$, there exists $K > 0$ such that, for all $n \geq K$,

$$d_\theta(x_n, x) < \frac{1}{n}, \quad d_\theta(y_n, y) < \frac{1}{n}. \quad (15)$$

Thus, by the continuity of θ with respect to each variable, we have

$$\begin{aligned} d_\theta(x, y) &\leq \theta(d_\theta(x, x_n), \theta(d_\theta(x_n, y_n), d_\theta(y_n, y))) \\ &< \theta\left(\frac{1}{n}, \theta\left(d_\theta(x_n, y_n), \frac{1}{n}\right)\right). \end{aligned} \quad (16)$$

Therefore,

$$\begin{aligned} d_\theta(x, y) &\leq \lim_{n \rightarrow \infty} \theta\left(\frac{1}{n}, \theta\left(d_\theta(x_n, y_n), \frac{1}{n}\right)\right) \\ &= \theta\left(0, \theta\left(\lim_{n \rightarrow \infty} d_\theta(x_n, y_n), 0\right)\right) \\ &\leq \lim_{n \rightarrow \infty} d_\theta(x_n, y_n) \\ &\leq \lim_{n \rightarrow \infty} \theta(d_\theta(x, x_n), \theta(d_\theta(x, y), d_\theta(y_n, y))) \\ &\leq \lim_{n \rightarrow \infty} \theta\left(\frac{1}{n}, \theta\left(d_\theta(x, y), \frac{1}{n}\right)\right) \\ &\leq d_\theta(x, y). \end{aligned} \quad (17)$$

Thus, $d_\theta(x_n, y_n) \rightarrow d_\theta(x, y)$. \square

Lemma 22. Let (X, d_θ) be a θ -metric space. Let $\{x_n\}$ be a sequence in X and $x_n \rightarrow x$. Then, x is unique.

Proof. Suppose that $x_n \rightarrow x$ and $x_n \rightarrow y$. We show that $x = y$. For each $n \in \mathbb{N}$, there exists $N > 0$ such that $d_\theta(x_n, x) < 1/n$ and $d_\theta(x_n, y) < 1/n$. By the continuity of θ , we have

$$\begin{aligned} 0 &\leq d_\theta(x, y) \leq \theta(d_\theta(x_n, x), d_\theta(x_n, y)) \\ &< \theta\left(\frac{1}{n}, \frac{1}{n}\right) \rightarrow 0, \quad (n \rightarrow \infty). \end{aligned} \quad (18)$$

Hence, we have $x = y$. \square

Definition 23. Let (X, d_θ) be a θ -metric space. Then, for a sequence $\{x_n\}$ in X , one says that $\{x_n\}$ is a Cauchy sequence if for each $\epsilon > 0$, there exists $N > 0$ such that, for all $m \geq n \geq N$, $d_\theta(x_n, x_m) < \epsilon$.

Definition 24. Let (X, d_θ) be a θ -metric space. One says that (X, d_θ) is complete θ -metric space if every Cauchy sequence $\{x_n\}$ is convergent in X .

Lemma 25 (see [12]). A Hausdorff topological space (X, τ_{d_θ}) is metrizable if and only if it admits a compatible uniformity with a countable base.

In the following theorems we apply the previous lemma and the concept of uniformity (see [12] for more information) to prove the metrizability of a topological space (X, τ_{d_θ}) .

Theorem 26. Let (X, d_θ) be a θ -metric space. Then, (X, τ_{d_θ}) is a metrizable topological space.

Proof. For each $n \in \mathbb{N}$, define

$$\mathcal{U}_n = \{(x, y) \in X \times X : d_\theta(x, y) < \frac{1}{n}\}. \quad (19)$$

We will prove that $\{\mathcal{U}_n : n \in \mathbb{N}\}$ is a base for a uniformity \mathcal{U} on X whose induced topology coincides with τ_{d_θ} .

We first note that for each $n \in \mathbb{N}$,

$$\{(x, x) : x \in X\} \subseteq \mathcal{U}_n, \quad \mathcal{U}_{n+1} \subseteq \mathcal{U}_n, \quad \mathcal{U}_n = \mathcal{U}_n^{-1}. \quad (20)$$

On the other hand, for each $n \in \mathbb{N}$, there is, by the continuity of θ , an $m \in \mathbb{N}$ such that

$$m > 2n, \quad \theta\left(\frac{1}{m}, \frac{1}{m}\right) < \frac{1}{n}. \quad (21)$$

Then, $\mathcal{U}_m \circ \mathcal{U}_m \subseteq \mathcal{U}_n$: indeed, let $(x, y) \in \mathcal{U}_m$ and $(y, z) \in \mathcal{U}_m$. Thus,

$$\begin{aligned} d_\theta(x, z) &\leq \theta(d_\theta(x, y), d_\theta(y, z)) \\ &< \theta\left(\frac{1}{m}, \frac{1}{m}\right) < \frac{1}{n}. \end{aligned} \quad (22)$$

Therefore, $(x, z) \in \mathcal{U}_n$. Hence, $\{\mathcal{U}_n : n \in \mathbb{N}\}$ is a base for a uniformity \mathcal{U} on X . Since for each $x \in X$ and each $n \in \mathbb{N}$, $\mathcal{U}_n(x) = \{y \in X : d_\theta(x, y) < 1/n\}$. We deduce from Lemma 25 that (X, τ_{d_θ}) is a metrizable topological space. \square

Let us recall that a metrizable topological space (X, τ) is said to be completely metrizable if it admits a complete metric [13].

Theorem 27. Let (X, d_θ) be a complete θ -metric space. Then, (X, τ_{d_θ}) is completely metrizable.

Proof. It follows from the proof of Theorem 26 that $\{\mathcal{U}_n : n \in \mathbb{N}\}$ is a base for a uniformity \mathcal{U} on X compatible with τ_{d_θ} , where $\mathcal{U}_n = \{(x, y) \in X \times X : d_\theta(x, y) < 1/n\}$, for every $n \in \mathbb{N}$. Then, there exists a metric d on X whose induced uniformity coincides with \mathcal{U} . We want to show that the metric d is complete. Indeed, given a Cauchy sequence $\{x_n\}$ in (X, d) , we will prove that $\{x_n\}$ is a Cauchy sequence in (X, d) . To this end, fix $\epsilon > 0$. Choose $k \in \mathbb{N}$ such that $1/k < \epsilon$. Then, there exists $n_0 \in \mathbb{N}$ such that $(x_m, x_n) \in \mathcal{U}_k$ for every $n, m \geq n_0$. Consequently, for each $n, m \geq n_0$, $d_\theta(x_n, x_m) \leq 1/k < \epsilon$. We have shown that $\{x_n\}$ is a Cauchy sequence in the complete θ -metric space (X, d) and so is convergent with respect to (X, d) . Thus, (X, d) is a complete metric space. \square

3. Two Fixed Point Theorems

In this section, we introduce two fixed point theorems in θ -metric spaces. First, we introduce the Banach fixed point and Caristi's fixed point theorems in such spaces.

3.1. Banach Fixed Point Theorem

Theorem 28. Let (X, d_θ) be a complete θ -metric space and $f : X \rightarrow X$ a mapping that satisfies the following:

$$d_\theta(fx, fy) \leq \alpha d_\theta(x, y), \quad (23)$$

for each $x, y \in X$, where $\alpha \in [0, 1)$. Then, f has a unique fixed point.

Proof. Let $x_0 \in X$ and $x_{n+1} = fx_n$. We divide our proof into four steps.

Step 1. We claim that $d_\theta(x_n, x_{n+1}) \rightarrow 0$. Indeed, we have

$$\begin{aligned} d_\theta(x_{n+1}, x_n) &\leq \alpha d_\theta(x_n, x_{n-1}) \\ &\leq \alpha^2 d_\theta(x_{n-1}, x_{n-2}) \\ &\vdots \\ &\leq \alpha^n d_\theta(x_1, x_0) \rightarrow 0, \quad \text{as } (n \rightarrow \infty). \end{aligned} \quad (24)$$

Thus, we have $d_\theta(x_n, x_{n+1}) \rightarrow 0$.

Step 2. We assert that the sequence $\{x_n\}$ is bounded. Suppose, on the contrary, that $\{x_n\}$ is an unbounded sequence. Thus, there exists subsequence $\{n(k)\}$ such that $n(1) = 1$ and for each $k \in \mathbb{N}$, $n(k+1)$ is minimal in the sense that the relation

$$d_\theta(x_{n(k+1)}, x_{n(k)}) > 1 \quad (25)$$

does not hold and

$$d_\theta(x_m, x_{n(k)}) \leq 1 \quad (26)$$

holds for all $m \in \{n(k) + 1, n(k) + 2, \dots, n(k+1) - 1\}$. Hence, by using the triangle inequality, we derive that

$$\begin{aligned} 1 &< d_\theta(x_{n(k+1)}, x_{n(k)}) \\ &\leq \theta(d_\theta(x_{n(k+1)}, x_{n(k+1)-1}), d_\theta(x_{n(k+1)-1}, x_{n(k)})) \\ &\leq \theta(d_\theta(x_{n(k+1)}, x_{n(k+1)-1}), 1). \end{aligned} \quad (27)$$

By taking the limit from two sides of (27) and using (II) of Definition 4, we derive that

$$d_\theta(x_{n(k+1)}, x_{n(k)}) \rightarrow 1^+ \quad \text{as } (k \rightarrow +\infty). \quad (28)$$

Also, we have

$$\begin{aligned} 1 &< d_\theta(x_{n(k+1)}, x_{n(k)}) \\ &\leq d_\theta(x_{n(k+1)-1}, x_{n(k)-1}) \\ &\leq \theta(d_\theta(x_{n(k+1)-1}, x_{n(k)}), d_\theta(x_{n(k)}, x_{n(k)-1})) \\ &\leq \theta(1, d_\theta(x_{n(k)}, x_{n(k)-1})), \end{aligned} \quad (29)$$

which implies that

$$d_\theta(x_{n(k+1)-1}, x_{n(k)-1}) \rightarrow 1^+ \quad \text{as } (k \rightarrow +\infty). \quad (30)$$

Since $d_\theta(x_{n(k+1)}, x_{n(k)}) \leq \alpha d_\theta(x_{n(k+1)-1}, x_{n(k)-1})$, we have $1 \leq \alpha 1$, a contradiction. Thus, the sequence $\{x_n\}$ is bounded.

Step 3. We will show that $\{x_n\}$ is a Cauchy sequence. Let $m, n \in \mathbb{N}$ with $m > n$

$$\begin{aligned} d_\theta(x_m, x_n) &= d_\theta(fx_{m-1}, fx_{n-1}) \\ &\leq \alpha d_\theta(fx_{m-2}, fx_{n-2}) \\ &\vdots \\ &\leq \alpha^n d_\theta(fx_{m-n}, x_0). \end{aligned} \quad (31)$$

Since $\{x_n\}$ is a bounded sequence, therefore, $\lim_{n,m \rightarrow \infty} d_\theta(x_m, x_n) = 0$; that is, $\{x_n\}$ is a Cauchy sequence. Thus, there exists $x \in X$ such that $x_n \rightarrow x$. Further, we derive that

$$\begin{aligned} d_\theta(x_{n+1}, fx) &= d_\theta(fx_n, fx) \\ &\leq \alpha d_\theta(x_n, x) \rightarrow 0, \quad (n \rightarrow \infty). \end{aligned} \quad (32)$$

It means that $x_{n+1} \rightarrow fx$; that is, $fx = x$.

Step 4. In the last step, we prove that the x is the unique fixed point of T . Suppose, on the contrary, that x, y are two distinct fixed points of f . So, we get that

$$d_\theta(y, x) = d_\theta(fy, fx) \leq \alpha d_\theta(y, x) \quad (33)$$

is a contradiction. This completes the proof. \square

3.2. Caristi-Type Fixed Point Theorem

Definition 29. Suppose that (X, d_θ) be a complete θ -metric space and \mathcal{P} the class of all maps $\psi : X \times X \rightarrow [0, +\infty)$ which satisfies the following conditions:

- (E₁) there exists $\hat{x} \in X$ such that $\psi(\hat{x}, \cdot)$ is bounded below and lower semicontinuous, and $\psi(\cdot, y)$ is upper semicontinuous for each $y \in X$,
- (E₂) $\psi(x, x) = 0$, for each $x \in X$,
- (E₃) $\theta(\psi(x, y), \psi(y, z)) \leq \psi(x, z)$, for each $x, y, z \in X$.

Lemma 30. By Definition 29, one has

$$\psi(x, y) \leq \eta(\psi(x, z), \psi(y, z)), \quad (34)$$

for each $x, y, z \in X$.

Proof. By Lemma 10, we obtain the desired result. \square

Example 31. Let $\theta(t, s) = ts/(1+ts)$; thus, $\text{Im}(\theta) = [0, 1)$. Now, let $\phi : X \rightarrow \mathbb{R}$ be a lower bounded, lower semicontinuous function and

$$\psi(x, y) = \begin{cases} e^{\phi(y)-\phi(x)} & x \neq y \\ 0 & x = y. \end{cases} \quad (35)$$

Then, ψ satisfies all conditions of Definition 29.

Example 32. Let $\theta(t, s) = \sqrt[2n+1]{t^{2n+1} + s^{2n+1}}$; thus, $\text{Im}(\theta) = [0, +\infty)$. Now, let $\phi : X \rightarrow \mathbb{R}$ be a lower bounded, lower semicontinuous function and

$$\psi(x, y) = \sqrt[2n+1]{\phi(y) - \phi(x)}. \quad (36)$$

Then, ψ satisfies all conditions of Definition 29. Also, $\eta(t, s) = \sqrt[2n+1]{t^{2n+1} - s^{2n+1}}$, and θ is regular.

From now to end, we assume that θ is regular (see Definition 8).

Lemma 33. Let (X, d_θ) be a complete θ -metric space and $\psi \in \mathcal{P}$. Let $\gamma : [0, +\infty) \rightarrow [0, +\infty)$ be θ -subadditive; that is, $\gamma(\theta(x, y)) \leq \theta(\gamma(x), \gamma(y))$, for each $x, y \in [0, +\infty)$, nondecreasing continuous map such that $\gamma^{-1}\{0\} = \{0\}$. Define the order \prec on X by

$$x \prec y \iff \gamma(d_\theta(x, y)) \leq \psi(x, y), \quad (37)$$

for any $x, y \in X$. Then, (X, \prec) is a partial order set which has minimal elements.

Proof. At first, we show that (X, \prec) is a partial ordered set. For each $x \in X$, we have $0 = \gamma(0) = \gamma(d_\theta(x, x)) \leq \psi(x, x) = 0$. Thus, $x \prec x$. If $x \prec y$ and $y \prec x$, then $\gamma(d_\theta(x, y)) \leq \psi(x, y)$ and $\gamma(d_\theta(x, y)) \leq \psi(y, x)$. Thus, we give

$$\begin{aligned} \gamma(\theta(d_\theta(x, y), d_\theta(x, y))) &\leq \theta(\gamma(d_\theta(x, y)), \gamma(d_\theta(x, y))) \\ &\leq \theta(\psi(x, y), \psi(y, x)) \\ &\leq \psi(x, x) = 0. \end{aligned} \quad (38)$$

It means that $x = y$. Finally, if $x \prec y$ and $y \prec z$, then $\gamma(d_\theta(x, y)) \leq \psi(x, y)$ and $\gamma(d_\theta(y, z)) \leq \psi(y, z)$. Thus, we give

$$\begin{aligned} \gamma(d_\theta(x, z)) &\leq \gamma(\theta(d_\theta(x, y), d_\theta(y, z))) \\ &\leq \theta(\gamma(d_\theta(x, y)), \gamma(d_\theta(y, z))) \\ &\leq \theta(\psi(x, y), \psi(y, z)) \\ &\leq \psi(x, z). \end{aligned} \quad (39)$$

It means that $x \prec z$. Thus, (X, \prec) is a partial ordered set.

To show that (X, \prec) has minimal elements, we show that any decreasing chain has a lower bound. Indeed, let $\{x_\alpha\}_{\alpha \in \Gamma}$ be a decreasing chain; then we have

$$\begin{aligned} 0 &\leq \gamma(d_\theta(x_\alpha, x_\beta)) \\ &\leq \psi(x_\alpha, x_\beta) \\ &\leq \eta(\psi(\hat{x}, x_\alpha), \psi(\hat{x}, x_\beta)); \end{aligned} \quad (40)$$

by definition of η , we have $\psi(\hat{x}, x_\alpha) \leq \psi(\hat{x}, x_\beta)$. Thus, $\{\psi(\hat{x}, x_\alpha)\}_{\alpha \in \Gamma}$ is decreasing net of reals which is bounded

below. Let $\{\alpha_n\}$ be an increasing sequence of elements from Γ such that

$$\lim_{n \rightarrow \infty} \psi(\hat{x}, x_{\alpha_n}) = \inf \{\psi(\hat{x}, x_\alpha) : \alpha \in \Gamma\} = \rho. \quad (41)$$

Then, for each $m \geq n$, we infer that

$$\begin{aligned} \gamma(d_\theta(x_{\alpha_n}, x_{\alpha_m})) &\leq \psi(x_{\alpha_n}, x_{\alpha_m}) \\ &\leq \eta(\psi(\hat{x}, x_{\alpha_n}), \psi(\hat{x}, x_{\alpha_m})). \end{aligned} \quad (42)$$

By taking limit from two sides of (42), the regularity of θ , and continuity of η , we give

$$\begin{aligned} \limsup_{n, m \rightarrow \infty} \gamma(d_\theta(x_{\alpha_n}, x_{\alpha_m})) &\leq \psi(x_{\alpha_n}, x_{\alpha_m}) \\ &\leq \limsup_{n, m \rightarrow \infty} \eta(\psi(\hat{x}, x_{\alpha_n}), \psi(\hat{x}, x_{\alpha_m})) \\ &\leq \eta(\rho, \rho) = 0. \end{aligned} \quad (43)$$

Then, our assumption on γ implies that $\{x_{\alpha_n}\}$ is a Cauchy sequence and therefore converges to some $x \in X$. Since γ is continuous and $\psi(\cdot, x_{\alpha_n})$ is upper semicontinuous, then we have

$$\begin{aligned} \gamma(d_\theta(x, x_{\alpha_n})) &= \limsup_{m \rightarrow \infty} \gamma(d_\theta(x_{\alpha_m}, x_{\alpha_n})) \\ &\leq \limsup_{m \rightarrow \infty} \psi(x_{\alpha_m}, x_{\alpha_n}) \\ &\leq \psi(x, x_{\alpha_n}). \end{aligned} \quad (44)$$

This shows that $x \prec x_{\alpha_n}$ for all $n \geq 1$, which means that x is lower bound for $\{x_{\alpha_n}\}$. In order to see that x is also a lower bound for $\{x_\alpha\}_{\alpha \in \Gamma}$, let $\beta \in \Gamma$ be such that $x_\beta \prec x_{\alpha_n}$ for all $n \geq 1$. Then, for each $n \in \mathbb{N}$, we have

$$\begin{aligned} 0 &\leq \gamma(d_\theta(x_\beta, x_{\alpha_n})) \\ &\leq \psi(x_\beta, x_{\alpha_n}) \\ &\leq \eta(\psi(\hat{x}, x_{\alpha_n}), \psi(\hat{x}, x_\beta)). \end{aligned} \quad (45)$$

Hence, for all $n \geq 1$,

$$\psi(\hat{x}, x_\beta) \leq \psi(\hat{x}, x_{\alpha_n}) \quad (46)$$

which implies that

$$\begin{aligned} \psi(\hat{x}, x_\beta) &= \inf \{\psi(\hat{x}, x_\alpha) : \alpha \in \Gamma\} \\ &= \lim_{n \rightarrow \infty} \psi(\hat{x}, x_{\alpha_n}). \end{aligned} \quad (47)$$

Thus, from (46), we get $\lim_{n \rightarrow \infty} x_{\alpha_n} = x_\beta$, which implies that $x_\beta = x$. Therefore, for any $\alpha \in \Gamma$, there exists $n \in \mathbb{N}$ such that $x_{\alpha_n} \prec x_\alpha$; that is, x is a lower bound of $\{x_\alpha\}$. Zorn's lemma will therefore imply that (X, \prec) has minimal elements. \square

Theorem 34. Let (X, d_θ) be a complete θ -metric space and $\psi \in \mathcal{P}$. Let $\gamma : [0, +\infty) \rightarrow [0, +\infty)$ be as in Lemma 33. Let $T : X \rightarrow X$ be a map satisfying the following:

$$\gamma(d_\theta(x, Tx)) \leq \psi(Tx, x), \quad (48)$$

for any $x \in X$. Then, T has a fixed point.

Proof. By Lemma 33, (X, \prec) has a minimal element say \bar{x} . Thus, $T\bar{x} \prec \bar{x}$. It means that $T\bar{x} = \bar{x}$. \square

Corollary 35. Let (X, d_θ) be a complete θ -metric space and $\psi \in \mathcal{P}$. Let $\gamma : [0, +\infty) \rightarrow [0, +\infty)$ be as in Lemma 33. Let $T : X \rightarrow 2^X$ be a multivalued mapping satisfying the following:

$$\gamma(d_\theta(x, y)) \leq \psi(y, x), \quad \forall y \in Tx. \quad (49)$$

Then, T has an endpoint; that is, there exists $\bar{x} \in X$ such that $T\bar{x} = \{\bar{x}\}$.

In Corollary 35, we can introduce many types of Caristi's fixed point theorem as follows.

If we set ψ as in Example 32, then (48) has the following form:

$$\gamma(d_\theta(x, Tx)) \leq \sqrt[2n+1]{\phi(Tx) - \phi(x)}. \quad (50)$$

References

- [1] S. Banach, "Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales," *Fundamenta Mathematicae*, vol. 3, pp. 133–181, 1922.
- [2] I. A. Rus, *Generalized Contractions and Applications*, Cluj University Press, Cluj-Napoca, Romania, 2001.
- [3] I. A. Rus, A. Petrușel, and G. Petrușel, *Fixed Point Theory*, Cluj University Press, Cluj-Napoca, Romania, 2008.
- [4] J. Caristi, "Fixed point theorems for mappings satisfying inwardness conditions," *Transactions of the American Mathematical Society*, vol. 215, pp. 241–251, 1976.
- [5] T. Abdeljawad and E. Karapinar, "Quasiconcave metric spaces and generalizations of Caristi Kirk's theorem," *Fixed Point Theory and Applications*, vol. 2009, Article ID 574387, 9 pages, 2009.
- [6] A. Amini-Harandi, "Some generalizations of Caristi's fixed point theorem with applications to the fixed point theory of weakly contractive set-valued maps and the minimization problem," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 72, no. 12, pp. 4661–4665, 2010.
- [7] E. Karapinar, "Generalizations of Caristi Kirk's theorem on partial metric spaces," *Fixed Point Theory and Applications*, vol. 2011, article 4, 2011.
- [8] R. P. Agarwal and M. A. Khamsi, "Extension of Caristi's fixed point theorem to vector valued metric spaces," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 74, no. 1, pp. 141–145, 2011.
- [9] I. Kramosil and J. Michálek, "Fuzzy metrics and statistical metric spaces," *Kybernetika*, vol. 11, no. 5, pp. 336–344, 1975.
- [10] M. Grabiec, "Fixed points in fuzzy metric spaces," *Fuzzy Sets and Systems*, vol. 27, no. 3, pp. 385–389, 1988.
- [11] A. George and P. Veeramani, "On some results in fuzzy metric spaces," *Fuzzy Sets and Systems*, vol. 64, no. 3, pp. 395–399, 1994.

- [12] J. L. Kelley, *General Topology*, D. Van Nostrand, New York, NY, USA, 1955.
- [13] R. Engelking, *General Topology*, PWN—Polish Scientific Publishers, Warsaw, Poland, 1977.

Research Article

Synchronization Analysis of Four Symbolic Complex Dynamical Systems for Future Biology Research

Yagang Zhang,^{1,2} Zheng Zhao,² and Yi Sun³

¹ Department of Mathematics and Physics, North China Electric Power University, Baoding, Hebei 071003, China

² State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, NCEPU, Beijing 102206, China

³ Hebei Electric Power Research Institute, Shijiazhuang 050022, China

Correspondence should be addressed to Yagang Zhang; yagangzhang@gmail.com

Received 22 June 2013; Accepted 2 July 2013

Academic Editor: Dumitru Baleanu

Copyright © 2013 Yagang Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The synchronization analysis of four symbolic complex dynamical systems will be discussed carefully in this paper. Grey system theory is mainly being used to study data sequences that are generated by 4-letter chaotic dynamical system, and the usual prediction accuracy has exceeded 90%. In this place we have found a generating rule that may at least realize chaotic synchronization in short and medium terms. Considering the current study of DNA base sequences A-G-C-T and the symbolic characteristic of four symbolic dynamical systems, which are formally in good corresponding relation. In this paper we have offered an effective research means to approach problems of this kind.

1. Introduction

The problems on uncertainty exist commonly in nature: the ones in myriad sample can be resolved by probability and statistics ways, the ones in recognizing uncertainty can be dealt with by fuzzy mathematics. However, there also exists another kind of uncertainty in less data, little samples, incomplete information, and devoid of experience, which is just suitable to be dealt with by grey system theory [1]. A system with both partially known information and certain unknown information can be defined as a grey system. That is to say, grey system is the system that lacks information, such as architecture, parameters, operation mechanism, and system behavior. Therefore, the key problem is how to make use of the finite available information to approximate the system dynamical behaviors [2].

Grey system theory mainly includes the theoretical system based on grey algebra system, grey equation, and grey matrix; the methodology system based on grey sequence generation; the analysis system based on grey correlation space; the grey-model-centered modeling system; and the technical system based on system analysis, evaluation, modeling, prediction and decision, and control and optimization.

Since there is not any certain physical model in many social and economic issues, it is quite difficult to specify all the factors, not to mention establishing definite mapping relations, though some influence factors are known. For instance, those factors that affect prices, such as psychological expectation and government orientation, are immeasurable. While some data are measurable yet short of detailed information. If the available data alone are taken into consideration, the analysis outcome will surely be inaccurate. Grey system theory can generate new data sequences which indicate the variation trend of the original data and eliminate the fluctuation as well. Grey system theory can also solve some complex system issues with unknown parameters. The applications of grey system theory lie in many scientific fields, such as agriculture, industry, energy, traffic, petroleum, geology, meteorology, hydrology, ecology, environment, medicine, military affairs, economy, and society, which succeeds in solving good many practical problems in production and daily life.

As early as in 1947, Ulam and Neumann studied the densities distribution [3] of surjective parabolic maps' orbital point $\{x_i\}_{i=0}^{\infty}$ and got famous Chebyshev distribution. The study of chaotic symbolic sequences is gradually developing in theory. However, the applied research of stochastic chaotic sequences

has not been fully carried out, for most of studies focus on controlling or avoiding chaos. Chaos, nevertheless, affords inherent stochastic properties that can be calculated, which is an important applied domain. The stochastic symbolic sequences bear the following three features. First, computer can generate them iteratively. Secondly, like false stochastic numbers, they can set up a stochastic sequence simulation (in contradiction, they are based on corresponding symbolic spaces). Thirdly, they can produce numerous symbolic spaces, which is not a characteristic of common stochastic numbers. Therefore, the symbolic dynamics [4–8] developed by this means are supposed to be very useful. We are familiar with two symbolic sequences and have clarified the inherent randomicity in 4 symbolic dynamics [9], the knot theory based on the minimal braid in Lorenz system [10]. In this paper, we are mainly using grey system theory to study the chaotic synchronization of four symbolic complex dynamical systems.

This paper is organized as follows. In Section 2, grey system theory is introduced. In Section 3, the symbolic dynamics of 4-letter surjective map are exhibited. In Section 4, synchronization analysis of 4-letter chaotic dynamical system based on grey system theory is clarified in detail.

2. Grey System Theory

Grey system theory provides an approach to investigate the relationships of input-output process with unclear inner relationships, uncertain mechanisms, and insufficient information. It deals directly with the original data and searches the intrinsic regularity of the data rather than using statistical method. In the grey system theory, grey relational analysis and grey prediction model account for the essential parts [11, 12]. The grey prediction model uses grey differential models to generate data series from the original data series of a dynamical system. And the data series generated by the grey prediction model are converted back to the original data series by a reverse procedure to predict the performance of the dynamical system. Grey prediction can reveal underlying rules within a random time sequence via a special data processing. Now it has been successfully used to model the dynamical systems in different fields, such as agriculture, ecology, economy, statistics, meteorology, and industry [13, 14].

The grey models have many different forms [15] that mainly include

$$x^{(0)}(k) + ax^{(1)}(k) = b,$$

$$x^{(0)}(k) + az^{(1)}(k) = b,$$

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b,$$

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-ak} + \frac{b}{a},$$

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k),$$

$$\begin{aligned} \hat{x}^{(0)}(k) &= \beta - \alpha x^{(1)}(k-1), \\ \left(k = 2, 3, \dots, n, \beta = \frac{b}{1+0.5a}, \alpha = \frac{a}{1+0.5a} \right), \\ \hat{x}^{(0)}(2) &= \beta - \alpha x^{(0)}(1), \\ \hat{x}^{(0)}(k) &= (1-\alpha)x^{(0)}(k-1), \quad (k = 3, 4, \dots, n), \\ \hat{x}^{(0)}(2) &= \beta - \alpha x^{(0)}(1), \\ \hat{x}^{(0)}(k) &= \frac{x^{(1)}(k-1) - 0.5b}{x^{(1)}(k-2) + 0.5b} x^{(0)}(k-1), \quad (k = 3, 4, \dots, n), \\ \hat{x}^{(0)}(k) &= \left(\frac{1-0.5a}{1+0.5a} \right)^{k-2} \left(\frac{b - ax^{(0)}(1)}{1+0.5a} \right), \\ &\quad (k = 2, 3, \dots, n), \\ \hat{x}^{(0)}(k) &= \frac{1}{(1-a)^3} x^{(0)}(3) e^{k \ln(1-a)}, \quad (k = 4, 5, \dots, n), \\ \hat{x}^{(0)}(k) &= (\beta - \alpha x^{(0)}(1)) e^{-a(k-2)}, \quad (k = 3, 4, \dots, n), \\ \hat{x}^{(0)}(k) &= (1 - e^a) \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-a(k-1)}, \\ &\quad (k = 2, 3, \dots, n), \\ \hat{x}^{(0)}(k) &= (-a) \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-a(k-1)}, \quad (k = 2, 3, \dots, n). \end{aligned} \tag{1}$$

One of the most important properties of chaotic system is *sensitivity to initial conditions*. Sensitivity to initial conditions means that an arbitrarily small perturbation of current trajectory may lead to significantly different future behavior. As a result of sensitivity, two different trajectories tend to depart exponentially with time. The behavior of chaotic systems appears to be random.

One of the primary motivations of our research is designed to discuss the corresponding relationships between the exponential separate characteristic in chaotic system and the exponential increasing law of grey prediction. For example, let us choose logistic equation as a model, which has been studied extensively in nonlinear system. One of its equivalent forms can be expressed as

$$x_{n+1} = F(\mu, x_n) = 1 - \mu x_n^2, \tag{2}$$

hereinto, x_n is defined on interval $[-1, 1]$ and $\mu \in (0, 2]$. Its bifurcation diagram is well known: see Figure 1.

Let us adopt unimodal surjective map as $\mu = 2$ to get chaotic iterative sequences, namely,

$$x_{n+1} = f(x_n) = 1 - 2x_n^2. \tag{3}$$

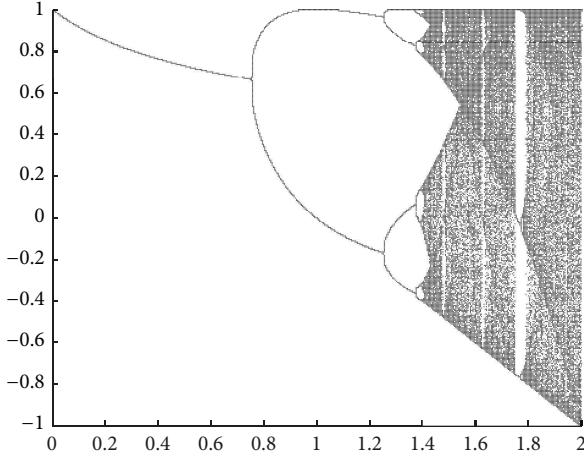
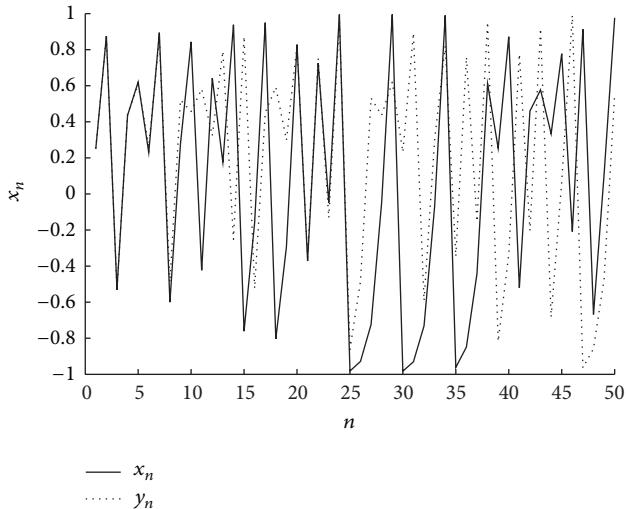


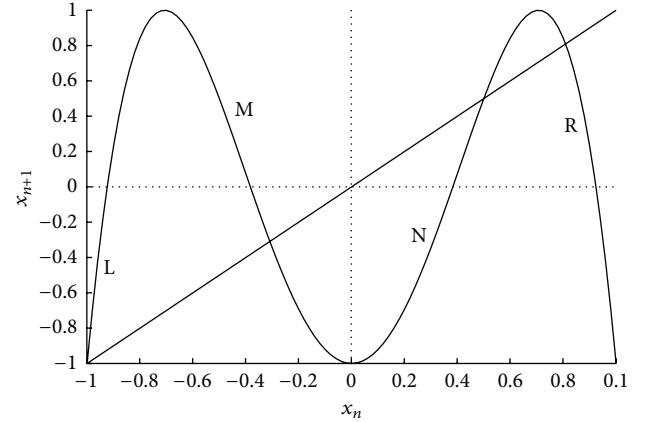
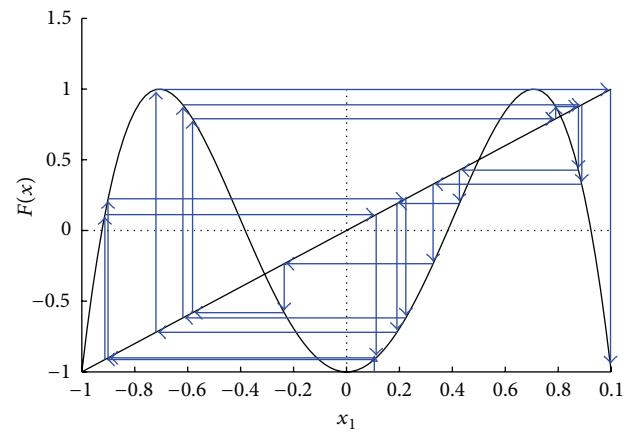
FIGURE 1: The bifurcation diagram of logistic map.

FIGURE 2: The exponential separate diagram of logistic map. (In this diagram, the solid line represents $f(x_n)$ and the dotted line represents $g(y_n)$).

Given initial values $x_1 = 0.25$, $y_1 = 0.25 + 0.001$, after 50 iterations, we can get the sequences of $f(x_n)$ and $g(y_n)$ ($n = 2, 3, 4, \dots$); see Figure 2 and Table 1.

According to Figure 2 and Table 1, it is not difficult to find some differences. Although the iterative original value only has a small deviation 0.001, during the iterative process, a remarkable change begins at the 7th step. Until at the 10th step, it has been very difficult to find that these two sequences came from small deviation original value. Classical researches have proved that these two trajectories generated from two adjacent original values will depart exponentially with time evolution. It is an important measure of the exponential separate characteristic in chaotic system that calculates Lyapunov exponent:

$$\lambda = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln \left| \frac{dF(x)}{dx} \right|_{x=x_i}. \quad (4)$$

FIGURE 3: The shape of 4-letter surjective map $x_{n+1} = -8x_n^4 + 8x_n^2 - 1$; four monotonic embankments are marked by L, M, N, and R.FIGURE 4: 16 iterations of 4-letter surjective map ($x_1 = 0.105$).

3. Symbolic Dynamics of 4-Letter Surjective Map

The generic iterative form of 4-letter surjective map is

$$x_{n+1} = F(A, x_n) = -8x_n^4 + 8x_n^2 - 1. \quad (5)$$

The shape of 4-letter surjective map is shown in Figure 3.

4. Synchronization Analysis of 4-Letter Chaotic System Based on Grey System Theory

4.1. Iterative Process of 4-Letter Chaotic System. The general dynamical iterative form of 4-letter surjective map is expressed in (5). Given an initial value $x_1 = 0.105$, the number of iterations is $n = 16$ and one can obtain Tables 2 and 3. Figure 4 is the concrete iterative process.

Let us transform these iterative values. First, each $F(x_n)$ is made as absolute value transformation. Then, let us carry out cumulative sum $Z(x_k) = \sum_{i=1}^k |F(x_i)|$ and get Tables 4 and 5.

TABLE 1: The exponential separate sequences of logistic map ($n = 2, 3, 4, \dots$).

n	2	3	4	5	6	7	8	9
$f(x_n)$	0.8750	-0.5313	0.4355	0.6206	0.2297	0.8945	-0.6001	0.2797
$g(y_n)$	0.8740	-0.5277	0.4430	0.6076	0.2618	0.8630	-0.4894	0.5209
n	10	11	12	13	14	15	16	17
$f(x_n)$	0.8435	-0.4231	0.6421	0.1755	0.9384	-0.7611	-0.1585	0.9498
$g(y_n)$	0.4573	0.5818	0.3231	0.7912	-0.2521	0.8728	-0.5237	0.4515
n	18	19	20	21	22	23	24	25
$f(x_n)$	-0.8041	-0.2931	0.8281	-0.3717	0.7237	-0.0476	0.9955	-0.9819
$g(y_n)$	0.5924	0.2982	0.8221	-0.3517	0.7526	-0.1328	0.9647	-0.8615
n	26	27	28	29	30	31	32	33
$f(x_n)$	-0.9283	-0.7235	-0.0469	0.9956	-0.9824	-0.9304	-0.7313	—
$g(y_n)$	-0.4843	0.5310	0.4361	0.6196	0.2322	0.8921	-0.5918	—

TABLE 2: The iterative values of 4-letter chaotic system ($n = 1, 2, \dots, 8$).

n	1	2	3	4	5	6	7	8
x_n	0.1050	-0.9128	0.1121	-0.9008	0.2242	-0.6182	0.8888	0.3272
$F(x_n)$	-0.9128	0.1121	-0.9008	0.2242	-0.6182	0.8888	0.3272	-0.2354

TABLE 3: The iterative values of 4-letter chaotic system ($n = 9, 10, \dots, 16$).

n	9	10	11	12	13	14	15	16
x_n	-0.2354	-0.5814	0.7902	0.8762	0.4265	0.1906	-0.7199	0.9973
$F(x_n)$	-0.5814	0.7902	0.8762	0.4265	0.1906	-0.7199	0.9973	-0.9579

TABLE 4: The cumulative sum results of $Z(x_k)$ ($k = 1, 2, \dots, 8$).

k	1	2	3	4	5	6	7	8
$Z(x_k)$	0.9128	1.0248	1.9256	2.1498	2.7680	3.6568	3.9840	4.2193

TABLE 5: The cumulative sum results of $Z(x_k)$ ($k = 9, 10, \dots, 16$).

k	9	10	11	12	13	14	15	16
$Z(x_k)$	4.8007	5.5909	6.4671	6.8936	7.0843	7.8041	8.8015	9.7593

4.2. Random Generating Sequence's Grey Prediction of 4-Letter Chaotic System. In the grey system theory, the expression of accumulated generating operator (AGO) is $x^{(1)}(k) = \sum_{m=1}^k x^{(0)}(m)$, and so one gets Table 6.

The formula of grey estimate value is

$$\text{Estim.} = \left[x^{(1)}(0) - \frac{u}{a} \right] \times e^{-at} \times (1 - e^a), \quad (6)$$

thereinto, $x^{(1)}(0) = 0.9128$, $a = -0.11155$, $u = 1.95557$, and $t = 1, 2, 3, \dots$, namely,

$$\text{Estim.} = 18.4437 \times e^{0.11155t} \times (1 - e^{-0.11155}). \quad (7)$$

So, the prediction values and corresponding error of 16 iterations can be calculated (see Table 7).

Figure 5 shows the actual values and the prediction values of 16 iterations.

Here are grey prediction results of the following 3 steps:

$$\begin{aligned} k = 17: \text{Estim. value} &= 11.60, \\ k = 18: \text{Estim. value} &= 12.97, \\ k = 19: \text{Estim. value} &= 14.50. \end{aligned} \quad (8)$$

It will be intuitionistic and natural to return these grey prediction results to actual symbolic space. Combining the exact position of peak and trough, the value of positive or negative can be confirmed.

Now let us present a research example.

Example 1. According to a group of DNA based sequences "CCGGGGTCGCGCAGCA", which double helix structure has been presented in Figure 6. Of course, the DNA based sequences have perfect matchings. In our researches, the dynamical matching rules of four symbolic dynamical

TABLE 6: Main variable values of grey prediction process.

k	$Z(x_k)$	AGO	AGO mean	Mean square	Mean AGO	Mean square AGO	Mean AGO square
1	0.9128	0.9128	*	*	*	*	*
2	1.0248	1.9376	1.4252	2.0312	1.4252	2.0312	2.0312
3	1.9256	3.8632	2.9004	8.4123	4.3256	10.4435	18.7108
4	2.1498	6.0130	4.9381	24.3848	9.2637	34.8283	85.8161
5	2.7680	8.7810	7.3970	54.7156	16.6607	89.5440	277.5789
6	3.6568	12.4378	10.6094	112.5594	27.2701	202.1033	743.6584
7	3.9840	16.4218	14.4298	208.2191	41.6999	410.3225	1738.882
8	4.2193	20.6411	18.5315	343.4146	60.2314	753.7371	3627.816
9	4.8007	25.4418	23.0415	530.9084	83.2728	1284.6455	6934.359
10	5.5909	31.0327	28.2373	797.3423	111.5101	2081.9878	12434.49
11	6.4671	37.4998	34.2663	1174.176	145.7763	3256.1637	21250.73
12	6.8936	44.3934	40.9466	1676.624	186.7229	4932.7877	34865.44
13	7.0843	51.4777	47.9356	2297.817	234.6585	7230.6047	55064.59
14	7.8041	59.2818	55.3798	3066.917	290.0382	10297.5214	84122.16
15	8.8015	68.0833	63.6826	4055.467	353.7208	14352.9886	125118.4
16	9.7593	77.8426	72.9630	5323.592	426.6837	19676.5806	182059.0

TABLE 7: Prediction values and corresponding error of 16 iterations.

k	$Z(x_k)$ Actual value	$Z(x_k)$ Estimate value	Error	Relative error	Precision
1	0.9128	0.9128	0	0	100%
2	1.0248	2.18	-1.1552	-112.724%	-12.7244%
3	1.9256	2.43	-0.5044	-26.1944%	73.80557%
4	2.1498	2.72	-0.5702	-26.5234%	73.4766%
5	2.7680	3.04	-0.272	-9.82659%	90.17341%
6	3.6568	3.40	0.2568	7.022533%	92.97747%
7	3.9840	3.80	0.184	4.618474%	95.38153%
8	4.2193	4.25	-0.0307	-0.72761%	99.27239%
9	4.8007	4.75	0.0507	1.056096%	98.9439%
10	5.5909	5.31	0.2809	5.024236%	94.97576%
11	6.4671	5.94	0.5271	8.150485%	91.84952%
12	6.8936	6.64	0.2536	3.678775%	96.32123%
13	7.0843	7.42	-0.3357	-4.73865%	95.26135%
14	7.8041	8.30	-0.4959	-6.35435%	93.64565%
15	8.8015	9.28	-0.4785	-5.43657%	94.56343%
16	9.7593	10.38	-0.6207	-6.36009%	93.63991%

systems and DNA sequences are “L \leftrightarrow A, M \leftrightarrow G, N \leftrightarrow C, and R \leftrightarrow T”. Based on the former synchronization analysis process and grey system theory, the prediction model can be expressed as

$$\text{Estim.} = 13.4428 \times e^{0.10910t} \times (1 - e^{-0.10910t}). \quad (9)$$

The final four symbolic prediction results are “NNMM-MNMMNMNMNL”, and the corresponding DNA based sequences are “CCGGGGCCGCGCGCGCA”. Figure 7 represents the DNA prediction structure. To sum up the previous synchronization analysis results, the error rate is

$$\hat{p}^* = \frac{1}{16} = 0.0625, \quad (10)$$

namely, the accuracy rate has reached 93.75% ($1 - \hat{p}^*$).

It should be stressed that the biological system is highly complex; we cannot look to one method or two methods to solve all problems. But at least, we have provided an interesting thought.

5. Conclusion and Discussion

The synchronization analysis of four symbolic complex dynamical systems has been clarified carefully in this paper. We are mainly using grey system theory to study data sequences that are generated by 4-letter chaotic dynamical system, and the usual prediction accuracy has exceeded 90%. In this place we have found a generating rule that may at least realize chaotic synchronization in short and medium terms.

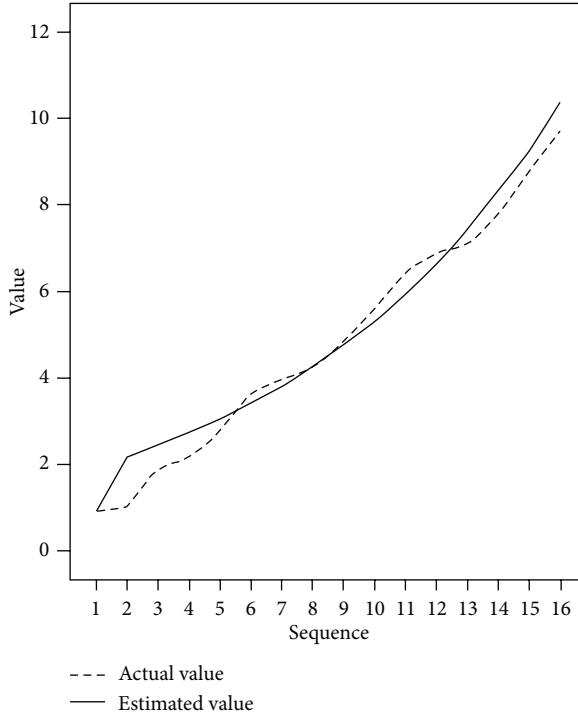


FIGURE 5: 16 iterations' actual values and prediction values of 4-letter chaotic system.

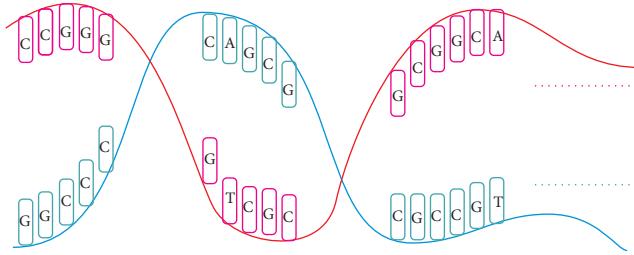


FIGURE 6: DNA double helix structure.

And in this way we can analyze and predict different kinds of complex systems.

The world is essentially nonlinear and highly complex. A small stochastic force is usually able to have an unexpected impact on deterministic equation, and in certain conditions, it could determine the evolution of systems and even change the fate of macroscopic system. Nonlinear dynamics theory, the combination of Newtonian mechanics and statistical mechanics method, linear stochastic equation and nonlinear deterministic equation, and periodic solution and chaotic solution have achieved the high inherent unity of determinacy and randomicity, determinism, and indeterminism. The change of weather, the growth of species, the diffusion of molecular, the fluctuation of electrocardiogram and electroencephalogram signals, the periodic reform of society, the growth, the sudden plunge, and even the collapse of stock markets all have inherent nonlinear dynamics rules [16]. The complexity of social economic system mainly lies in the dynamic evolution of multivariable system,

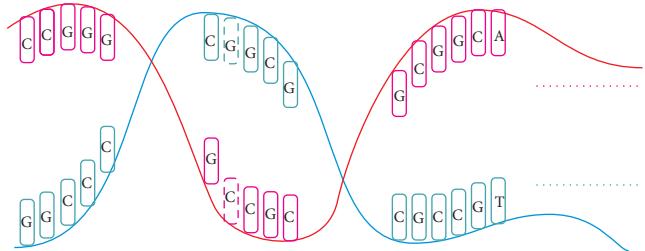


FIGURE 7: DNA prediction structure.

the incompleteness and uncertainty of social information. The multilevels of social economic system result in each level having its own structure composition. Each social unit, whether it is a whole unit or the smallest fundamental sector, is realizing its functions according to its structural property. The various randomicity and uncertainties in social system make it impossible to fully master the entire system information. Therefore, it has been the main task of nonlinear dynamics theory and grey system theory to analyze the complexity of social system in the conditions of incomplete information and uncertainty and to study the intrinsic laws of social development.

Considering the current study of DNA [17–25] based sequences A-G-C-T and the symbolic characteristic of four symbolic dynamical systems, which are formally in good corresponding relation. In this paper we have offered an effective research means. A part of our current work is studying certain properties of different kinds of data sequences, such as DNA based sequences, 20 amino acids symbolic sequences in proteid structures, and time series that can be symbolic in finance market. The grey system theory provides a set of effective methods to approach problems of this kind, which will open up a vast vista.

Acknowledgments

This research was supported partly by the NSFC (51277193), the Specialized Research Fund for the Doctoral Program of Higher Education (20110036110003), the Fundamental Research Funds for the Central Universities (11MG37, 12MS120), and the Natural Science Foundation of Hebei Province.

References

- [1] J. L. Deng, *Basic Theory of Grey System*, Huazhong University of Science and Technology Press, Wuhan, China, 2002.
- [2] S. F. Liu and Y. Lin, *Grey Information: Theory and Practical Applications*, Springer, London, UK, 2006.
- [3] S. M. Ulam and J. V. Neumann, "On combination of stochastic and deterministic processes," *Bulletin of the American Mathematical Society*, vol. 53, no. 11, p. 1120, 1947.
- [4] B.-L. Hao and W.-M. Zheng, *Applied Symbolic Dynamics and Chaos*, vol. 7 of *Directions in Chaos*, World Scientific, River Edge, NJ, USA, 1998.
- [5] K.-F. Cao, X.-S. Zhang, Z. Zhou, and S.-L. Peng, "Devil's carpet of topological entropy and complexity of global dynamical

- behavior," *Chaos, Solitons & Fractals*, vol. 16, no. 5, pp. 709–726, 2003.
- [6] S.-L. Peng, K.-F. Cao, and Z.-X. Chen, "Devil's staircase of topological entropy and global metric regularity," *Physics Letters A*, vol. 193, no. 5-6, pp. 437–443, 1994.
- [7] Q. Liu, K.-F. Cao, and S.-L. Peng, "A generalized Kolmogorov-Sinai-like entropy under Markov shifts in symbolic dynamics," *Physica A*, vol. 388, no. 20, pp. 4333–4344, 2009.
- [8] K. Wang, W. Pei, X. Hou, Y. Shen, and Z. He, "Symbolic dynamics approach to parameter estimation without initial value," *Physics Letters A*, vol. 374, no. 1, pp. 44–49, 2009.
- [9] Y. Zhang, C. Wang, and Z. Zhou, "Inherent randomicity in 4-symbolic dynamics," *Chaos, Solitons & Fractals*, vol. 28, no. 1, pp. 236–243, 2006.
- [10] Y. Zhang and Z. Wang, "Knot theory based on the minimal braid in Lorenz system," *International Journal of Theoretical Physics*, vol. 47, no. 4, pp. 873–880, 2008.
- [11] H. V. Trivedi and J. K. Singh, "Application of grey system theory in the development of a runoff prediction model," *Biosystems Engineering*, vol. 92, no. 4, pp. 521–526, 2005.
- [12] S. He, Y. Li, and R. Z. Wang, "A new approach to performance analysis of ejector refrigeration system using grey system theory," *Applied Thermal Engineering*, vol. 29, no. 8-9, pp. 1592–1597, 2009.
- [13] L.-C. Hsu and C.-H. Wang, "Forecasting integrated circuit output using multivariate grey model and grey relational analysis," *Expert Systems with Applications*, vol. 36, no. 2, pp. 1403–1409, 2009.
- [14] L. Wei, M. Fei, and H. Hu, "Modeling and stability analysis of grey-fuzzy predictive control," *Neurocomputing*, vol. 72, no. 1–3, pp. 197–202, 2008.
- [15] S. F. Liu and N. M. Xie, *Grey System Theory and Its Applications*, Science Press, Beijing, China, 2008.
- [16] Y. Zhang, Y. Xu, and Z. Wang, "Dynamical randomicity and predictive analysis in cubic chaotic system," *Nonlinear Dynamics*, vol. 61, no. 1-2, pp. 241–249, 2010.
- [17] M. Daniel and V. Vasumathi, "Nonlinear molecular excitations in a completely inhomogeneous DNA chain," *Physics Letters A*, vol. 372, no. 31, pp. 5144–5151, 2008.
- [18] Z.-B. Wu, "Global transposable characteristics in the complete DNA sequence of the yeast," *Physica A*, vol. 389, no. 24, pp. 5698–5705, 2010.
- [19] J. Lu, W. L. Dentler, and E. A. Lundquist, "FLI-1 Flightless-1 and LET-60 Ras control germ line morphogenesis in *C. elegans*," *BMC Developmental Biology*, vol. 8, article 54, pp. 1–15, 2008.
- [20] S. R. C. Lopes and M. A. Nunes, "Long memory analysis in DNA sequences," *Physica A*, vol. 361, no. 2, pp. 569–588, 2006.
- [21] M. L. E. Gracey, J. M. Maniar, Z. Y. Chen, J. Lu, A. Z. Fire, and M. A. Kay, "Minicircle DNA vectors achieve sustained expression reflected by active chromatin and transcriptional level," *Molecular Therapy*, vol. 21, no. 1, pp. 131–138, 2013.
- [22] J. Lu, F. Zhang, and M. Kay, "A mini-intronic plasmid (MIP): a novel robust transgene expression vector in vivo and in vitro," *Molecular Therapy*, vol. 21, no. 5, pp. 954–963, 2013.
- [23] V. R. Chechetkin and V. V. Lobzin, "Spectral entropy criteria for structural segmentation in genomic DNA sequences," *Physics Letters A*, vol. 328, no. 1, pp. 79–86, 2004.
- [24] Y. Yang, J. Lu, J. Rovnak, S. L. Quackenbush, and E. A. Lundquist, "SWAN-1, a *Caenorhabditis elegans* WD repeat protein of the AN11 family, is a negative regulator of Rac GTPase function," *Genetics*, vol. 174, no. 4, pp. 1917–1932, 2006.
- [25] J. Lu, F. Zhang, S. Xu, A. Z. Fire, and M. A. Kay, "The extragenic spacer length between the 5' and 3' ends of the transgene expression cassette affects transgene silencing from plasmid-based vectors," *Molecular Therapy*, vol. 20, no. 11, pp. 2111–2119, 2012.

Research Article

Swarm Stability Analysis of High-Order Linear Time-Invariant Singular Multiagent Systems

Mingyu Fu,¹ Shimin Wang,¹ and Yirui Cong²

¹ Department of Control Theory, College of Automation, Harbin Engineering University, Harbin 150001, China

² Department of Control Theory, College of Automation, Northeastern University, Shenyang 110000, China

Correspondence should be addressed to Shimin Wang; wangshimin@hrbeu.edu.cn

Received 5 May 2013; Revised 15 June 2013; Accepted 24 June 2013

Academic Editor: Dumitru Baleanu

Copyright © 2013 Mingyu Fu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The swarm stability problem of high-order linear time-invariant (LTI) singular multiagent systems with directed graph topology is investigated extensively. Consensus of multiagent systems can be regarded as a specific case of swarm stability problem. Necessary and sufficient conditions for both swarm stability and consensus are presented. These conditions depend on the graph topology and generalized inverse theory, the dynamics of agents, and interaction among the neighbours. Several examples to illustrate the effectiveness of theoretical results are given.

1. Introduction

Recently the study of multiagent systems has attracted considerable attention from biologists (see [1, 2]), physical scientists (see [3]), and engineers (see [4–8]). Due to its broad application in many areas such as satellites, spacecrafts (see [4]), aircrafts (see [5]), mobile robots, UAV, AUV (see [6]), flocking of social insects (see [7]), and design of sensor networks (see [8]). Stability is usually a basic requirement for a control system. Stability of multiagent systems is special because of the configuration difference between these systems and isolated dynamic systems. Models of these systems originated from bioscience. Studies on them have aroused great interest in the field of automatic control theory since about a decade ago. The research on stability problems of multiagent systems has been studied extensively for years. Researchers have endeavoured to give criteria for checking whether a specific multiagent system is stable. For example, Jin et al. [9] paid attention to the stability of a simple swarm system rather early. The background of their research came from engineering systems such as multiagent supporting system [10]. Based on the work of Jin, Liu et al. [11, 12] tried to extend the research to more general and critical situations, and they studied the so-called asynchronous swarm systems. However, the technical setting of their model is complicated, and the results are restrictive. Compared with the research of

Liu et al., Gazi and Passino [13, 14] and Gazi [15] studied the stability problem of some first-order models abstracted from biological systems. Chu et al. [16] extended the discussions in [13] to certain anisotropic model. Recently, Li [17] extended the results in [13] to the situation with general graph topology instead of a complete graph. Cai et al. [18] studied the swarm stability of high-order linear time-invariant multiagent systems in 2011. Besides, there were also other studies related to the stability of swarm systems, for example, the study about string stability [19, 20]. Zhou et al. [21] studied the stability of switched linear singular systems.

Consensus of multiagent systems can be regarded as a specific case of swarm stability problem; our results shows that consensus achievement is a specific case of swarm stability, which will be called asymptotic swarm stability. Consensus problems for multiagent systems have been studied by lots of researchers. Vicsek et al. [3] proposed a discrete-time first-order model of n agents moving in the horizontal plane with different headings. The systems can achieve consensus on heading via updating their heading according to their neighbor's heading. Jadbabaie et al. [22] used graph theory to give an explanation. Olfati-Saber and Murray [23] showed that the first-order consensus problems can be solved if the topology was strongly connected. Ren and Beard [24] further proved that the first-order consensus can be achieved if interaction graphs have a directed spanning tree. Xiao et al. [25]

studied the general first-order consensus problems multiagent discrete-time systems with fixed topology. Several dynamic consensus algorithms for second-order multiagent systems and sufficient conditions for state consensus of systems have been proposed in [26, 27]. Tian and Zhang [28] discussed the high-order consensus of heterogeneous multiagent systems. Zheng and Wang [29] investigated finite-time consensus problems of heterogeneous multiagent systems. Xiao and Wang [30] have investigated the consensus problem for high-order general nonsingular multiagent systems based on the structure of certain high-dimensional matrices. Cheng et al. [31] solved a modified consensus problems of high-order normal multiagent systems with fixed fixed communication topologies. Xi et al. [32] studied the consensus problems for general high-order LTI multiagent systems which have time-varying consensus function.

Recently, more attention has been paid to the consensus problem for the singular multiagent systems. It is well known that singular systems can better describe physical systems than normal ones (see [33]). When there are algebraic constraints among coordinated variables or when coordinated variables include both fast variation components and slow variation components, each agent can only be modeled as a singular system rather than a normal one. One typical example is three-link manipulator show in [34], which can be used to clean the facade of a very large building. Another example is multiagent supporting systems in earthquake damage-preventing buildings, water-floating plants, large-diameter parabolic antennae, or telescopes. As shown in [35], a MASS can only be described by a singular multiagent system if it consists of a lot of independent blocks and each block is supported by several pillars. Xi et al. (see [35–37]) addressed consensus problems of high-order singular multiagent systems via linear matrix inequality (LMI) constraints and matrix equality constraints. Yang et al. (see [34]) studied consensus problems for singular multiagent systems with output feedback consensus protocols. Zhou et al. [21] analyzed the stability of switched linear singular systems. These study transformed singular system into two restricted equivalent subsystems: slow subsystem and quick subsystem, which is complicated. The current paper uses the Drazin inverse theory to analyze the singular multiagent systems which can easily extend the normal system results to singular system. The Drazin inverse theorem which is a kind of generalized inverse was firstly presented by Drazin (see [38, 39]). Campbell et al. (see [40, 41]) used the Drazin inverse to give the solutions of singular system. More knowledge about the Drazin inverse readers can get from [41–44].

Compared with the existing works related to consensus, the current paper is characterized with the following novel features. First the current paper focuses on swarm stability problem of high-order singular multiagent systems; our results show that consensus achievement is a specific case of swarm stability, which will be called asymptotic swarm stability. Second the main contribution of this paper is the presentation of necessary and sufficient conditions for the swarm stability of LTI singular systems with a general high-order model. The model in [34, 35] can be seen as special case of our results. By far, most of the linear high-order swarm

system models studied by other scholars on consensus problems are more restrictive than the one in the current paper. Third some necessary and sufficient conditions based on the firstly utilization of the Drazin inverse theory to analyze the singular multiagent system are given, and a tool is presented to analyze the singular multiagent systems as well as the nonsingular systems. When index of the Drazin inverse is zero, these results (see [18, 32, 35, 45]) in normal system are actually the current paper's corollary. The initial condition with nonadmissible bounded value is illustrated.

This paper is organised as follows. In Section 2, some basic lemmas and problem descriptions are introduced. In Section 3, the swarm stability problem and the LTI singular swarm system model are described. In Section 4, a necessary and sufficient condition for swarm stability is expounded. Section 4 provides a necessary and sufficient condition for asymptotic swarm stability. Some numerical simulations are discussed in Section 5. Finally, Section 6 concludes this paper.

2. Preliminaries and Problem Descriptions

2.1. Some Lemmas and Definitions

Definition 1 (see [33]). For any given two matrices $E, A \in \mathbb{R}^{n \times n}$, (E, A) is called regular if there exists a constant scalar $\alpha \in \mathbb{C}$ such that $\det(\alpha E + A) \neq 0$.

Lemma 2 (see [33]). For any given two matrices $E, A \in \mathbb{R}^{n \times n}$, (E, A) is regular if and only if there exist two nonsingular matrices $Q, P \in \mathbb{R}^{n \times n}$ such that

$$QEP = \text{diag}\{I_{n_1}, N\}, \quad QAP = \text{diag}\{A_1, I_{n_2}\}, \quad (1)$$

where $n_1 + n_2 = n$, $A_1 \in \mathbb{R}^{n_1 \times n_1}$, $N \in \mathbb{R}^{n_2 \times n_2}$ is nilpotent.

Lemma 3 (see [33]). For the regular singular system $E\dot{x}(t) = Ax(t) + Bu(t)$ with $E, A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times d}$,

- (1) the system (E, A, B) does not have impulse-modes called system impulse-free if and only if $\text{rank}[\begin{smallmatrix} E & 0 \\ A & E \end{smallmatrix}] = n + \text{rank}(E)$ or $\deg \det(sE - A) = \text{rank}(E)$,
- (2) the system (E, A, B) is impulse controllable if and only if $\text{rank}[\begin{smallmatrix} E & 0 & 0 \\ A & E & B \end{smallmatrix}] = n + \text{rank}(E)$.

Definition 4 (see [40]). Let $A \in \mathbb{C}^{n \times n}$, where the index of A is the smallest nonnegative integer k such that $\text{rank}(A^{k+1}) = \text{rank}(A^k)$ is denoted by $\text{Ind}(A)$. For a square matrix A , the unique matrix X satisfying the following relations

$$XAX = X, \quad AX = XA, \quad A^{k+1}X = A^k \quad (2)$$

is called the Drazin inverse of A and denoted by $X = A^D$. Particularly, when $\text{Ind}(A) = 0$, the matrix X is A^{-1} .

Some Drazin inverse lemmas will be introduced.

Lemma 5 (see [46]). If $\text{Index}(A) = k$, $\text{rank}(A^k) = r$ and P , Q , and C are invertible, N is nilpotent of index k , the Drazin inverse of B is B^D , and the Drazin inverse of A is A^D , then

$$(1) (QAQ^{-1})^D = QA^DQ^{-1},$$

$$(2) (A \otimes B)^D = A^D \otimes B^D,$$

(3) if A have a Jordan canonical form $A = P \begin{bmatrix} C & 0 \\ 0 & N \end{bmatrix} P^{-1}$, then the Drazin inverse can be computed as

$$A^D = P \begin{bmatrix} C^{-1} & 0 \\ 0 & 0 \end{bmatrix} P^{-1}. \quad (3)$$

Lemma 6. For any given $E, A \in \mathbb{R}^{n \times n}$ and if (E, A) is regular, let $\tilde{E} = (\alpha E + A)^{-1}E$ and $\tilde{A} = (\alpha E + A)^{-1}A$. There exist two nonsingular matrices $Q, P \in \mathbb{R}^{n \times n}$, such that

$$\begin{aligned} (\tilde{E})^D &= P \begin{bmatrix} I_{n_1} & \\ & O_{n_2} \end{bmatrix} Q(\alpha E + A), \\ (\tilde{E})^D \tilde{A} &= P \begin{bmatrix} A_1 & \\ & O_{n_2} \end{bmatrix} P^{-1} = \left(P \begin{bmatrix} I_{n_1} & \\ & O_{n_2} \end{bmatrix} Q \right) A, \quad (4) \\ (\tilde{E})^D \tilde{E} &= P \begin{bmatrix} I_{n_1} & \\ & O_{n_2} \end{bmatrix} P^{-1} \end{aligned}$$

with $A_1 \in \mathbb{R}^{n_1 \times n_1}$, $n_1 + n_2 = n$, $\alpha \in \mathbb{C}$ is a constant scalar, and O_{n_2} is a zero matrix.

Proof. For (E, A) is regular, by Lemma 2, there exist two nonsingular matrices $Q, P \in \mathbb{R}^{n \times n}$; we can get $E = Q^{-1} \text{diag}\{I_{n_1}, N\}P^{-1}$, $A = Q^{-1} \text{diag}\{A_1, I_{n_2}\}P^{-1}$, where $A_1 \in \mathbb{R}^{n_1 \times n_1}$, $N \in \mathbb{R}^{n_2 \times n_2}$, and $n_1 + n_2 = n$, N is nilpotent matrix. We get that

$$\begin{aligned} \tilde{E} &= (\alpha E + A)^{-1}E \\ &= P \begin{bmatrix} (\alpha I_{n_1} + A_1)^{-1} & \\ & (\alpha N + I_{n_2})^{-1} \end{bmatrix} Q Q^{-1} \begin{bmatrix} I_{n_1} & \\ & N \end{bmatrix} P^{-1} \\ &= P \begin{bmatrix} (\alpha I_{n_1} + A_1)^{-1} & \\ & (\alpha N + I_{n_2})^{-1} N \end{bmatrix} P^{-1} \quad (5) \end{aligned}$$

for $(\alpha N + I_{n_2})^{-1}N$ is nilpotent, and with Lemma 5 we can get

$$\begin{aligned} (\tilde{E})^D &= P \begin{bmatrix} \alpha I_{n_1} + A_1 & \\ & O_{n_2} \end{bmatrix} P^{-1} \\ &= P \begin{bmatrix} I_{n_1} & \\ & O_{n_2} \end{bmatrix} Q Q^{-1} \begin{bmatrix} \alpha I_{n_1} + A_1 & \\ & \alpha N + I_{n_2} \end{bmatrix} P^{-1} \quad (6) \\ &= P \begin{bmatrix} I_{n_1} & \\ & O_{n_2} \end{bmatrix} Q(\alpha E + A), \end{aligned}$$

and the other results are easily to get. \square

Lemma 7. If E is a nonsingular matrix in Lemma 6, suppose that $\tilde{E} = (\alpha E + A)^{-1}E$, $\tilde{A} = (\alpha E + A)^{-1}A$. Then

$$\begin{aligned} (\tilde{E})^D &= E^{-1}(\alpha E + A), \\ (\tilde{E})^D \tilde{A} &= E^{-1}A, \\ (\tilde{E})^D \tilde{E} &= I. \end{aligned} \quad (7)$$

Lemma 8 (see [39]). For $A, B \in \mathbb{R}^{n \times n}$, regular system $Ax(t) + Bx(t) = f(t)$ has unique solutions for admissible conditions. Let α be a number such that $\hat{A} = (\alpha A + B)^{-1}A$, $\hat{B} = (\alpha A + B)^{-1}B$, $\hat{f}(t) = (\alpha A + B)^{-1}f(t)$, and $k = \text{Ind}(\hat{A})$. Then

$$\begin{aligned} x(t) &= e^{-\hat{A}^D \hat{B} t} \hat{A}^D \hat{A} x(0) \\ &\quad + \hat{A}^D e^{-\hat{A}^D \hat{B} t} \int_0^t e^{\hat{A}^D \hat{B} s} \hat{f}(s) ds \\ &\quad + (I - \hat{A}^D \hat{A}) \sum_{n=0}^{k-1} (-1)^n (\hat{A} \hat{B}^D)^n \hat{A}^D \hat{f}^{(n)}(t). \end{aligned} \quad (8)$$

A directed weighted graph G consists of a note set $V(G) = \{v_1, v_2, \dots, v_n\}$ and an edge set $E(G) \subseteq \{(v_i, v_j) : v_i, v_j \in V(G)\}$, an adjacency matrix of graph G is defined by $W = [w_{ij}] \in \mathbb{R}^{n \times n}$, and w_{ij} is weight of (v_i, v_j) , $(v_i, v_j) \in E(G)$ if only if $w_{ij} > 0$. Moreover, it is supposed that $w_{ij} = 0$ for all $i \in \{1, 2, \dots, n\}$. The set of neighbors of v_i is denoted by $N_i = \{v_j \in V(G) : (v_i, v_j) \in E(G)\}$. The Laplacian matrix of graph G is defied as $L = D - W$, where D is the in-degree matrix of G . A directed weighted graph having a spanning tree means that there exists at least one node having a directed path to all other nodes.

Lemma 9 (see [24]). Let $L \in \mathbb{R}^{n \times n}$ be the Laplacian matrix of a directed graph G and $1_n = [1, 1, \dots, 1]^T \in \mathbb{R}^n$. Then

- (1) L at least has a zero eigenvalue and 1_n is the associated eigenvector, that is, $L1_n = 0$,
- (2) if graph G has a spanning tree, then 0 is a simple eigenvalue of L , and all the other $n-1$ eigenvalues have positive real parts.

2.2. Description of Singular Multiagent Continuous-Time Systems

Assumption 10. The systems which the current paper studies are regular and impulse-free.

The regularity condition could guarantee that the system has a unique solution. The regularity condition is deemed as a kind of singular system (see [33]). If the continuous-time system has strong impulse modes, the impulse may stop the system from working or even destroy it. Readers can get the example from [47].

This section introduces our multiagent model. Suppose that the multiagent system consists of n autonomous agents, labeled i through n , and all these agents share a common state

space \mathbb{R}^m . Let $x_i(t)$ denote the state of agent i , $i \in \{1, 2, \dots, n\}$, and let $x(t) = [x_1^T(t), x_2^T(t), \dots, x_n^T(t)]^T$. We say that i th and j th agents on a network agree if and only if $x_i(t) = x_j(t)$ and that the agents on a network reached a consensus if and only if $x_i(t) = x_j(t)$ for all $i, j \in \{1, 2, \dots, n\}, i \neq j$ [23]. Whenever the agents of a network are all in agreement, the common value of all agents is called group decision or consensus state.

Now we take a typical example, and we also consider the dynamics of the n agents where the communication topology is represented by G and each vertex corresponds to an agent. Suppose that these agents are with the following dynamics:

$$E\dot{x}_i(t) = Ax_i(t) + Bu_i(t) \quad (9)$$

and system with state feedback:

$$E\dot{x}_i(t) = Ax_i(t) + BK_1x_i(t) + Bu_i(t), \quad (10)$$

where $E \in \mathbb{R}^{m \times m}$, $A \in \mathbb{R}^{m \times m}$, $B \in \mathbb{R}^{m \times d}$, and $K_1 \in \mathbb{R}^{d \times m}$. As we can see when system with state feedback is actually special case of system (9) for $E\dot{x}_i(t) = (A + BK_1)x_i(t) + Bu_i(t)$ with $\text{rank}(E) = c \leq m$ and $u_i(t)$ is the consensus protocol which is designed according to the information available to agent i , the following consensus protocol is applied:

$$u_i(t) = K_2 \sum_{j \in n_i} w_{ij} (x_j(t) - x_i(t)), \quad (11)$$

where $i \in \{1, 2, \dots, n\}$, $K_1, K_2 \in \mathbb{R}^{d \times m}$. Let $x(t) = [x_1^T(t), x_2^T(t), \dots, x_n^T(t)]^T$, and let L be the Laplacian matrix of graph G . Then the dynamics of a singular swarm system with n agents can be described by

$$(I_n \otimes E)\dot{x}(t) = Hx(t) = (I_n \otimes A - L \otimes BK_2)x(t). \quad (12)$$

For the Laplacian matrix $L \in \mathbb{R}^{n \times n}$ of the interaction topology G , there exists a nonsingular matrix $U \in \mathbb{C}^{n \times n}$ such that $U^{-1}LU = J_L$ where J_L is a Jordan canonical form of L . Let $\bar{x}(t) = (\bar{x}_1^H(t), \bar{x}_2^H(t), \dots, \bar{x}_n^H(t))^H = (U^{-1} \otimes I_n)x(t)$, where I_n is identity matrix. Let $H_{\lambda_i} = A - \lambda_i BK_2, i \in \{1, 2, \dots, n\}$. Then by Lemma 20 system, (12) can be transformed in to

$$\begin{aligned} (I_n \otimes E)\dot{\bar{x}}(t) &= \tilde{H}\bar{x}(t) = (I_n \otimes A - J_L \otimes BK_2)\bar{x}(t) \\ &= \begin{bmatrix} H_{\lambda_1} & * & & & \\ & H_{\lambda_2} & * & & \\ & & \ddots & \ddots & \\ & & & H_{\lambda_{n-1}} & * \\ & & & & H_{\lambda_n} \end{bmatrix} \bar{x}(t), \end{aligned} \quad (13)$$

where $*$ denotes a block in $\mathbb{R}^{m \times m}$ that may either be $-BK_2$ or 0 .

Definition 11 (out stability). For an LTI singular system

$$\begin{aligned} E\dot{x} &= Ax, \\ y &= Cx, \end{aligned} \quad (14)$$

suppose that, for $\forall \varepsilon > 0$, $\exists \delta(\varepsilon) > 0$ s.t. $\|y(t)\| < \varepsilon$ ($t > 0$) as $\|y(0)\| < \delta(\varepsilon)$. Then the system is said to be output stable with respect to y .

Define the auxiliary vectors $\eta_i(t) = \sum_{j \in n_i} w_{ij}(x_i(t) - x_j(t))$ ($i \in \{1, 2, \dots, n\}$) and the stack of them as $\eta(t) = [\eta_1^T(t), \eta_2^T(t), \dots, \eta_n^T(t)]^T$. It follows that

$$\eta = (L \otimes I_m)x(t). \quad (15)$$

As a result of (15) and $\bar{x}(t) = (U^{-1} \otimes I_n)x(t)$,

$$\begin{aligned} (U^{-1} \otimes I_n)\eta &= (U^{-1} \otimes I_n)(L \otimes I_m)x(t) = (U^{-1}L \otimes I_m)x(t) \\ &= (U^{-1}L \otimes I_m)(U \otimes I_n)\bar{x}(t) = (J \otimes I_n)\bar{x}(t). \end{aligned} \quad (16)$$

For the system (9), we said that the system is stable like the normal linear system case; when studying stability of singular linear systems, we need only to consider the following homogeneous equation:

$$E\dot{x}_i(t) = Ax_i(t), \quad (17)$$

where $E \in \mathbb{R}^{m \times m}$, $A \in \mathbb{R}^{m \times m}$, $B \in \mathbb{R}^{m \times d}$, with $\text{rank}(E) = c \leq m$, and $x_i(0)$ is the admissible initial condition.

System (17) satisfies the regularity condition, and there exists a scalar α such that $\det(\alpha E + A) \neq 0$. The system (17) can be transformed into the equivalent system

$$\tilde{E}\dot{x}(t) = \tilde{A}x(t), \quad (18)$$

where $\tilde{E} = (\alpha E + A)^{-1}E$, $\tilde{A} = (\alpha E + A)^{-1}A$. The regularity condition assures that the system (17) has solution (see [33])

$$x(t) = e^{\tilde{E}^D \tilde{A}t} \tilde{E}^D \tilde{E}x(0) \quad (19)$$

and $x(0)$ is admissible initial condition.

By Lemmas 2 and 6, there exist two nonsingular matrices $Q, P \in \mathbb{R}^{mn \times mn}$; we can get

$$\begin{aligned} x(t) &= Pe^{\begin{bmatrix} A_1 & \\ O_{n_2} & \end{bmatrix} t} \begin{bmatrix} I_{n_1} & \\ O_{n_2} & \end{bmatrix} P^{-1}x(0) \\ &= P \begin{bmatrix} e^{A_1 t} & \\ O_{n_2} & \end{bmatrix} P^{-1}x(0), \end{aligned} \quad (20)$$

where $A_1 \in \mathbb{R}^{n_1 \times n_1}$, $n_1 + n_2 = nm$.

Remark 12. From the structure of the solution we can see that $\tilde{E}^D \tilde{A}$ has two-part eigenvalues, eigenvalues of part A_1 and eigenvalues of part O_{n_2} . Eigenvalues of part A_1 are effective eigenvalues. In this current paper slow subsystem eigenvalues of (E, A) represent eigenvalues of part A_1 . Let (E, A) represent $\tilde{E}^D \tilde{A}$.

Definition 13 (see [48]). Consider the system (17) or the matrix pair (E, A) where $E \in \mathbb{R}^{m \times m}$, $A \in \mathbb{R}^{m \times m}$.

- (1) A pole of the system, or a finite eigenvalue of the matrix pair (E, A) , is called stable if it has negative real part, is called critically stable if it has zero real part, and is called unstable if it has positive real part.

(2) The matrix pair (E, A) is called stable if it has only stable finite eigenvalues, is called critically stable if it has critical finite eigenvalues but not unstable ones, and is called unstable if it has unstable finite eigenvalues.

Remark 14. For a clear expression, the LTI singular systems is said to be stable if 0 is a stable equilibrium point of this system.

Definition 15 (swarm stability). Consider LTI singular multiagent systems with n agents and $x_1, x_2, \dots, x_n \in \mathbb{R}^m$ as the agent states. The system is said to be **swarm stable** if for $\forall \varepsilon > 0$, $\exists \delta(\varepsilon) > 0$ s.t. $\|x_i(t) - x_j(t)\| < \varepsilon$ ($t > 0$) as $\|x_i(t) - x_j(t)\| < \delta(\varepsilon)$ ($\forall i, j \in \{1, 2, \dots, n\}$).

Definition 16 (asymptotically swarm stability). Consider that LTI singular multiagent systems achieve full state consensus, that is, for $\forall \varepsilon > 0$, $\exists T > 0$ s.t. when $t > T$, $\|x_i(t) - x_j(t)\| < \varepsilon$ ($\forall i, j \in \{1, 2, \dots, n\}$). Then it is said to be **asymptotically swarm stable**.

The swarm stability discussed in this paper is LTI singular systems, and the definition of swarm stability and asymptotic swarm stability seem similar to [18]; however, they are actually different. The current paper focuses on the more general cases which is LTI singular systems. As well known for the system (10), when $\text{rank}(E) = m$ the normal systems are actually spacial cases of the system (10).

Besides consensus achievement, some phenomena investigated by the researchers, for example, flocking and formation keeping, also require a system to be swarm stable. A bird flock, vehicle platoon, or robot crew may navigate to anywhere, but the distances among its members should not go unbounded. The states of agents in a swarm stable system might still oscillate or even diverge. Swarm stability is a kind of nonequilibrium stability [18].

3. Swarm Stability

The major purpose of this section is to present a necessary and sufficient condition for the swarm stability of high-order LTI singular swarm systems. For this end, several preparations are needed.

Lemma 17. *For multiagent systems (9), if it is asymptotically swarm stable, then as $t \rightarrow \infty$, the trajectory of each agent satisfies $E\dot{\xi} - A\xi \rightarrow 0$.*

Proof. Without loss of generality, consider an agent i . The difference between $E\dot{x}_i$ and Ax_i is $E\dot{x}_i - Ax_i = BK_2 \sum_{j \in N_i} w_{ij}(x_j - x_i)$. Since the system is asymptotically swarm stable, $\lim_{t \rightarrow \infty} (x_j - x_i) = 0$ ($j \in \{1, 2, \dots, n\}$) and $\lim_{t \rightarrow \infty} BK_2 \sum_{j \in N_i} w_{ij}(x_j - x_i) = 0$. Thus, as $t \rightarrow \infty$, the difference $(E\dot{x}_i - Ax_i) \rightarrow 0$. \square

Lemma 17 naturally leads to the following two corollaries.

Corollary 18. *For multiagent System (9), if each slow subsystem eigenvalue of (E, A) has negative real part, then the system*

is asymptotically swarm stable if and only if it is asymptotically stable and the limit of all agents state that is the consensus state, is zero.

Proof. For the LTI singular systems (9), like the normal linear system case, when studying stability of singular linear systems, we need only to consider the following homogeneous equation:

$$E\dot{x}_i = Ax_i \quad (21)$$

for $i \in \{1, 2, \dots, n\}$. The regularity condition assures that the system (9), by the Lemma 8 we can get

$$x_i(t) = e^{\tilde{E}^D \tilde{A}} \tilde{E}^D \tilde{E} x_i(0) \quad (22)$$

with $\tilde{E} = (\alpha E + A)^{-1} E$ and $\tilde{A} = (\alpha E + A)^{-1} A$, $x_i(0)$ is the admissible initial condition. As we can see that each slow subsystem eigenvalue of (E, A) has negative real part, the limit $\lim_{t \rightarrow \infty} x_i(t) = \lim_{t \rightarrow \infty} e^{\tilde{E}^D \tilde{A}} \tilde{E}^D \tilde{E} x_i(0) = 0$, so the consensus state, is zero and it is asymptotically stable with Lemma 17 from which we can get the results. \square

Corollary 19. *For singular multi agent system (9), the system is asymptotically swarm stable and the consensus state is not zero. Then not each slow subsystem eigenvalues of (E, A) has negative real part.*

From the proof of Corollary 18, we can easily get the results.

Lemma 20. *For LTI singular multi agent system (9), if (E, A) is stable, then swarm stability of the system is equivalent to stable.*

Proof. Suppose that the system is swarm stable. If $\|x_i(0)\| < \delta/2$ ($i \in \{1, 2, \dots, n\}$), then $\|x_i(0) - x_j(0)\| < \delta$ ($i, j \in \{1, 2, \dots, n\}$). Owning to the swarm stability, for $\forall \varepsilon > 0$, $\exists \delta(\varepsilon)$ s.t. $\|x_i(t) - x_j(t)\| < \varepsilon$ ($j \in \{1, 2, \dots, n\}$) as $\|x_i(0)\| < \delta/2$ ($i \in \{1, 2, \dots, n\}$). ε can be arbitrarily small if with proper δ . Let $\varepsilon \rightarrow 0$. According to the similar idea in the proof of Lemma 17, the motion of each agent can satisfy $E\dot{\xi} - A\xi \rightarrow 0$. Assume that the swarm system is unstable. Then the agent states may diverge since the system is LTI. This contradicts the fact that (E, A) is stable. Thus, the multi agent system must be stable.

Suppose that the system is stable. Assume that the system is swarm unstable. There may exist a pair of agents i and j with $(x_i - x_j) \rightarrow \infty$ as $t \rightarrow \infty$. This implicates that at least the state of one agent goes unbounded. The assumption contradicts the fact that the system is stable. Thus, the system is swarm stable. \square

Lemma 21. *If the graph G includes a spanning tree, then a necessary and sufficient condition for the swarm stability of the singular system is that it is output stable with respect to η .*

Proof. Owing to the definition of η the system is output stable with respect to η if it is swarm stable. According to Lemma 20,

if the graph G includes a spanning tree, then there exists certain nonsingular matrix $P \in \mathbb{R}^{n \times n}$ s.t.

$$PL = \Gamma = \begin{bmatrix} 1 & -1 & 0 & & \\ 0 & 1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 & 0 \\ & & 0 & 1 & -1 \\ * & * & \cdots & \cdots & * \end{bmatrix}, \quad (23)$$

where the first $n - 1$ rows in Γ form a basis for the row space of L and $*$ can be any value. Thus, (15) can be transformed into $(P \otimes I_d)\eta = (\Gamma \otimes I_d)x$. All the vectors $x_1 - x_2, x_2 - x_3, \dots, x_{n-1} - x_n$ are dependent on η . The value of any $\|x_i - x_j\|$ ($i, j \in \{1, 2, \dots, n\}$) can be arbitrarily small if η is sufficiently close to zero. Therefore, the system is swarm stable if it is output stable with respect to η . \square

This section will establish some necessary and sufficient conditions guaranteeing that the system has a consensus property.

Theorem 22. For swarm system (9) with $\lambda_1 = 0, \lambda_2, \dots, \lambda_n \in \mathbb{C}$ as the eigenvalues of $L(G)$, if (E, A) is unstable, the system is swarm stable if and only if both (1) and (2) below are true:

- (1) G includes a spanning tree;
- (2) all the $(E, A - \lambda_i BK_2)$, $(\lambda_i \neq 0)$, are stable. Besides, if for some i , $(E, A - \lambda_i BK_2)$ is critically stable and L is not diagonalisable, each submatrix in $(I_n \otimes E, I_n \otimes A - J \otimes BK_2)$ corresponding to a Jordan block of λ_i which has the form

$$\left(\begin{bmatrix} E & & & \\ & E & & \\ & & \ddots & \\ & & & E \end{bmatrix}, \begin{bmatrix} A - \lambda_i BK_2 & -BK_2 & 0 & & \\ 0 & A - \lambda_i BK_2 & \ddots & & 0 \\ \vdots & & \ddots & & -BK_2 \\ 0 & \dots & 0 & A - \lambda_i BK_2 & \end{bmatrix} \right) \quad (24)$$

is stable.

If (E, A) is stable, then the system is swarm stable if and only if condition (2) is true.

Proof. **(E, A) Is Unstable.** Assume that G does not include a spanning tree while the system is still swarm stable. $\exists \delta(\varepsilon) > 0$ s.t. $\|x_i(t) - x_j(t)\| < \varepsilon$ if $\|x_i(0) - x_j(0)\| < \delta$ ($\forall i, j \in \{1, 2, \dots, n\}$). G must contain $k \geq 2$ different subgraphs $\widehat{G}_1, \widehat{G}_2, \dots, \widehat{G}_k$ each receiving no information. According to the basic idea in the proof of Lemmas 17 and 20, ε can be arbitrarily small if δ is small enough. Let $\varepsilon \rightarrow 0$. Then the motion of each agent associated with \widehat{G}_1 can satisfy $E\xi_1 - A\xi_1 \rightarrow 0$, whereas the motion of each agent associated with \widehat{G}_2 can satisfy $E\xi_2 - A\xi_2 \rightarrow 0$. As ξ_1 and ξ_2 are independent, $E\xi_1 - E\xi_2 - A(\xi_1 - \xi_2) \rightarrow 0$. Since (E, A) is unstable and the system is LTI, there exist $\xi_1(0) \neq \xi_2(0)$ s.t. $\xi_1(t) - \xi_2(t) \rightarrow \infty$

($t \rightarrow \infty$). This contradicts the fact that the system is swarm stable. Thus, Condition (1) is necessary.

Owing to Lemma 21, when condition (1) is true, swarm stability is equivalent to that the system is output stable with respect to η . Notice that $\lambda_2, \dots, \lambda_n$ are all nonzero. According to the linear equation (16) and the upper triangular structure of J , the output stability with respect to η is equivalent to the output stability with respect to $\tilde{x}_2, \tilde{x}_3, \dots, \tilde{x}_n$. Evidently, one can infer that condition (1) is a necessary and sufficient condition for such an output stability by considering the structure of (13).

(E, A) Is Stable. According to Lemma 20, the swarm stability of the system is equivalent to the asymptotically stability of (10). One can draw the conclusion that (10) is stable if and only if Condition (2) is true by considering the structure of (13). \square

Remark 23. For system (9), if $\text{Index}(E) = 0$, the results in [18] is a special case of the Theorem 22.

Corollary 24. For swarm system (10) with $\lambda_1 = 0, \lambda_2, \dots, \lambda_n \in \mathbb{C}$ as the eigenvalues of $L(G)$, if $(E, A + BK_1)$ is unstable, the system is swarm stable if and only if both (1) and (2) below are true:

- (1) G includes a spanning tree;
- (2) all the $(E, A + BK_1 - \lambda_i BK_2)$, $\lambda_i \neq 0$, are stable. Besides, if for some i , $(E, A + BK_1 - \lambda_i BK_2)$ is critically stable and L is not diagonalisable, each submatrix in $(I_n \otimes E, I_n \otimes (A + BK_1) - J \otimes BK_2)$ corresponding to a Jordan block of λ_i which has the form

$$\left(\begin{bmatrix} E & & & \\ & E & & \\ & & \ddots & \\ & & & E \end{bmatrix}, \begin{bmatrix} A + BK_1 - \lambda_i BK_2 & -BK_2 & 0 & & \\ 0 & A + BK_1 - \lambda_i BK_2 & \ddots & & 0 \\ \vdots & & \ddots & & -BK_2 \\ 0 & \dots & 0 & A + BK_1 - \lambda_i BK_2 & \end{bmatrix} \right) \quad (25)$$

is stable.

If $(E, A + BK_1)$ is stable, then the system is swarm stable if and only if condition (2) is true.

The proof is the same as that Theorem 22.

4. Asymptotic Swarm Stability

Theorem 25. For swarm system (9) with $\lambda_1 = 0, \lambda_2, \dots, \lambda_n \in \mathbb{C}$ as the eigenvalues of the Laplacian matrix $L(G)$, if not each slow subsystem eigenvalues of (E, A) has negative real part, the system is asymptotically swarm stable if and only if both (1) and (2) below are true:

- (1) the graph topology G includes a spanning tree.
- (2) each slow subsystem eigenvalue of $(E, A - \lambda_i BK_2)$ ($\lambda_i \neq 0$) has negative real part.

If each slow subsystem eigenvalue of (E, A) has negative real part, then the system is asymptotically swarm stable if and only if condition (2) is true.

Proof. Part 1. Not each slow subsystem eigenvalue of (E, A) has negative real part. According to the same approach in the proof of Theorem 22, it is easy to prove that Condition (1) is necessary. With Condition (1), according to Lemma 20, any vector in the null space of $L \otimes I_d$ is of the form $\phi \otimes \xi$ where $\phi = [1, 1, \dots, 1]^T \in \mathbb{R}^m$ and $\xi \in \mathbb{R}^m$. Therefore $x_1 = x_2 = \dots = x_n$ if and only if $\eta = 0$. According to Lemma 20, $\lambda_2, \dots, \lambda_n$ are all nonzero. From the linear equation (16), one knows that $\eta = 0$ if and only if $\tilde{x}_2 = \dots = \tilde{x}_n = 0$ with considering the triangular structure of J . According to (13), $I_n \otimes A - J \otimes BK_2$ is a strictly upper triangular block matrix and the stability of (10) is determined by the eigenvalues of $(E, A - \lambda_i BK_2)$. Meanwhile, the output stability of the system (9) respect to $\tilde{x}_2, \tilde{x}_3, \dots, \tilde{x}_n$ is independent of \tilde{x}_1 . Consequently, $\tilde{x}_2, \tilde{x}_3, \dots, \tilde{x}_n \rightarrow 0$ ($t \rightarrow \infty$) if and only if each slow subsystem eigenvalue of $(E, A - \lambda_2 BK_2), \dots, (E, A - \lambda_n BK_2)$ has negative real part.

Part 2. Each slow subsystem eigenvalue of (E, A) has negative real part. When each slow subsystem eigenvalue of (E, A) has negative real part, according to Corollary 18, the multiagent system's asymptotic swarm stability is equivalent to its asymptotic stability. The stability of system (9) is equivalent to that of (10). By observing the structure of $I_n \otimes A - J \otimes BK_2$ in (13), it is easy to infer that (10) is asymptotically stable if and only if condition (2) is true. \square

Remark 26. For system (9), if $\text{Index}(E) = 0$, the results in [18, 34] is a special case of Theorem 25.

Corollary 27. For swarm system (10) with $\lambda_1 = 0, \lambda_2, \dots, \lambda_n \in \mathbb{C}$ as the eigenvalues of the Laplacian matrix $L(G)$, if not each slow subsystem eigenvalue of $(E, A + BK_1)$ has negative real part, the system is asymptotically swarm stable if and only if both (1) and (2) below are true:

- (1) the graph topology G includes a spanning tree.
- (2) each slow subsystem eigenvalue of $(E, A + BK_1 - \lambda_i BK_2)$ ($\lambda_i \neq 0$) has negative real part.

If each slow subsystem eigenvalue of $(E, A + BK_1)$ has negative real part, then the system is asymptotically swarm stable if and only if condition (2) is true.

Remark 28. For system (10), the result in [18, 32, 35] is a special case of the Corollary 27.

5. Simulation

In this section, numerical instances will be exhibited to illustrate the theoretical results in the previous section, three graphs G_1 , G_2 , and G_3 .

5.1. First Example. A network with three nodes and four edges, whose topology is shown in Figure 1, the eigenvalues

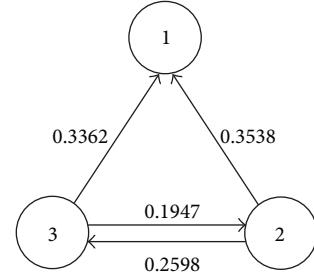


FIGURE 1: Direction interaction topology graph G_1 .

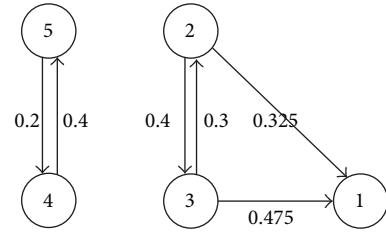


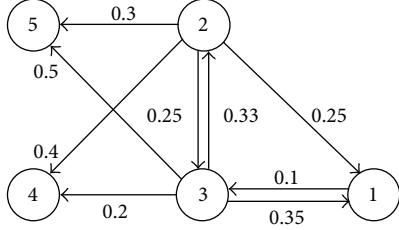
FIGURE 2: Direction interaction topology graph G_2 .

of the Laplacian matrices for the two graphs are $\lambda(G_1) = \{0, 0.4545, 0.69\}$ and

$$\begin{aligned} E &= \begin{bmatrix} 0 & 1 & 0 \\ -5 & 2 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \\ A &= \begin{bmatrix} 7.4197 & 0.152 & 0 \\ 5.5994 & 3.4841 & 0 \\ 1 & 2 & 1 \end{bmatrix}, \\ BK_2 &= \begin{bmatrix} 1 & 4 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned} \quad (26)$$

In this example, (E, A) is unstable with slow subsystem eigenvalues $1 \pm 2i$, $\text{rank}(E) = 2 < 3$ the multiagent system is singular multiagent system, whereas the slow subsystem eigenvalues of $(E, A + \lambda_i BK_2)$ ($i = 2, 3$) are $-0.518 \pm 0.917i$ and $\pm 1.512i$. According to Theorems 22 and 25, the multiagent system is swarm stable but not asymptotically swarm stable.

Figure 4 shows the three trajectories of the agents with nonadmissible initial condition in the three dimension space. The three agents diverge, while they neither escape from each other nor achieve consensus. For clearer observation, Figure 5 shows the three trajectories representing the relative motion, that is, $x_1 - x_2$, $x_2 - x_3$ and $x_1 - x_3$. It is clear that the relative motions are Lyapunov stable. The dots \bullet represent the nonadmissible initial condition position which we set, and the dots $*$ is the admissible initial condition which the system actually starts. Figure 4 indicates that the nonadmissible initial bounded has no effect on swarm stable if the multiagent can accept the nonadmissible initial bounded value.

FIGURE 3: Direction interaction topology graph G_3 .

5.2. Second Example. For a network with five nodes and six edges, whose topology is shown in Figure 2, the eigenvalues of the Laplacian matrices for the two graphs are $\lambda(G_2) = \{0, 0, 0.6, 0.7, 0.8\}$ and

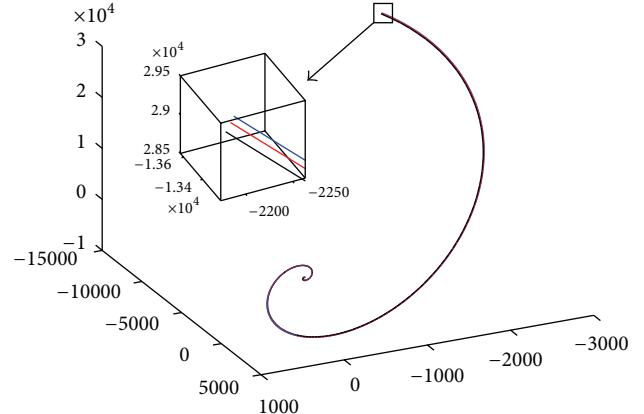
$$\begin{aligned} E &= \begin{bmatrix} 0 & 1 & 0 \\ -5 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\ A &= \begin{bmatrix} 7.4197 & 0.152 & 0 \\ 5.5994 & 3.4841 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad (27) \\ BK_2 &= \begin{bmatrix} 1 & 4 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

In this example, (E, A) is unstable with slow subsystem eigenvalues $1 \pm 2i$, $\text{rank}(E) = 2 < 3$ the multiagent system is singular swarm systems, whereas the slow subsystem eigenvalues of $(E, A + \lambda_i BK_2)$ ($\lambda_i \neq 0$) are $-0.32 \pm 1.1994i, -0.54 \pm 0.8780i, -0.76 \pm 0.1857i$, Although the the slow subsystem eigenvalues of $(E, A + \lambda_i BK_2)$ ($\lambda_i \neq 0$) are all with negative real parts, the system is swarm unstable according to Theorem 22 for (E, A) is unstable and G_2 includes no spanning tree. For G_2 has two subgraph, each subgraph has a spanning tree, so they depart two groups. * is the admissible initial bounded value position (see Figure 6).

5.3. Third Example. For a network with five nodes and nine edges, whose topology is shown in Figure 3, the eigenvalues of the Laplacian matrices for the two graphs are $\lambda(G_3) = \{0, 0.7, 0.58, 0.6, 0.8\}$ and

$$\begin{aligned} E &= \begin{bmatrix} 0 & 1 & 0 \\ -5 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\ A &= \begin{bmatrix} 7.4197 & 0.152 & 0 \\ 5.5994 & 3.4841 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad (28) \\ BK_2 &= \begin{bmatrix} 1 & 4 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

In this example, (E, A) is unstable with slow subsystem eigenvalues $1 \pm 2i$, $\text{rank}(E) = 2 < 3$ the multiagent system is sin-

FIGURE 4: Trajectories of three agents in the first instance with non-admissible bounded initial $t \in [0 \ 7]$.

gular swarm systems, whereas the slow subsystem eigenvalues of $(E, A + \lambda_i BK_2)$ ($\lambda_i \neq 0$) are $-0.276 \pm 1.2505i, -0.54 \pm 0.8780i, -0.32 \pm 0.1194i, -0.76 \pm 0.1857i$ although the slow subsystem eigenvalue of (E, A) ($\lambda_i \neq 0$) is unstable and the system matches the condition of Theorem 25, and it is asymptotically swarm stable. Figure 7 shows the five trajectories of the agents in the three-dimensional space. For a clear observation Figure 8 shows the four trajectories representing the relative motions that is $x_1 - x_2, x_2 - x_3, x_3 - x_4, x_4 - x_5$. It clear that the relative motions are asymptotically swarm stable.

5.4. Forth Example. For a network with five nodes and nine edges, whose topology is shown in Figure 3, the eigenvalues of the Laplacian matrices for the two graphs are $\lambda(G_3) = \{0, 0.7, 0.58, 0.6, 0.8\}$ and

$$\begin{aligned} E &= \begin{bmatrix} 0 & 1 & 0 \\ -5 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\ A &= \begin{bmatrix} 7.4197 & 0.152 & 0 \\ 10.5994 & 1.4841 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad (29) \\ BK_2 &= \begin{bmatrix} 1 & 4 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

In this example, (E, A) is critically stable with slow subsystem eigenvalues $\pm 2i$, $\text{rank}(E) = 2 < 3$ the multiagent system is singular swarm systems, whereas the slow subsystem eigenvalues of $(E, A + \lambda_i BK_2)$ ($\lambda_i \neq 0$) are $-1.54 \pm 0.878i, -1.276 \pm 1.2505i, -1.32 \pm 1.994i, -1.76 \pm 0.1857i$ although the the slow subsystem eigenvalues of (E, A) ($\lambda_i \neq 0$) is critically stable and the system matches the conditions of Theorem 25. It is asymptotically swarm stable and they achieve consensus with the consensus function is time varying. Figure 9 shows the five trajectories of the agents in the three-dimensional space. * is the starting positions of agents.

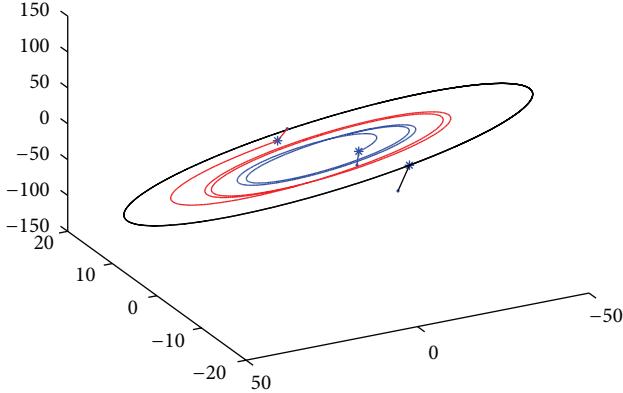


FIGURE 5: Relative motions in the first instance with non-admissible bounded initial $t \in [0, 7]$.

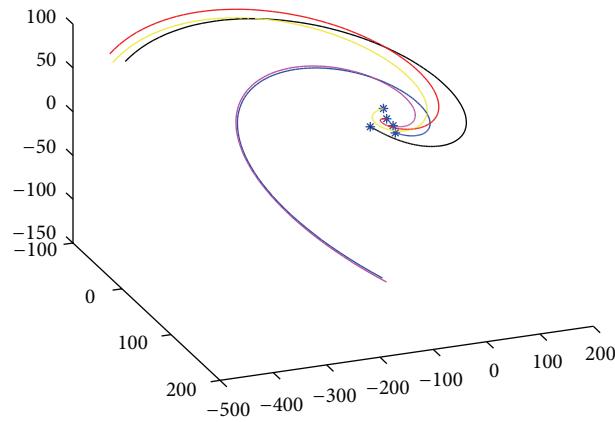


FIGURE 6: Trajectories of five agents in the second instance with admissible bounded initial $t \in [0, 3.2]$.

5.5. Fifth Example. For a network with three nodes and four edges, whose topology is shown in Figure 1, the eigenvalues of the Laplacian matrices for the two graphs are $\lambda(G_1) = \{0, 0.4545, 0.69\}$ and

$$E = \begin{bmatrix} 0 & 1 & 8 & 2 & 1 & 0 \\ -5 & 2 & 6 & 1 & 1 & -1 \\ -2 & -4 & -8 & 8 & 0 & 0 \\ -1 & -2 & -4 & -2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$BK_2 = \begin{bmatrix} -1 & -4 & -5 & 4 & 5 & -6 \\ -20 & 7 & 4 & 8 & 0 & -5 \\ -2 & -2 & 8 & -2 & 0 & 0 \\ -8 & -1 & -5 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A = \begin{bmatrix} 1.2817 & -7.6558 & -19.2807 & -1.3580 & -2.1439 & -7.6522 \\ 3.1890 & -8.7816 & -54.2009 & -45.3840 & 19.6020 & -4.2240 \\ -11.1773 & 0.6922 & -9.5160 & -34.5783 & 15.7559 & 1.6546 \\ -0.0788 & 1.0435 & 2.3770 & -13.6672 & 1.4261 & -1.7988 \\ 2 & 1 & -2 & 4 & 4 & 2 \\ 2 & 6 & 8 & 8 & 5 & 5 \end{bmatrix}. \quad (30)$$

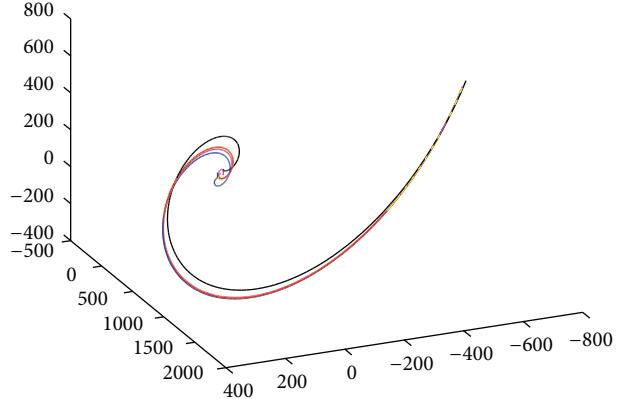


FIGURE 7: Trajectories of five agents in the third instance with admissible bounded initial $t \in [0, 5]$.

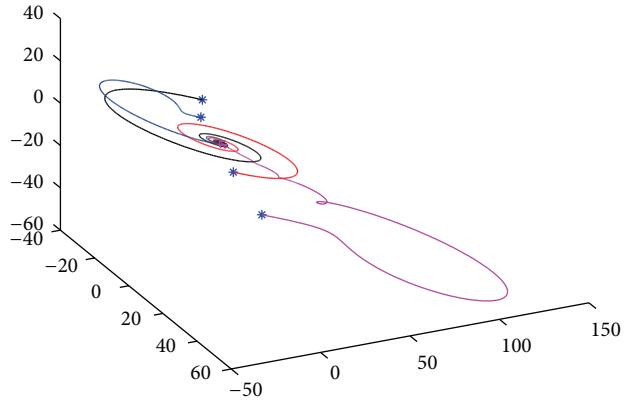


FIGURE 8: Relative motions in the third instance with admissible bounded initial $t \in [0, 5]$.

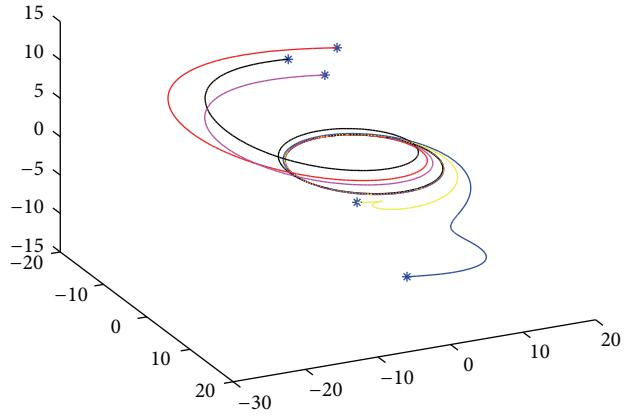


FIGURE 9: Trajectories of five agents in the forth instance with admissible bounded initial $t \in [0, 7]$.

In this example, (E, A) is critically stable with slow subsystem eigenvalues $\pm 1.5773i, \pm 3.2057i$, $\text{rank}(E) = 4 < 6$ the multiagent system is singular swarm systems, whereas the slow subsystem eigenvalue of $(E, A + \lambda_i BK_2)$ ($\lambda_i \neq 0$) are $-1.0998 \pm 1.0758i, -0.5158 \pm 4.1064i, -1.7073 \pm 0.6370i, -0.7455 \pm 4.4668i$, the slow subsystem eigenvalues of (E, A) ($\lambda_i \neq 0$) is

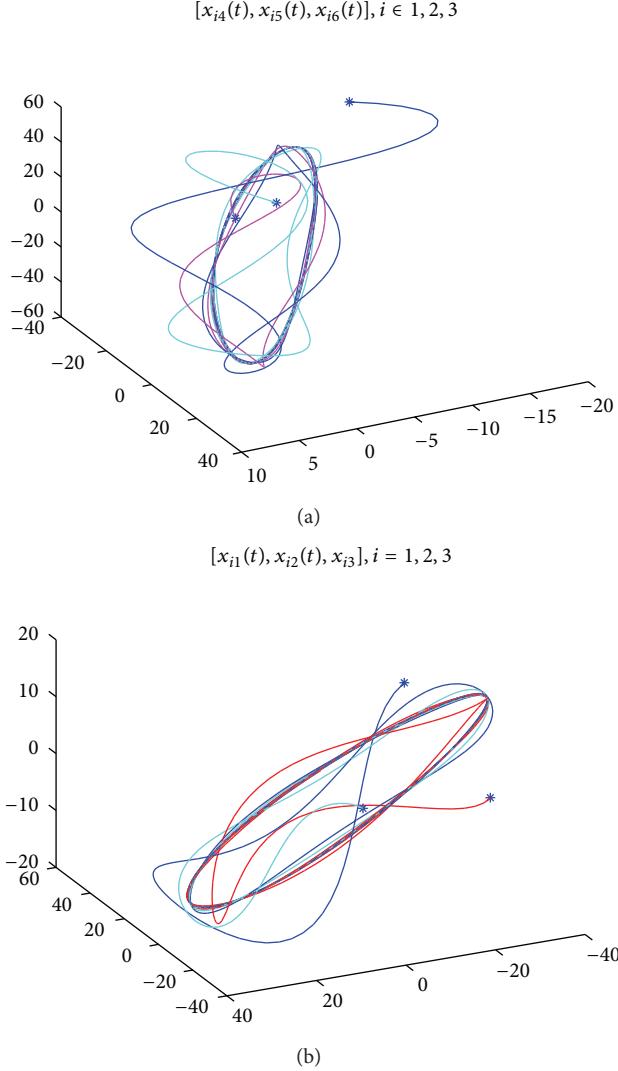


FIGURE 10: Trajectories of three agent's $[x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}, x_{i6}]$ in the fifth instance with admissible bounded initial.

critically stable and the system matches the conditions of Theorem 25. It is asymptotically swarm stable and they achieve consensus with the consensus function is time varying. Figure 10, shows the trajectories of the i th agent's state $[x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}, x_{i6}]^T$. * is the starting positions of agents.

6. Conclusions

The swarm stability problem of high-order LTI singular multiagent systems was solved through a method based on generalized inverse theory. Consensus is regarded as a specific type of swarm stability. The model considered in this paper is more general than the one in [18], and most of LTI homogeneous normal or singular multiagent system models appeared in the literature are specific cases of this one. A necessary and sufficient condition for swarm stability was presented via analyzing of the solution of the singular multiagents systems with the *Drazin* inverse theory. Swarm stability is a kind of

nonequilibrium stability. Meanwhile Simulation results for reaching consensus are presented. From the simulation we can get that if the agent system can accept nonadmissible initial bounded value, the multiagent systems can be swarm stability with any initial bounded value.

References

- [1] W. K. Potts, "The chorus-line hypothesis of manoeuvre coordination in avian flocks," *Nature*, vol. 309, no. 5966, pp. 344–345, 1984.
- [2] J. K. Parrish, S. V. Viscido, and D. Grünbaum, "Self-organized fish schools: an examination of emergent properties," *Biological Bulletin*, vol. 202, no. 3, pp. 296–305, 2002.
- [3] T. Vicsek, A. Czirok, E. B. Jacob, I. Cohen, and O. Schonch, "Novel typeof phase transitions in a system of self-driven particles," *Physical Review Letters*, vol. 75, pp. 1226–1229, 1995.
- [4] A. Abdessameud and A. Tayebi, "Attitude synchronization of a group of spacecraft without velocity measurements," *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2642–2648, 2009.
- [5] R. Olfati-Saber, "Global configuration stabilization for the VTOL aircraft with strong input coupling," *IEEE Transactions on Automatic Control*, vol. 47, no. 11, pp. 1949–1952, 2002.
- [6] J. A. Fax and R. M. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1465–1476, 2004.
- [7] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: algorithms and theory," *IEEE Transactions on Automatic Control*, vol. 51, no. 3, pp. 401–420, 2006.
- [8] P. Ögren, E. Fiorelli, and N. E. Leonard, "Cooperative control of mobile sensor networks: adaptive gradient climbing in a distributed environment," *IEEE Transactions on Automatic Control*, vol. 49, no. 8, pp. 1292–1302, 2004.
- [9] K. Jin, P. Liang, and G. Beni, "Stability of synchronized distributed control of discrete swarm structures," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 1033–1038, San Diego, Calif, USA, May 1994.
- [10] S. Ma, S. Hackwood, and G. Beni, "Multi-agent supporting systems (MASS): control with centralized estimator of disturbance," in *Proceedings of the IEEE/RSJ/GI International Conference on Intelligent Robots and Systems, Advanced Robotic Systems and the Real World (IROS '94)*, pp. 679–686, Munich, Germany.
- [11] Y. Liu, K. M. Passino, and M. Polycarpou, "Stability analysis of one-dimensional asynchronous swarms," *IEEE Transactions on Automatic Control*, vol. 48, no. 10, pp. 1848–1854, 2003.
- [12] Y. Liu, K. M. Passino, and M. M. Polycarpou, "Stability analysis of M -dimensional asynchronous swarms with a fixed communication topology," *IEEE Transactions on Automatic Control*, vol. 48, no. 1, pp. 76–95, 2003.
- [13] V. Gazi and K. M. Passino, "A class of attraction/repulsion functions for stable swarm aggregations," in *Proceedings of the 41st IEEE Conference on Decision and Control*, pp. 2842–2847, Las Vegas, Nev, USA, December 2002.
- [14] V. Gazi and K. M. Passino, "Stability analysis of swarms," *IEEE Transactions on Automatic Control*, vol. 48, no. 4, pp. 692–697, 2003.
- [15] V. Gazi, "Stability of an asynchronous swarm with time-dependent communication links," *IEEE Transactions on Systems, Man, and Cybernetics B*, vol. 38, no. 1, pp. 267–274, 2008.

- [16] T. Chu, L. Wang, and T. Chen, "Self-organized motion in anisotropic swarms," *Journal of Control Theory and Applications*, vol. 1, no. 1, pp. 77–81, 2003.
- [17] W. Li, "Stability analysis of swarm with general topology," *IEEE Transactions on Systems, Man, and Cybernetics B*, vol. 38, no. 4, pp. 1084–1097, 2008.
- [18] N. Cai, J.-X. Xi, and Y.-S. Zhong, "Swarm stability of high-order linear time-invariant swarm systems," *IET Control Theory & Applications*, vol. 5, no. 2, pp. 402–408, 2011.
- [19] D. Swaroop, "Stability analysis of swarms with general topology," *IEEE Transactions on Systems, Man, and Cybernetics B*, vol. 4, no. 4, pp. 1084–1097, 2008.
- [20] A. Pant, P. Seiler, and J. k. Hedrick, "String stability for a class of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 2, pp. 349–357, 1996.
- [21] L. Zhou, D. W. C. Ho, and G. S. Zhai, "Stability analysis of switched linear singular systems," *Automatica*, vol. 49, no. 5, pp. 1481–1487, 2013.
- [22] A. Jadbabaie, J. 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.
- [23] 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.
- [24] 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.
- [25] F. Xiao, L. Wang, and A. Wang, "Consensus problems in discrete-time multiagent systems with fixed topology," *Journal of Mathematical Analysis and Applications*, vol. 322, no. 2, pp. 587–598, 2006.
- [26] P. Lin and Y. Jia, "Consensus of second-order discrete-time multi-agent systems with nonuniform time-delays and dynamically changing topologies," *Automatica*, vol. 45, no. 9, pp. 2154–2158, 2009.
- [27] W. Yu, G. Chen, and M. Cao, "Some necessary and sufficient conditions for second-order consensus in multi-agent dynamical systems," *Automatica*, vol. 46, no. 6, pp. 1089–1095, 2010.
- [28] Y.-P. Tian and Y. Zhang, "High-order consensus of heterogeneous multi-agent systems with unknown communication delays," *Automatica*, vol. 48, no. 6, pp. 1205–1212, 2012.
- [29] Y. Zheng and L. Wang, "Finite-time consensus of heterogeneous multi-agent systems with and without velocity measurements," *Systems & Control Letters*, vol. 61, no. 8, pp. 871–878, 2012.
- [30] F. Xiao and L. Wang, "Consensus problems for high-dimensional multi-agent systems," *IET Control Theory and Applications*, vol. 1, no. 3, pp. 830–837, 2007.
- [31] L. Cheng, Z.-G. Hou, Y. Lin, M. Tan, and W. Zhang, "Solving a modified consensus problem of linear multi-agent systems," *Automatica*, vol. 47, no. 10, pp. 2218–2223, 2011.
- [32] J. Xi, N. Cai, and Y. Zhong, "Consensus problems for high-order linear time-invariant swarm systems," *Physica A*, vol. 389, no. 24, pp. 5619–5627, 2010.
- [33] L. Dai, *Singular Control Systems*, vol. 118 of *Lecture Notes in Control and Information Sciences*, Springer, Berlin, Germany, 1989.
- [34] X.-R. Yang and G.-P. Liu, "Necessary and sufficient consensus conditions of descriptor multi-agent systems," *IEEE Transactions on Circuits and Systems*, vol. 59, no. 11, pp. 2669–2677, 2012.
- [35] J. X. Xi, Z. Y. Shi, and Y. S. Zhong, "Admissible consensus and consensualization of high order linear time-invariant singular swarm systems," *Physica A*, vol. 391, pp. 5839–5849, 2012.
- [36] X. Dong, J. Xi, G. Lu, and Y. Zhong, "Containment analysis and design for high-order linear time-invariant singular swarm systems with time delays," *International Journal of Robust and Nonlinear Control*, vol. 01, 2012.
- [37] J. Xi, F. Meng, Z. Shi, and Y. Zhong, "Delay-dependent admissible consensualization for singular time-delayed swarm systems," *Systems & Control Letters*, vol. 61, no. 11, pp. 1089–1096, 2012.
- [38] M. P. Drazin, "Pseudo-inverses in associative rings and semi-groups," *The American Mathematical Monthly*, vol. 65, pp. 506–514, 1958.
- [39] S. L. Campbell and C. D. Meyer, *Generalized Inverse of Linear Transformations*, Pitman, London, 1979, SIAM, Philadelphia, Pa, USA, 2009.
- [40] S. L. Campbell, C. D. Meyer, and N. J. Rose, "Applications of the Drazin inverse to linear systems of differential equations with singular constant coefficients," *SIAM Journal on Applied Mathematics*, vol. 31, no. 3, pp. 411–425, 1976.
- [41] C. Bu, K. Zhang, and J. Zhao, "Representations of the Drazin inverse on solution of a class singular differential equations," *Linear and Multilinear Algebra*, vol. 59, no. 8, pp. 863–877, 2011.
- [42] J. Zhou, C. Bu, and Y. Wei, "Some block matrices with signed Drazin inverses," *Linear Algebra and Its Applications*, vol. 437, no. 7, pp. 1779–1792, 2012.
- [43] C. Bu, C. Feng, and S. Bai, "Representations for the Drazin inverses of the sum of two matrices and some block matrices," *Applied Mathematics and Computation*, vol. 218, no. 20, pp. 10226–10237, 2012.
- [44] C. Bu, J. Zhao, and J. Tang, "Representation of the Drazin inverse for special block matrix," *Applied Mathematics and Computation*, vol. 217, no. 10, pp. 4935–4943, 2011.
- [45] J. Wang, D. Cheng, and X. Hu, "Consensus of multi-agent linear dynamic systems," *Asian Journal of Control*, vol. 10, no. 2, pp. 144–155, 2008.
- [46] C. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia, Pa, USA, 2000.
- [47] D. Cobb, "Controllability, observability, and duality in singular systems," *IEEE Transactions on Automatic Control*, vol. 29, no. 12, pp. 1076–1082, 1984.
- [48] G. R. Duan, H. H. Yu, A. G. Wu, and X. Zhang, *Analysis and Design of Descriptor Linear Systems*, vol. 23 of *Advances in Mechanics and Mathematics*, Springer, New York, NY, USA, 2010.

Research Article

A New Neumann Series Method for Solving a Family of Local Fractional Fredholm and Volterra Integral Equations

Xiao-Jing Ma,¹ H. M. Srivastava,² Dumitru Baleanu,^{3,4,5} and Xiao-Jun Yang⁶

¹ College of Electrical Engineering, Xinjiang University, Urumqi, Xinjiang 830046, China

² Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada V8W 3R4

³ Department of Mathematics and Computer Sciences, Faculty of Arts and Sciences, Cankaya University, 06530 Ankara, Turkey

⁴ Department of Chemical and Materials Engineering, Faculty of Engineering, King Abdulaziz University, P.O. Box 80204, Jeddah 21589, Saudi Arabia

⁵ Institute of Space Sciences, Magurele, 077125 Bucharest, Romania

⁶ Department of Mathematics and Mechanics, China University of Mining and Technology, Xuzhou Campus, Xuzhou, Jiangsu 221008, China

Correspondence should be addressed to Xiao-Jing Ma; jingcici@gmail.com

Received 30 May 2013; Accepted 13 June 2013

Academic Editor: J. A. Tenreiro Machado

Copyright © 2013 Xiao-Jing Ma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose a new Neumann series method to solve a family of local fractional Fredholm and Volterra integral equations. The integral operator, which is used in our investigation, is of the local fractional integral operator type. Two illustrative examples show the accuracy and the reliability of the obtained results.

1. Introduction

Many initial- and boundary-value problems associated with ordinary differential equations (ODEs) and partial differential equations (PDEs) can be transformed into problems of solving the corresponding approximate integral equations. However, some initial- and boundary-value domains are fractal curves, which are everywhere continuous, but nowhere differentiable. As a result, we cannot employ the classical calculus, which requires that the defined functions should be differentiable, in order to process various classes of ordinary differential equations (ODEs) and partial differential equations (PDEs). Applications of fractional calculus, in general, and fractional differential equations [1–10], in particular, as well as various transport phenomena in complex and disordered media and fractional systems, have attracted considerable attention during the past two decades or so [11–22].

Recently, local fractional calculus [23–40], processing local fractional continuous non-differential functions, was successfully applied to model the stress-strain relation in

fractal elasticity [26, 27], fractal release equation [32], wave equations on Cantor sets [34], fractal heat equation [34], diffusion equation arising in discontinuous heat transfer in fractal media [35], Laplace equation within local fractional operators [36], Schrödinger equation in fractal time-space [37], damped wave equation and dissipative wave equation in fractal strings [38], heat-conduction equation on Cantor sets without heat generation in fractal media [39], and so on. There are some analytical and numerical methods for solving local fractional ODEs and PDEs, such as fractional complex transform method with local fractional operator [35], local fractional variational iteration method [37], Cantor-type cylindrical-coordinate method [38], local fractional Fourier series method [39], local fractional series expansion method [40], Fourier and Laplace transforms with local fractional operator [39], and reference therein.

The Neumann series method was applied to solve the integral equations [41, 42]. Recently, the fractional Neumann series method was considered in [43, 44]. This paper focuses on a new Neumann series method for solving the local fractional Fredholm and Volterra integral equation being

here facts in mind. This paper is structured as follows. Section 2 introduces the notations and the basic concepts. Section 3 is devoted to a new Neumann series method via local fractional integral operator. Two illustrative examples are explained in Section 4. Finally, conclusions are reported in Section 5.

2. Preliminaries

In order to investigate the local fractional continuity of non-differential functions, we suggest the result derived from fractal geometry [34, 39].

Let $f(x)$ be local fractional continuous on interval (a, b) ; then we write [34, 35]

$$f(x) \in C_\alpha(a, b). \quad (1)$$

If $f : (F, d) \rightarrow (\Omega', d')$ is a bi-Lipschitz mapping, then

$$\rho^s H^s(F) \leq H^s(f(F)) \leq \tau^s H^s(F), \quad (2)$$

which leads to

$$\rho^\alpha |x_1 - x_2|^\alpha \leq |f(x_1) - f(x_2)| \leq \tau^\alpha |x_1 - x_2|^\alpha, \quad (3)$$

so that

$$|f(x_1) - f(x_2)| < \varepsilon^\alpha, \quad (4)$$

where $\rho, \tau > 0$ and $x_1, x_2 \in F$.

The result deduced from fractal geometry is related to fractal coarse-grained mass function $\gamma^\alpha[F, a, b]$, which reads [34] as

$$\gamma^\alpha[F, a, b] = \frac{H^\alpha(F \cap (a, b))}{\Gamma(1 + \alpha)}, \quad (5)$$

with

$$H^\alpha(F \cap (a, b)) = (b - a)^\alpha, \quad (6)$$

where H^α is an α -dimensional Hausdorff measure.

Notice that we consider that the dimensions of any fractal spaces (e.g., Cantor spaces or the Cantor-like spaces) are positive numbers. It looks like the Euclidean space because its dimension is also positive number. The detailed results were considered in [34].

For $f(x) \in C_\alpha(a, b)$, local fractional integral of $f(x)$ of order α in the interval $[a, b]$ is given by [34, 37, 39]

$$\begin{aligned} {}_a I_b^{(\alpha)} f(x) &= \frac{1}{\Gamma(1 + \alpha)} \int_a^b f(t) (dt)^\alpha \\ &= \frac{1}{\Gamma(1 + \alpha)} \lim_{\Delta t \rightarrow 0} \sum_{j=0}^{j=N-1} f(t_j) (\Delta t_j)^\alpha, \end{aligned} \quad (7)$$

where $\Delta t_j = t_{j+1} - t_j$, $\Delta t = \max\{\Delta t_1, \Delta t_2, \Delta t_3, \dots\}$ and $[t_j, t_{j+1}], j = 0, \dots, N - 1$, $t_0 = a$, $t_N = b$, is a partition of the interval $[a, b]$.

For any $x \in (a, b)$, we have [34]

$${}_a I_x^{(\alpha)} f(x), \quad (8)$$

denoted by

$$f(x) \in I_x^{(\alpha)}(a, b). \quad (9)$$

If $f(x) \in I_x^{(\alpha)}(a, b)$, then we have [34]

$$f(x) \in C_\alpha(a, b). \quad (10)$$

For detailed content of fractal geometrical explanation of local fractional integral, we can see [34, 35]. Some properties of local fractional integral operator were suggested in (A.1)–(A.5).

3. A New Neumann Series Method to Deal with the Local Fractional Fredholm and Volterra Integral Equations

In this section, we consider a new Neumann series method to process the local fractional Fredholm and Volterra integral equations.

A new Neumann series method to deal with the local fractional Fredholm integral equation is written in the following form:

$$u(x) = f(x) + \frac{\lambda^\alpha}{\Gamma(1 + \alpha)} \int_a^b K(x, t) u(t) (dt)^\alpha. \quad (11)$$

It is obtained if we set

$$u_0(x) = f(x), \quad (12)$$

such that

$$\begin{aligned} u_1(x) &= u_0(x) + \frac{\lambda^\alpha}{\Gamma(1 + \alpha)} \int_a^b K(x, t) u_0(t) (dt)^\alpha \\ &= f(x) + \lambda^\alpha \psi_1(x), \end{aligned} \quad (13)$$

where $\psi_1(x) = (1/\Gamma(1 + \alpha)) \int_a^b K(x, t) f(t) (dt)^\alpha$.

The zeroth approximation can be written as

$$\begin{aligned} u_2(x) &= f(x) + \frac{\lambda^\alpha}{\Gamma(1 + \alpha)} \int_a^b K(x, t) u_1(t) (dt)^\alpha \\ &= f(x) + \frac{\lambda^\alpha}{\Gamma(1 + \alpha)} \\ &\quad \times \int_a^b K(x, t) \{f(t) + \lambda^\alpha \psi_1(t)\} (dt)^\alpha \\ &= f(x) + \lambda^\alpha \psi_1(x) + \lambda^{2\alpha} \psi_2(x), \end{aligned} \quad (14)$$

where $\psi_2(x) = (1/\Gamma(1 + \alpha)) \int_a^b K(x, t) \psi_1(t) (dt)^\alpha$.

Proceeding in this manner, the final solution $u(x)$ can be obtained as

$$\begin{aligned} u(x) &= f(x) + \lambda^\alpha \psi_1(x) + \lambda^{2\alpha} \psi_2(x) + \dots + \lambda^{n\alpha} \psi_n(x) + \dots \\ &= f(x) + \sum_{n=1}^{\infty} \lambda^{n\alpha} \psi_n(x), \end{aligned} \quad (15)$$

where $\psi_n(x) = (1/\Gamma(1 + \alpha)) \int_a^b K(x, t) \psi_{n-1}(t) (dt)^\alpha$, $n \geq 1$.

Now we structure a new Neumann series method to handle the local fractional Volterra integral equation, which reads as

$$u(x) = f(x) + \frac{\lambda^\alpha}{\Gamma(1+\alpha)} \int_a^x K(x,t) u(t) (dt)^\alpha. \quad (16)$$

The method is applicable provided that $u(x)$ is a local fractional analysis function; that is, $u(x)$ have a local fractional Taylor's expansion around $x = 0$.

$u(x)$ can be expressed by a local fractional series expansion; which reads as

$$u(x) = \sum_{n=1}^{\infty} a_n x^{n\alpha}, \quad (17)$$

where the coefficients a_n and x are constants that are required to be determined.

We have

$$\sum_{n=1}^{\infty} a_n x^{n\alpha} = f(x) + \frac{\lambda^\alpha}{\Gamma(1+\alpha)} \int_a^x K(x,t) \sum_{n=1}^{\infty} a_n x^{n\alpha} (dt)^\alpha. \quad (18)$$

Thus, using a few terms of the expansion in both sides, we find that

$$\begin{aligned} a_0 + a_1 x^\alpha + a_2 x^{2\alpha} + \cdots + a_n x^{n\alpha} + \cdots \\ = f(x) + \frac{\lambda^\alpha}{\Gamma(1+\alpha)} \int_a^x K(x,t) a_0 (dt)^\alpha \\ + \frac{\lambda^\alpha}{\Gamma(1+\alpha)} \int_a^x K(x,t) a_1 x^\alpha (dt)^\alpha \\ + \frac{\lambda^\alpha}{\Gamma(1+\alpha)} \int_a^x K(x,t) a_2 x^{2\alpha} (dt)^\alpha + \cdots \\ + \frac{\lambda^\alpha}{\Gamma(1+\alpha)} \int_a^x K(x,t) a_n x^{n\alpha} (dt)^\alpha + \cdots. \end{aligned} \quad (19)$$

We then write the local fractional Taylor's expansions for $f(x)$ and count the first few integrals in (19). After the integration is performed, we equate the coefficients of the same powers of x^α in both sides of (19). By this way, we can determine completely the unknown coefficients and produce solution in a local fractional series form.

4. Examples

Example 1. Solve the following local fractional Fredholm integral equation:

$$u(x) = \Gamma(1+\alpha) + \frac{1}{\Gamma(1+\alpha)} \int_0^1 x^\alpha u(t) (dt)^\alpha. \quad (20)$$

Let us consider the zeroth approximation given by

$$u_0(t) = \Gamma(1+\alpha). \quad (21)$$

The first approximation can be computed as follows:

$$\begin{aligned} u_1(x) &= \Gamma(1+\alpha) + \frac{1}{\Gamma(1+\alpha)} \int_0^1 x^\alpha \Gamma(1+\alpha) (dt)^\alpha \\ &= \Gamma(1+\alpha) + x^\alpha. \end{aligned} \quad (22)$$

Proceeding in this manner, we find the following local fractional series approximation:

$$\begin{aligned} u_2(x) &= \Gamma(1+\alpha) + \frac{1}{\Gamma(1+\alpha)} \int_0^1 x^\alpha (\Gamma(1+\alpha) + t^\alpha) (dt)^\alpha \\ &= \Gamma(1+\alpha) + x^\alpha \left(1 + \frac{\Gamma(1+\alpha)}{\Gamma(1+2\alpha)} \right). \end{aligned} \quad (23)$$

Similarly, the third approximation reads as follows:

$$\begin{aligned} u_3(x) &= \Gamma(1+\alpha) + \frac{1}{\Gamma(1+\alpha)} \\ &\times \int_0^1 x^\alpha \left(\Gamma(1+\alpha) + t^\alpha \left(1 + \frac{\Gamma(1+\alpha)}{\Gamma(1+2\alpha)} \right) \right) (dt)^\alpha \\ &= \Gamma(1+\alpha) + x^\alpha \left(1 + \frac{\Gamma(1+\alpha)}{\Gamma(1+2\alpha)} + \frac{\Gamma^2(1+\alpha)}{\Gamma^2(1+2\alpha)} \right). \end{aligned} \quad (24)$$

The fourth approximation yields

$$\begin{aligned} u_4(x) &= \Gamma(1+\alpha) + \frac{1}{\Gamma(1+\alpha)} \\ &\times \int_0^1 x^\alpha \left(\Gamma(1+\alpha) + t^\alpha \right. \\ &\quad \left. \times \left(1 + \frac{\Gamma(1+\alpha)}{\Gamma(1+2\alpha)} + \frac{\Gamma^2(1+\alpha)}{\Gamma^2(1+2\alpha)} \right) \right) (dt)^\alpha \\ &= \Gamma(1+\alpha) + x^\alpha \left(1 + \frac{\Gamma(1+\alpha)}{\Gamma(1+2\alpha)} \right. \\ &\quad \left. + \frac{\Gamma^2(1+\alpha)}{\Gamma^2(1+2\alpha)} + \frac{\Gamma^3(1+\alpha)}{\Gamma^3(1+2\alpha)} \right). \end{aligned} \quad (25)$$

In conclusion, we get

$$u_n(x) = \Gamma(1+\alpha) + x^\alpha \sum_{i=0}^n \left(\frac{\Gamma(1+\alpha)}{\Gamma(1+2\alpha)} \right)^n. \quad (26)$$

Hence,

$$\begin{aligned} u(x) &= \lim_{n \rightarrow \infty} u_n(x) \\ &= \Gamma(1+\alpha) + x^\alpha \lim_{n \rightarrow \infty} \sum_{i=0}^n \left(\frac{\Gamma(1+\alpha)}{\Gamma(1+2\alpha)} \right)^n \\ &= \Gamma(1+\alpha) + \frac{x^\alpha \Gamma(1+\alpha)}{\Gamma(1+2\alpha) - \Gamma(1+\alpha)}. \end{aligned} \quad (27)$$

Example 2. Obtain the solution of the following local fractional Volterra equation:

$$u(x) = 1 + \frac{x^\alpha}{\Gamma(1+\alpha)} + \frac{1}{\Gamma(1+\alpha)} \int_0^x \frac{(t-x)^\alpha}{\Gamma(1+\alpha)} u(t) (dt)^\alpha. \quad (28)$$

Suppose that there exists the solution in the following local fractional series form:

$$u(x) = \sum_{n=1}^{\infty} a_n x^{n\alpha}. \quad (29)$$

Then, upon substituting the local fractional series into the equation, we find that

$$\begin{aligned} & \sum_{n=1}^{\infty} a_n x^{n\alpha} \\ &= \frac{x^\alpha}{\Gamma(1+\alpha)} + \frac{1}{\Gamma(1+\alpha)} \int_0^x \frac{(t-x)^\alpha}{\Gamma(1+\alpha)} \sum_{n=1}^{\infty} a_n x^{n\alpha} (dt)^\alpha \\ &= \frac{x^\alpha}{\Gamma(1+\alpha)} - \sum_{n=1}^{\infty} \frac{\Gamma(n\alpha+1) a_n x^{(n+2)\alpha}}{\Gamma((n+2)\alpha+1)}. \end{aligned} \quad (30)$$

Comparing the coefficients of the same powers of x^α , we get

$$\begin{aligned} a_0 &= 1, & a_1 &= \frac{1}{\Gamma(1+\alpha)}, \\ a_2 &= -\frac{a_0}{\Gamma(2\alpha+1)}, & a_3 &= -\frac{\Gamma(\alpha+1)a_1}{\Gamma(3\alpha+1)}, \\ &\vdots \\ a_n &= -\frac{\Gamma((n-2)\alpha+1)a_{n-2}}{\Gamma(n\alpha+1)}, \end{aligned} \quad (31)$$

and so on. Thus, the values of the coefficients can be calculated as follows:

$$\begin{aligned} a_0 &= 1, & a_1 &= \frac{1}{\Gamma(1+\alpha)}, \\ a_2 &= -\frac{1}{\Gamma(2\alpha+1)}, & a_3 &= -\frac{1}{\Gamma(3\alpha+1)}, \\ a_4 &= \frac{1}{\Gamma(4\alpha+1)}, & a_5 &= \frac{1}{\Gamma(5\alpha+1)}, \\ a_6 &= -\frac{1}{\Gamma(6\alpha+1)}, & a_7 &= -\frac{1}{\Gamma(7\alpha+1)}, \\ &\vdots \end{aligned} \quad (32)$$

Hence, the local fractional series solution is given by

$$\begin{aligned} u(x) &= \sum_{n=1}^{\infty} a_n x^{n\alpha} \\ &= \left(1 - \frac{x^{3\alpha}}{\Gamma(3\alpha+1)} + \frac{x^{5\alpha}}{\Gamma(5\alpha+1)} - \frac{x^{7\alpha}}{\Gamma(7\alpha+1)} + \dots \right) \\ &\quad + \left(\frac{x^\alpha}{\Gamma(1+\alpha)} - \frac{x^{2\alpha}}{\Gamma(2\alpha+1)} + \frac{x^{4\alpha}}{\Gamma(4\alpha+1)} \right. \\ &\quad \left. - \frac{x^{6\alpha}}{\Gamma(6\alpha+1)} + \dots \right) \\ &= \cos_\alpha x^\alpha + \sin_\alpha x^\alpha, \end{aligned} \quad (33)$$

which are satisfied with the condition given by [34, 39]

$$E_\alpha(i^\alpha x^\alpha) = \cos_\alpha x^\alpha + i^\alpha \sin_\alpha x^\alpha, \quad (34)$$

where the Mittag-Leffler function defined on fractal set of fractal dimension α is suggested by [34, 39]

$$E_\alpha(x^\alpha) = \sum_{k=0}^{\infty} \frac{x^{\alpha k}}{\Gamma(1+k\alpha)}. \quad (35)$$

5. Conclusions

Local fractional differential and integral operators have proven to be useful tools to deal with everywhere continuous (but nowhere differentiable) functions in fractal areas ranging from fundamental science to engineering. In this paper, it is proven that a new Neumann series method can be used for solving the local fractional Fredholm and Volterra integral equations, and their solutions are fractal functions. The proposed method is efficient and leads to accurate, approximately convergent solutions to local fractional Fredholm and Volterra integral equations. It is demonstrated that the solutions of local fractional Fredholm and Volterra integral equations are fractal functions, which are equipped with local fractional continuities. However, the classical and fractional Neumann series methods [41–44] were only applied to continuous functions.

Appendix

The following properties of local fractional integral operator are valid [34].

- (a) For any $f(x) \in C_\alpha(a, b)$, $0 < \alpha \leq 1$, we have local fractional multiple integrals, which are written as [34]

$${}_{x_0} I_x^{(k\alpha)} f(x) = \overbrace{{}_{x_0} I_x^{(\alpha)} \cdots {}_{x_0} I_x^{(\alpha)}}^{k \text{ times}} f(x). \quad (\text{A.1})$$

- (b) If $\psi(x, y) \in C_\alpha(a, b) \times C_\alpha(c, d)$, then [34]

$${}_a I_b^{(\alpha)} {}_c I_b^{(\alpha)} \psi(x, y) = {}_c I_d^{(\alpha)} {}_a I_b^{(\alpha)} \psi(x, y). \quad (\text{A.2})$$

(c) The sine and cosine subfunctions can, respectively, be written as follows [34, 39]:

$$\begin{aligned}\sin_{\alpha} x^{\alpha} &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{\alpha(2k+1)}}{\Gamma[1 + \alpha(2k+1)]}, \\ \cos_{\alpha} x^{\alpha} &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2\alpha k}}{\Gamma(1 + 2\alpha k)}, \quad 0 < \alpha \leq 1.\end{aligned}\quad (\text{A.3})$$

(d) Suppose that $f(t)$ is local fractional continuous on the interval $[a, b]$. Then

$${}_a I_x^{(\alpha)} {}_a I_{\tau}^{(\alpha)} f(t) = {}_a I_x^{(\alpha)} \frac{(x-t)^{\alpha} f(t)}{\Gamma(1+\alpha)} \quad (x \in [a, b]). \quad (\text{A.4})$$

(e) We have

$${}_0 I_x^{(\alpha)} {}_0 I_{\tau}^{(\alpha)} \frac{t^{k\alpha}}{\Gamma(k\alpha+1)} = \frac{t^{(k+2)\alpha}}{\Gamma((k+2)\alpha+1)}. \quad (\text{A.5})$$

References

- [1] K. B. Oldham and J. Spanier, *The Fractional Calculus*, Academic Press, London, UK, 1974.
- [2] K. S. Miller and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*, John Wiley & Sons Inc., New York, NY, USA, 1993.
- [3] S. G. Samko, A. A. Kilbas, and O. I. Marichev, *Fractional Integrals and Derivatives*, Gordon and Breach Science Publishers, Yverdon, Switzerland, 1993.
- [4] V. Kiryakova, *Generalized Fractional Calculus and Applications*, vol. 301, Longman Scientific & Technical, Harlow, UK, 1994.
- [5] I. Podlubny, *Fractional Differential Equations*, vol. 198, Academic Press, San Diego, Calif, USA, 1999.
- [6] R. Hilfer, Ed., *Applications of Fractional Calculus in Physics*, World Scientific Publishing, River Edge, NJ, USA, 2000.
- [7] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, *Theory and Applications of Fractional Differential Equations*, vol. 204, Elsevier Science B.V., Amsterdam, The Netherlands, 2006.
- [8] J. Sabatier, O. P. Agrawal, and J. A. Tenreiro Machado, *Advances in Fractional Calculus: Theoretical Developments and Applications in Physics and Engineering*, Springer, New York, NY, USA, 2007.
- [9] F. Mainardi, *Fractional Calculus and Waves in Linear Viscoelasticity*, Imperial College Press, London, UK, 2010.
- [10] G. M. Zaslavsky, “Chaos, fractional kinetics, and anomalous transport,” *Physics Reports*, vol. 371, no. 6, pp. 461–580, 2002.
- [11] J. Hristov, “Heat-balance integral to fractional (half-time) heat diffusion sub-model,” *Thermal Science*, vol. 14, no. 2, pp. 291–316, 2010.
- [12] B. J. West, M. Bologna, and P. Grigolini, *Physics of Fractal Operators*, Springer, New York, NY, USA, 2003.
- [13] R. Metzler and J. Klafter, “The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics,” *Journal of Physics A*, vol. 37, no. 31, pp. R161–R208, 2004.
- [14] L. M. Zelený and A. V. Milovanov, “Fractal topology and strange kinetics: from percolation theory to problems in cosmic electrodynamics,” *Physics-Uspekhi*, vol. 47, no. 8, pp. 749–788, 2004.
- [15] G. M. Zaslavsky, *Hamiltonian Chaos and Fractional Dynamics*, Oxford University Press, Oxford, UK, 2008.
- [16] J. A. T. Machado, “Analysis and design of fractional-order digital control systems,” *Systems Analysis Modelling Simulation*, vol. 27, no. 2-3, pp. 107–122, 1997.
- [17] J. A. T. Machado, “Fractional-order derivative approximations in discrete-time control systems,” *Systems Analysis Modelling Simulation*, vol. 34, no. 4, pp. 419–434, 1999.
- [18] R. Herrmann, *Fractional Calculus: An Introduction for Physicists*, World Scientific, 2011.
- [19] J. Klafter, S. C. Lim, and R. Metzler, *Fractional Dynamics: Recent Advances*, World Scientific, 2012.
- [20] M. D. Ortigueira, *Fractional Calculus for Scientists and Engineers*, Springer, 2011.
- [21] D. Baleanu, K. Diethelm, E. Scalas, and J. J. Trujillo, *Fractional Calculus Models and Numerical Methods*, Complexity, Nonlinearity and Chaos, World Scientific, Boston, Mass, USA, 2012.
- [22] I. Petras, *Fractional-Order Nonlinear Systems: Modeling, Analysis and Simulation*, Springer, Berlin, Germany, 2011.
- [23] K. M. Kolwankar and A. D. Gangal, “Fractional differentiability of nowhere differentiable functions and dimensions,” *Chaos*, vol. 6, no. 4, pp. 505–513, 1996.
- [24] K. M. Kolwankar and A. D. Gangal, “Local fractional Fokker-Planck equation,” *Physical Review Letters*, vol. 80, pp. 214–217, 1998.
- [25] F. B. Adda and J. Cresson, “About non-differentiable functions,” *Journal of Mathematical Analysis and Applications*, vol. 263, no. 2, pp. 721–737, 2001.
- [26] A. Carpinteri, B. Chiaia, and P. Cornetti, “Static-kinematic duality and the principle of virtual work in the mechanics of fractal media,” *Computer Methods in Applied Mechanics and Engineering*, vol. 191, no. 1-2, pp. 3–19, 2001.
- [27] A. Carpinteri, B. Chiaia, and P. Cornetti, “The elastic problem for fractal media: basic theory and finite element formulation,” *Computers and Structures*, vol. 82, no. 6, pp. 499–508, 2004.
- [28] Y. Chen, Y. Yan, and K. Zhang, “On the local fractional derivative,” *Journal of Mathematical Analysis and Applications*, vol. 362, no. 1, pp. 17–33, 2010.
- [29] A. Babakhani and V. D. Gejji, “On calculus of local fractional derivatives,” *Journal of Mathematical Analysis and Applications*, vol. 270, no. 1, pp. 66–79, 2002.
- [30] G. Jumarie, “Table of some basic fractional calculus formulae derived from a modified Riemann-Liouville derivative for non-differentiable functions,” *Applied Mathematics Letters*, vol. 22, no. 3, pp. 378–385, 2009.
- [31] G. Jumarie, “Probability calculus of fractional order and fractional Taylor’s series application to Fokker-Planck equation and information of non-random functions,” *Chaos, Solitons and Fractals*, vol. 40, no. 3, pp. 1428–1448, 2009.
- [32] W. Chen, H. Sun, X. Zhang, and D. Korošak, “Anomalous diffusion modeling by fractal and fractional derivatives,” *Computers and Mathematics with Applications*, vol. 59, no. 5, pp. 1754–1758, 2010.
- [33] W. Chen, “Time-space fabric underlying anomalous diffusion,” *Chaos, Solitons and Fractals*, vol. 28, no. 4, pp. 923–925, 2006.
- [34] X.-J. Yang, *Advanced Local Fractional Calculus and Its Applications*, World Science, New York, NY, USA, 2012.
- [35] M.-S. Hu, D. Baleanu, and X.-J. Yang, “One-phase problems for discontinuous heat transfer in fractal media,” *Mathematical Problems in Engineering*, vol. 2013, Article ID 358473, 3 pages, 2013.

- [36] A. Liangprom and K. Nonlaopon, "On the convolution equation related to the diamond Klein-Gordon operator," *Abstract and Applied Analysis*, vol. 2011, Article ID 908491, 16 pages, 2011.
- [37] W.-H. Su, D. Baleanu, X.-J. Yang, and H. Jafari, "Damped wave equation and dissipative wave equation in fractal strings within the local fractional variational iteration method," *Fixed Point Theory and Applications*, vol. 2013, no. 1, pp. 1–11, 2013.
- [38] X.-J. Yang, H. M. Srivastava, J.-H. He, and D. Baleanu, "Cantor-type cylindrical-coordinate method for differential equations with local fractional derivatives," *Physics Letters A*, vol. 377, no. 38–30, pp. 1696–1700.
- [39] X.-J. Yang, *Local Fractional Functional Analysis and Its Applications*, Asian Academic, Hong Kong, China, 2011.
- [40] A.-M. Yang, X.-J. Yang, and Z.-B. Li, "Local fractional series expansion method for solving wave and diffusion equations on Cantor sets," *Abstract and Applied Analysis*, vol. 2013, Article ID 351057, 5 pages, 2013.
- [41] D. Medková, "On the convergence of Neumann series for noncompact operators," *Czechoslovak Mathematical Journal*, vol. 41, no. 2, pp. 312–316, 1991.
- [42] P. Stefanov and G. Uhlmann, "Thermoacoustic tomography with variable sound speed," *Inverse Problems*, vol. 25, no. 7, Article ID 075011, 2009.
- [43] H. M. Srivastava and R. K. Saxena, "Operators of fractional integration and their applications," *Applied Mathematics and Computation*, vol. 118, no. 1, pp. 1–52, 2001.
- [44] A. Ebaid, D. M. M. ElSayed, and M. D. Aljoufi, "Fractional calculus model for damped Mathieu equation: approximate analytical solution," *Applied Mathematical Sciences*, vol. 6, no. 81–84, pp. 4075–4080, 2012.

Research Article

Multidimensional Scaling for Orthodontic Root Resorption

**Cristina Teodora Preoteasa,¹ Ecaterina Ionescu,¹ Elena Preoteasa,¹
J. A. Tenreiro Machado,² Mihaela Cristina Baleanu,³ and Dumitru Baleanu^{4,5,6}**

¹ Faculty of Dental Medicine, University of Medicine and Pharmacy Carol Davila, 010232 Bucharest, Romania

² Institute of Engineering, Polytechnic of Porto, Department of Electrical Engineering, 431 Dr. Antonio Bernardino de Almeida Street, 4200-072 Porto, Portugal

³ Mihail Sadoveanu Theoretical High School, District 2, Popa Lazar Street, No. 8, 021586 Bucharest, Romania

⁴ Department of Mathematics and Computer Sciences, Faculty of Arts and Sciences, Cankaya University, Balgat, 06530 Ankara, Turkey

⁵ Department of Chemical and Materials Engineering, Faculty of Engineering, King Abdulaziz University, P.O. Box 80204, Jeddah 21589, Saudi Arabia

⁶ Institute of Space Sciences, P.O. Box MG-23, Magurele, 76900 Bucharest, Romania

Correspondence should be addressed to J. A. Tenreiro Machado; jtm@isep.ipp.pt

Received 26 March 2013; Accepted 23 May 2013

Academic Editor: Bashir Ahmad

Copyright © 2013 Cristina Teodora Preoteasa et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The paper investigates the risk factors for the severity of orthodontic root resorption. The multidimensional scaling (MDS) visualization method is used to investigate the experimental data from patients who received orthodontic treatment at the Department of Orthodontics and Dentofacial Orthopedics, Faculty of Dentistry, “Carol Davila” University of Medicine and Pharmacy, during a period of 4 years. The clusters emerging in the MDS plots reveal features and properties not easily captured by classical statistical tools. The results support the adoption of MDS for tackling the dentistry information and overcoming noise embedded into the data. The method introduced in this paper is rapid, efficient, and very useful for treating the risk factors for the severity of orthodontic root resorption.

1. Introduction

Root resorption is defined as the biological process characterized by destruction of hard structure of the tooth root. Damaging may involve cementum, dentin, or both structures. Orthodontic treatment is associated with a higher frequency and severity of the pathological process of external root resorption. Frequency of orthodontic root resorption is about 100% when diagnostic techniques based on microscopy are used and around 70% when periapical or panoramic radiographs are used [1]. But generally, orthodontic root resorption severity is moderate and low, not interfering with the positive outcomes of this particular medical intervention. The meta-analysis conducted by Segal [2] identified a mean value of root shortening after orthodontic treatment of 1.421 ± 0.448 mm. Of interest are those cases where severe root resorption with root shortening is beyond 4 mm or 1/3 of root length, which

are noticed in 1–5% of orthodontic patients [3]. Due to the fact that the considerable root shortening has a negative impact on tooth's prognosis, orthodontic root resorption is nowadays one of the most discussed complications of the orthodontic treatment: efforts are made in order to establish the proper preventive treatment conduct [4].

Etiopathogeny of orthodontic root resorption presents several uncertainties. It seems that higher incidence and severity of orthodontic root resorption are related mainly to patients' characteristics (individual susceptibility has the main role in root resorption appearance) and particularities of the orthodontic treatment applied [5]. It seems that a higher risk of developing external root resorption presents patients with allergies, asthma, diabetes, arthritis, endocrine disorders, Paget's disease, tooth eruption disorders (more frequently caused by pressure of the canine or wisdom teeth during eruption), hypodontia, open bite, increased alveolar

bone density, particular dental morphology like reduced root diameter, and abnormal root shape in the apical part of the root, especially eroded, pointed, deviated, or bottle shape [3, 5–9]. Among particularities of the orthodontic technique seen as risk factors for root resorption are increased orthodontic treatment duration, treatment with tooth extraction, and tooth intrusion (especially when vestibular coronal torque was associated) [10, 11].

The aim of this study is to investigate the impact of several risk factors (sex of patients, orthodontic extractions, and duration of treatment) on the severity of orthodontic root resorption. The analysis is based on the formation of clusters for experimental data analysis that may indicate similar behavior in some particular clinical situations.

Given the characteristics of the biomedical data, with a plethora of different influential factors, the numerical extraction of characteristics poses difficulties to classical statistical and computer tools. In this line of thought, the adoption of advanced computational tools capable of handling the incertitude implicit in the application is imperative. Therefore, the multidimensional scaling (MDS) method which is an algorithm that does not require initial assumptions about the data is tested. MDS is a computational visualization tool that constructs maps based on comparison criteria [12–19]. The MDS plot consists of a series of points where each one represents an item. The maps can be rotated, translated, and zoomed in order to provide a better visualization of a given area [20–28]. The interpretation of the chart is based on relative position of clusters of points and takes advantage of the user intuition and experience in the particular field of application, making MDS a powerful tool for the type of data handled in this study.

Bearing these ideas in mind, the paper is organized as follows. In Section 2, the material and methods are presented. Section 3 is devoted to multidimensional scaling method for patients with and without orthodontic extractions. Section 4 presents our results and discussions. Finally, Section 5 outlines the main conclusions.

2. Materials and Methods

In order to achieve the proposed objectives, we designed and implemented a retrospective observational clinical study.

The sample was composed of patients receiving orthodontic treatment at the Department of Orthodontics and Dentofacial Orthopedics, Faculty of Dentistry, “Carol Davila” University of Medicine and Pharmacy, during October 2005–October 2009. In this study, patients with fixed metallic orthodontic appliances, standard edgewise, or straight-wire technique, applied in both jaws for a period of at least 6 months were included. From this study, patients with radiological signs of root resorption before the treatment start were excluded. According to the protocol established in the Department of Orthodontics and Dentofacial Orthopedics, all patients sign an informed consent for the use of their medical documents for teaching and scientific purposes.

External root resorption was assessed in terms of root shortening, with the changes of root length being recorded, compared to the situation before applying the orthodontic device. Measurements of upper and lower incisors were made

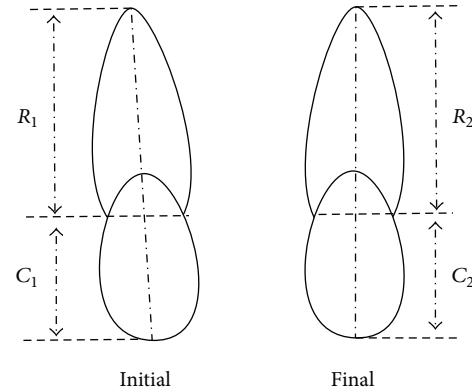


FIGURE 1: Calculation of the amount of root shortening (root resorption) on serial panoramic radiographs.

on serial panoramic radiographs, and changes in root length were being assessed using a mathematical formula based on the one proposed by L. Linge and B. O. Linge [29] (see Figure 1):

$$\text{Root resorption} = R_1 - R_2 \frac{C_1}{C_2}, \quad (1)$$

where R_1 and C_1 are the initial (pretreatment) root and crown lengths in the first radiograph and R_2 and C_2 are root and crown lengths in the second radiograph, respectively.

3. Experimental Setup and Mathematical Tools

In this section, the experimental cases and the mathematical tools to be adopted are briefly described.

3.1. Experiments. In the experiments, two cases were considered, namely, 1 (orthodontic treatment with tooth extraction) and 2 (orthodontic treatment without tooth extraction), denoted as cases 1 and 2 in the sequel, involving $p = 18$ patients for case 1 and $p = 37$ patients for case 2, respectively, which are studied separately. For all patients considered, $k_{\max} = 8$ measures the teeth absolute deviations that were compared by means of indices (3)–(4). In order to identify points in the MDS maps, each patient was labelled as GPP-AA-TT with $G = \{M, F\}$ for gender male/female, PP = $\{1, \dots, p\}$ for patient number, AA = $\{10, \dots, 30\}$ for age, in years, and TT = $\{6, \dots, 51\}$ for treatment time, in months. For providing a better visualization, the male and female points have distinct marks (filled circle for male and open rectangle for female).

3.2. Multidimensional Scaling. Multidimensional scaling (MDS) has its origins in psychometrics and psychophysics, where it is used as a tool for perceptual and cognitive modeling. From the beginning, MDS has been applied in many fields, such as psychology, sociology, anthropology, economy, and educational research. In the last decades, this

technique has been applied also in other areas such as music, finance, and biology.

MDS is a statistical technique used for visualization of information in the perspective of exploring similarities in data. MDS assigns a point to each item in an m -dimensional space and arranges the p objects in a space with a given number of dimensions m , in order to reproduce the observed similarities. Often, instead of similarities distances between the investigated p objects are considered. For two or three dimensions, the resulting locations may be displayed in a “map” that can be analyzed. We can rotate or translate the map, but the similarities between items remain the same. Therefore, the final orientation of axes in space is mostly the result of a subjective decision by the researcher, who will choose the one that can be most easily interpreted or that leads to a better visualization.

An MDS algorithm starts by defining a measure of similarity for constructing a $p \times p$ matrix \mathbf{R} of item-to-item similarities. In classical MDS, the matrix is symmetric and its main diagonal is composed of “1” for similarities or “0” for distances. MDS rearranges objects so as to arrive at a configuration that best approximates the observed similarities (or, alternatively, the measured distances). For this purpose, MDS uses a function minimization algorithm that evaluates different configurations with the goal of optimizing the goodness of fit.

The most common measure used to evaluate how well a particular configuration reproduces the observed distance matrix is the raw stress defined by

$$S = [d_{ij} - f(\delta_{ij})]^2, \quad i, j = 1, \dots, p, \quad (2)$$

where d_{ij} denotes the reproduced distances, given the respective number of dimensions and δ_{ij} stands for the input data (i.e., the observed distances). The expression $f(\delta_{ij})$ represents a nonmetric, monotone transformation of the input data.

There are several measures that are commonly used, but most of them amount to the computation of the sum of squared deviations of observed values from the reproduced distances. Thus, the smaller the stress value S , the better the fit between the reproduced and the observed distance matrices.

We can plot S versus the number of dimensions m of the visualization, for deciding the most adequate. Usually, we get a monotonic decreasing plot, and we chose the “best dimension” as a compromise between stress reduction and dimension for the map representation. In practical terms, we chose a low dimension at the region where we have a significant variation in the stress plot.

We can also plot the reproduced distances, for a particular number of dimensions, against the observed input data (distances). This scatter plot, referred to as a Shepard diagram, shows the distances between points versus the original dissimilarities. In the Shepard plot, a narrow scatter that is around a 45-degree line indicates a good fit of the distances to the dissimilarities, while a large scatter indicates a lack of fit.

Since MDS is fed with relative distances, the maps are insensitive to rotation and translation. This means that the user can view and zoom the plots interactively in order to

interpret the clusters of points that emerge in the map. Furthermore, distinct indices, capturing different characteristics, produce MDS charts, better or worse, merely in the viewpoint of easiness of interpretation in conjunction with the user own experience.

For the comparison of objects i and j , the cosine correlation [30], r_{ij} is adopted, defined as

$$r_{ij} = \frac{\sum_{k=1}^{k_{\max}} x_i(k) x_j(k)}{\sqrt{\sum_{k=1}^{k_{\max}} x_i^2(k) \cdot \sum_{k=1}^{k_{\max}} x_j^2(k)}}, \quad i, j = 1, \dots, p, \quad (3)$$

where $x_i(k)$ denotes the k th measured component of the i th item and k_{\max} represents the maximum number of components.

In the sequel will be adopted GGobi [31] for calculating MDS and for visualization. Other packages often used are Matlab [32] and R [33]. The GGobi package requires a $p \times p$ matrix $\Delta = [\delta_{ij}]$ of item-to-item distances (instead of similarities) and, therefore, is used:

$$\delta_{ij} = \frac{\pi}{2} - \cos^{-1}(r_{ij}). \quad (4)$$

In the experiments other metrics that lead to inferior results were tested and, therefore, are not analyzed here.

4. Results and Discussion

In this section, the sample of patients is described and the MDS results are analysed.

4.1. Description of the Sample. The sample included 55 patients, of which 74.5% ($n = 41$) were female and 25.5% ($n = 14$) were male, with the mean age being 15.92 years. Corresponding to the 55 patients, a total of 440 incisors were measured.

Associated with the orthodontic treatment, there was a mean reduction of 1.32 mm of tooth length. Most of the patients (67.27%, $n = 37$) had at least one incisor with root resorption exceeding 2 mm; this is often considered as the limit between low severity and medium severity of this pathological process.

4.2. Multidimensional Scaling of Patients with and without Orthodontic Extractions. Figures 2 and 3 depict the 3-dimensional MDS maps for cases 1 and 2, respectively. The figures show two rotations views and two zoomed areas. This paper examines the formation of clusters for experimental data analysis.

In case 1 (with orthodontic extractions) mean root resorption was 1.59 mm. We cannot say anything about the implications of sex on root resorption severity due to the insignificant number of women ($n = 1$). This aspect is equivalent to the fact that more frequently the treatment plan included orthodontic extractions in male patients. Subjects presented ages between 12–28 years (mean = 16 years). Orthodontic treatment time was between 10 and 51 months (mean = 24.5 months). Important cluster points were

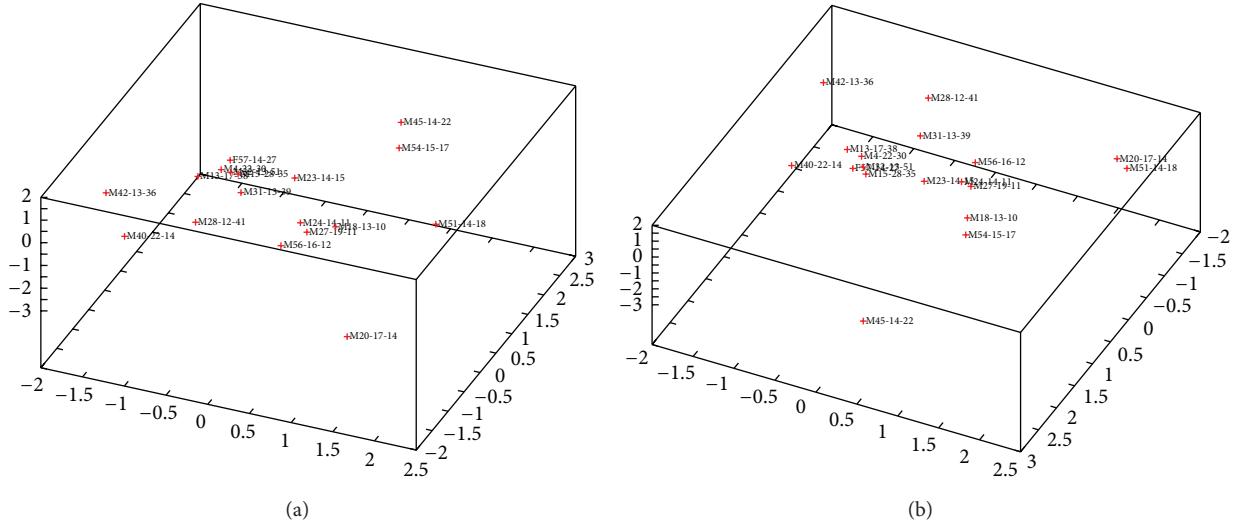


FIGURE 2: Projections of the 3-dimensional MDS map for case 1 (with tooth extraction) involving $p = 18$ patients.

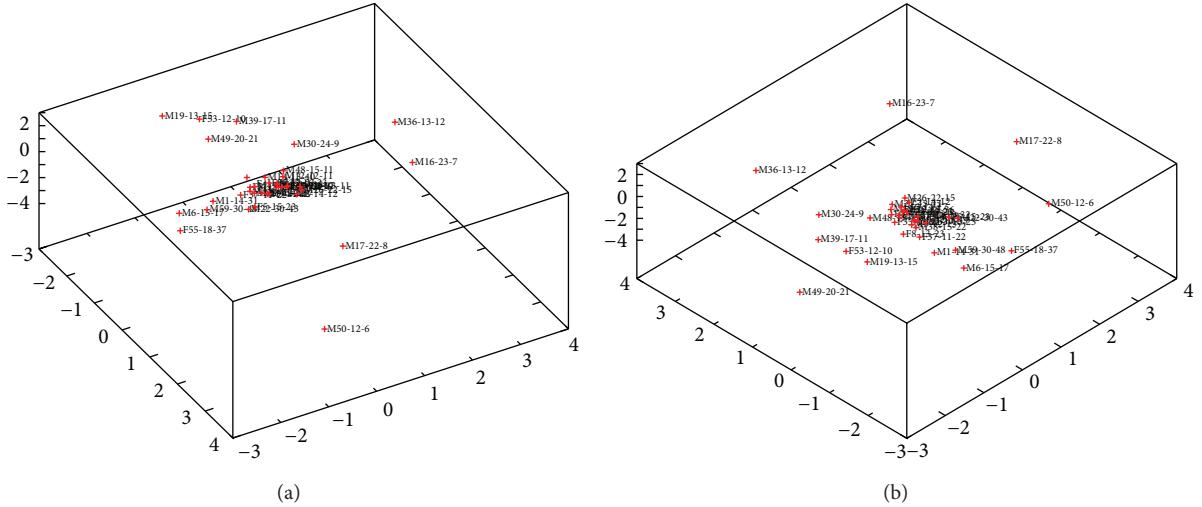


FIGURE 3: Projections of the 3-dimensional MDS map for case 2 (without teeth extraction) involving $p = 37$ patients.

observed, namely, M 13-17-38, M 4-22-30, M 15-28-35, M 32-13-51, and F 57-14-27. These patients presented moderate values of root shortening (mean = 2.26 mm), indicating a higher severity than usual. It can be observed that these patients had one particular aspect in common, the extent of treatment time, which is above average. In this cluster, age did not present some clear tendency, being variable (between 13 and 28 years), but it can be noticed that mean age for the patients with orthodontic extractions was higher in subjects from the cluster (mean = 19 years) than that for the others (mean = 15 years). In consequence, we observe a tendency to form clusters with higher severity of root resorption in cases with orthodontic extractions with a greater period of orthodontic treatment time.

In case 2 (without orthodontic extractions) 37 patients were included with mean root resorption being 1.18 mm, indicating a lower severity of this pathological process. In this

group, the number of males and females are comparable (13 males and 24 females). Patients presented ages between 10–30 years (mean = 16 years). Treatment time was between 6–48 months (mean = 20.21 months). We report the appearance of the clusters which shows some similarities (Figure 3). We observed a more uniform behavior compared to that of case 1. A large cluster was observed, including about 24 patients. We mention that generally patients inside the cluster presented a tendency to a more severe root resorption (mean = 1.40 mm), compared to those outside the cluster (mean = 0.77 mm). We observed also that outside the cluster, there were mostly male patients and only 2 female patients. That may suggest that females may present a tendency to a more uniform and predictive behavior related to orthodontic root resorption appearance. Also, even if there were patients of different ages outside the cluster, most of them were older than those inside the cluster. Also the mean age of patients outside the cluster

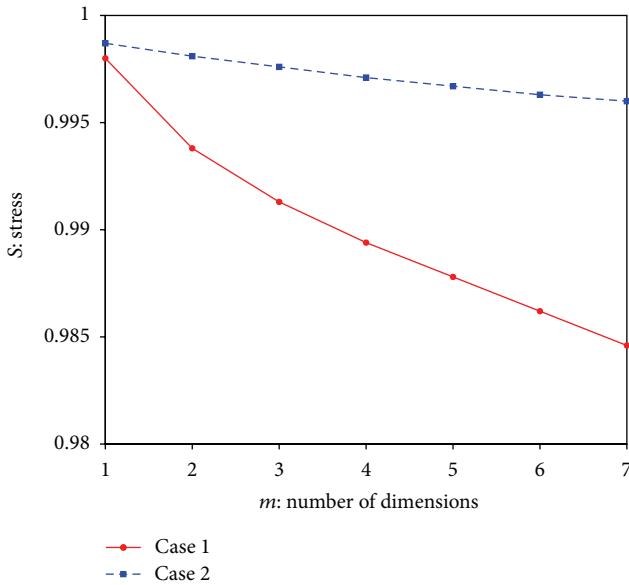


FIGURE 4: Stress versus the number of dimensions of the MDS representations for cases 1 and 2.

(18 years) was higher compared to that of the subjects inside the cluster (15 years). This may suggest that in older patients root resorption is a process that is less predictable.

Figure 4 shows the variation of the stress S versus the number of dimensions m of the MDS representations for cases 1 and 2. We observe that the adopted dimension $m = 3$ establishes a good compromise between precision and dimensionality reduction for visualization feasibility. Comparing the 2 situations mentioned (with and without orthodontic extractions), we observed that stress presents a higher variability in cases with extractions and a more uniform behavior in those cases without extractions. That may be related to the increased complexity of the medical intervention in the first case.

Orthodontic root resorption is one of the complications of the orthodontic treatment, knowing its etiology being an important factor in prevention of those forms with moderate and severe root shortening.

According to our results, the orthodontic extraction may be a risk for apical root resorption. More frequently it is associated with a more severe root resorption and also a more unpredictable behavior. This aspect is concordant with several clinical studies conducted on this topic [34]. Marques et al. reported that those receiving orthodontic treatment with extraction of the four first premolars, compared to patients treated without extractions, have a chance of 6.7 : 1 to present root resorption [35]. de Freitas et al. also identify a statistically significant difference in the severity of root resorption between cases with and without orthodontic extractions, with root shortening being more severe in the first case [36]. Mohandesan et al. support the same behavior that differentiated between cases with and without orthodontic extractions [37]. These observations may be explained by the larger root apex displacement in cases with orthodontic extractions and

also by the increase of treatment time, with both favouring this condition. We mention that increased treatment time of the orthodontic intervention is seen, in accordance with the current knowledge, as the main risk factor of root resorption related to the orthodontic intervention [38].

5. Conclusions

Nowadays, preventive methods are seen as ensuring the best medical outcome. In this context, it is extremely important to accurately identify diseases risk factors. Interdisciplinary approach of these aspects, interpreting medical data using advanced statistical tools can offer extra knowledge. The paper studied the risk factors for the severity of orthodontic root resorption. The MDS visualization technique was adopted for exploring the data from patients receiving orthodontic treatment at the Department of Orthodontics and Dentofacial Orthopedics, Faculty of Dentistry, Carol Davila University of Medicine and Pharmacy, during a period of 4 years. The clusters in the MDS charts reveal features not easily captured by classical statistical tools and overcome noise effects embedded into the data.

The method introduced in this paper is rapid, efficient, and very useful for identifying the risk factors for the severity of orthodontic root resorption.

References

- [1] J. E. Lupi, C. S. Handelman, and C. Sadowsky, "Prevalence and severity of apical root resorption and alveolar bone loss in orthodontically treated adults," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 109, no. 1, pp. 28–37, 1996.
- [2] G. R. Segal, P. H. Schiffman, and O. C. Tuncay, "Meta analysis of the treatment-related factors of external apical root resorption," *Orthodontics & Craniofacial Research*, vol. 7, no. 2, pp. 71–78, 2004.
- [3] K. Lopatiene and A. Dumbravaite, "Risk factors of root resorption after orthodontic treatment," *Stomatologija*, vol. 10, no. 3, pp. 89–95, 2008.
- [4] E. Lim, G. Samemshima, P. Petocz, and A. Darendeliler, "Comparison of Australian and American orthodontic clinical approaches towards root resorption," *Australian Orthodontic Journal*, vol. 28, no. 2, pp. 181–189, 2012.
- [5] T. Gruber, T. Eliades, and A. Athanasiou, *Risk Management in Orthodontics: Experts' Guide to Malpractice*, Quintessence, Chicago, Ill, USA, 2004.
- [6] S. Ericson and J. Kurol, "Resorption of incisors after ectopic eruption of maxillary canines: a CT study," *Angle Orthodontist*, vol. 70, no. 6, pp. 415–423, 2000.
- [7] J. Artun, V. R. Hullenaar, D. Doppel, and A. M. Kuijpers-Jagtman, "Identification of orthodontic patients at risk of severe apical root resorption," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 135, no. 4, pp. 448–455, 2009.
- [8] I. Smale, J. Årtun, F. Behbehani, D. Doppel, M. Van't Hof, and A. M. Kuijpers-Jagtman, "Apical root resorption 6 months after initiation of fixed orthodontic appliance therapy," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 128, no. 1, pp. 57–67, 2005.

- [9] C. T. Preoteasa, Ionescu, and E. Preoteasa, "Risks and complications associated with orthodontic treatment," in *Orthodontics-Basic Aspects and Clinical Considerations*, F. Bourzgui, Ed., InTech, Rijeka, Croatia, 2012.
- [10] N. Fox, "Longer orthodontic treatment may result in greater external apical root resorption," *Evidence-Based Dentistry*, vol. 6, no. 1, p. 21, 2005.
- [11] B. Weltman, K. W. L. Vig, H. W. Fields, S. Shanker, and E. E. Kaizar, "Root resorption associated with orthodontic tooth movement: a systematic review," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 137, no. 4, pp. 462–476, 2010.
- [12] J. Kruskal, "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis," *Psychometrika*, vol. 29, no. 1, pp. 1–27, 1964.
- [13] J. Sammon, "A nonlinear mapping for data structure analysis," *IEEE Transactions on Computers C*, vol. 18, no. 5, pp. 401–409, 1969.
- [14] J. Kruskal and M. Wish, *Multidimensional Scaling*, Sage, Newbury Park, Calif, USA, 1978.
- [15] J. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.
- [16] T. Cox and M. Cox, *Multidimensional Scaling*, Chapman & Hall/CRC, New York, NY, USA, 2nd edition, 2001.
- [17] I. Fodor, "A survey of dimension reduction techniques," Technical Report, Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, Livermore, Calif, USA, 2002.
- [18] I. Borg and P. Groenen, *Modern Multidimensional Scaling: Theory and Applications*, Springer, New York, NY, USA, 2nd edition, 2005.
- [19] R. N. Shepard, "The analysis of proximities: multidimensional scaling with an unknown distance function. II," *Psychometrika*, vol. 27, no. 3, pp. 219–246, 1962.
- [20] W. Glunt, T. L. Hayden, and M. Raydan, "Molecular conformation from distance matrices," *Journal of Computational Chemistry*, vol. 14, no. 1, pp. 114–120, 1993.
- [21] J. Tzeng, H. Lu, and W. Li, "Multidimensional scaling for large genomic data sets," *BMC Bioinformatics*, vol. 9, p. 179, 2008.
- [22] G. Mao and B. Fidan, *Localization Algorithms and Strategies for Wireless Sensor Networks*, Igi-Global, Hershey, Pa, USA, 2009.
- [23] J. A. Tenreiro Machado, A. C. Costa, and M. F. M. Lima, "Dynamical analysis of compositions," *Nonlinear Dynamics*, vol. 65, no. 4, pp. 399–412, 2011.
- [24] J. A. T. Machado, G. M. Duarte, and F. B. Duarte, "Identifying economic periods and crisis with the multidimensional scaling," *Nonlinear Dynamics*, vol. 63, no. 4, pp. 611–622, 2011.
- [25] C. M. Ionescu, J. T. Machado, and R. De Keyser, "Is multidimensional scaling suitable for mapping the input respiratory impedance in subjects and patients?" *Computer Methods and Programs in Biomedicine*, vol. 104, no. 3, pp. e189–e200, 2011.
- [26] J. A. T. Machado, A. C. Costa, and M. D. Quelhas, "Analysis and visualization of chromosome information," *Gene*, vol. 491, no. 1, pp. 81–87, 2012.
- [27] M. C. Hout, M. H. Papesh, and S. D. Goldinger, "Multidimensional scaling," *Wiley Interdisciplinary Reviews*, vol. 4, no. 1, pp. 93–103, 2013.
- [28] C. Bécavin, N. Tchitchev, C. Mintsa-Eya, A. Lesne, and A. Benecke, "Improving the efficiency of multidimensional scaling in the analysis of high-dimensional data using singular value decomposition," *Bioinformatics*, vol. 27, no. 10, Article ID btr143, pp. 1413–1421, 2011.
- [29] L. Linge and B. O. Linge, "Patient characteristics and treatment variables associated with apical root resorption during orthodontic treatment," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 99, no. 1, pp. 35–43, 1991.
- [30] E. Deza and M. M. Deza, *Dictionary of Distances*, Elsevier, New York, NY, USA, 2006.
- [31] GGobi, "Interactive and dynamic graphics," <http://www.ggobi.org/>.
- [32] W. Martinez and A. Martinez, *Exploratory Data Analysis with MATLAB*, Chapman & Hall/CRC, New York, NY, USA, 2005.
- [33] J. de Leeuw and P. Mair, "Multidimensional scaling using majorization: SMACOF in R," *Journal of Statistical Software*, vol. 31, no. 3, pp. 1–30, 2009.
- [34] K. Nanekrungsan, V. Patanaporn, A. Janhom, and N. Kornwanich, "External apical root resorption in maxillary incisors in orthodontic patients: associated factors and radiographic evaluation," *Imaging Science in Dentistry*, vol. 42, no. 3, pp. 147–154, 2012.
- [35] L. S. Marques, M. L. Ramos-Jorge, A. C. Rey, M. C. Armond, and A. C. de Oliveira Ruellas, "Severe root resorption in orthodontic patients treated with the edgewise method: prevalence and predictive factors," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 137, no. 3, pp. 384–388, 2010.
- [36] M. R. de Freitas, R. T. Soares Beltrão, G. Janson, J. F. Castanha Henriques, and K. Chiqueto, "Evaluation of root resorption after open bite treatment with and without extractions," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 132, no. 2, pp. 143.e15–143.e22, 2007.
- [37] H. Mohandesan, H. Ravanmehr, and N. Valaei, "A radiographic analysis of external apical root resorption of maxillary incisors during active orthodontic treatment," *European Journal of Orthodontics*, vol. 29, no. 2, pp. 134–139, 2007.
- [38] M. Motokawa, T. Sasamoto, M. Kaku et al., "Association between root resorption incident to orthodontic treatment and treatment factors," *European Journal of Orthodontics*, vol. 35, no. 2, p. 273, 2013.

Research Article

Systems of Navier-Stokes Equations on Cantor Sets

Xiao-Jun Yang,¹ Dumitru Baleanu,^{2,3,4} and J. A. Tenreiro Machado⁵

¹ Department of Mathematics and Mechanics, China University of Mining and Technology, Xuzhou Campus, Xuzhou, Jiangsu 221008, China

² Department of Mathematics and Computer Sciences, Faculty of Arts and Sciences, Cankaya University, 06530 Ankara, Turkey

³ Institute of Space Sciences, Magurele, Bucharest, Romania

⁴ Department of Chemical and Materials Engineering, Faculty of Engineering, King Abdulaziz University, P.O. Box 80204, Jeddah 21589, Saudi Arabia

⁵ Department of Electrical Engineering, Institute of Engineering, Polytechnic of Porto, Rua Dr. Antonio Bernardino de Almeida, 431, 4200-072 Porto, Portugal

Correspondence should be addressed to Xiao-Jun Yang; dyangxiaojun@163.com

Received 31 March 2013; Revised 3 June 2013; Accepted 10 June 2013

Academic Editor: Bashir Ahmad

Copyright © 2013 Xiao-Jun Yang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We present systems of Navier-Stokes equations on Cantor sets, which are described by the local fractional vector calculus. It is shown that the results for Navier-Stokes equations in a fractal bounded domain are efficient and accurate for describing fluid flow in fractal media.

1. Introduction

The Navier-Stokes equations [1–3] are commonly used in describing motion of fluids in models relevant to weather, ocean currents, water flow in pipes, and so forth.

However, the turbulent flows may be of fractal character, and the smoothness required in the common problem modelled by the Navier-Stokes equations may be violated in fractal flows. Such distortions exist in both small-scale [4] and large-scale [5, 6] turbulent flows. In this context, Benzi et al. [7] investigated the scaling properties of turbulent flows while She and Leveque [8] stressed the attention on universal scaling laws in fully developed turbulence. Shraiman and Siggia also reported scalar turbulence [9], while the self-similarity in turbulent flows was analyzed by [10]. The fractal dimension of turbulent flows was pointed out in several studies [10–12], and some applications involving the Navier-Stokes equations were suggested [13–15] taking into account the multifractal nature of fully developed turbulence and chaotic systems [16]. In this context, the alpha model of turbulence [17, 18] and its applications to Navier Stokes have been discussed [19].

Based on fractional calculus theory [20–23] time-fractional Navier-Stokes equation has been recently proposed

by El-Shahed and Salem [24]. Several analytical solutions have been developed by the Adomian decomposition method [25] and the homotopy perturbation method [26].

Recently, the element of fractal arc length squared in fractal time-space was discussed [27]. Based on it, the Schrödinger equation [28], the heat conduction equation [27, 29], the wave equation [27, 30], and the diffusion equation on Cantor time-space [31] were suggested based on local fractional calculus theory [27, 32, 33]. Carpinteri and Sapora had reported diffusion equation on Cantor space [34]. Kolwankar and Gangal had proposed the Fokker-Planck equation on Cantor space [35].

In this paper we employ the local fractional vector calculus [27], which is set up on fractals, to model systems of Navier-Stokes equations on Cantor sets. The paper is organized as follows. In Section 2, a short introduction to local fractional vector calculus theory is given. Fractal kinematics, continuity equations, and constitutive equations for Newtonian fluid on Cantor sets are discussed in Section 3. Cauchy's equations of motion and mechanical energy equations on Cantor sets are investigated in Section 4. Systems of Navier-Stokes equations on Cantor sets with Cantorian coordinates systems are presented in Section 5. Finally, in Section 6, the conclusions are presented.

2. Mathematical Tool

Local fractional vector form of function gives [27]

$$\mathbf{r}(x) = M(x)\mathbf{e}_1^\alpha + N(x)\mathbf{e}_2^\alpha \quad (1)$$

which is expressed by

$$\mathbf{r}(x) = (M(x), N(x)). \quad (2)$$

For a function of three variables, the vector form can be written in the form [27]

$$\mathbf{r}(x, y, z) = L(x, y, z)\mathbf{e}_1^\alpha + M(x, y, z)\mathbf{e}_2^\alpha + N(x, y, z)\mathbf{e}_3^\alpha, \quad (3)$$

or

$$\mathbf{r}(x, y) = (L(x, y, z), M(x, y, z), N(x, y, z)). \quad (4)$$

Let $\mathbf{u}(x, y, z) = u(x, y, z)\mathbf{e}_1^\alpha + u_2(x, y, z)\mathbf{e}_2^\alpha + u_3(x, y, z)\mathbf{e}_3^\alpha$ a local fractional vector field and $\varphi(x, y, z)$ be a local fractional scalar field, the computing rules of Hamilton operator are valid as follows [27].

- (1) The local fractional gradient and Laplace operator of a local fractional scalar field are, respectively, defined as [27]

$$\begin{aligned} \nabla^\alpha \varphi &= \frac{\partial^\alpha \varphi}{\partial x_1^\alpha} \mathbf{e}_1^\alpha + \frac{\partial^\alpha \varphi}{\partial x_2^\alpha} \mathbf{e}_2^\alpha + \frac{\partial^\alpha \varphi}{\partial x_3^\alpha} \mathbf{e}_3^\alpha, \\ \nabla^{2\alpha} \varphi &= \frac{\partial^{2\alpha} \varphi}{\partial x_1^{2\alpha}} + \frac{\partial^{2\alpha} \varphi}{\partial x_2^{2\alpha}} + \frac{\partial^{2\alpha} \varphi}{\partial x_3^{2\alpha}}, \end{aligned} \quad (5)$$

where local fractional partial derivatives of high order are defined as [27, 31–33]

$$\frac{\partial^{k\alpha}}{\partial x^{k\alpha}} f(x) = \overbrace{\frac{\partial^\alpha}{\partial x^\alpha} \cdots \frac{\partial^\alpha}{\partial x^\alpha}}^{k \text{ times}} f(x). \quad (6)$$

- (2) The local fractional divergence and curl of a local fractional vector field are written in the form [27, 31]

$$\begin{aligned} \operatorname{div}^\alpha \mathbf{u} &= \nabla^\alpha \cdot \mathbf{u} = \frac{\partial^\alpha u_1}{\partial x_1^\alpha} + \frac{\partial^\alpha u_2}{\partial x_2^\alpha} + \frac{\partial^\alpha u_3}{\partial x_3^\alpha}, \\ \operatorname{curl}^\alpha \mathbf{u} &= \nabla^\alpha \times \mathbf{u} = \left(\frac{\partial^\alpha u_3}{\partial x_2^\alpha} - \frac{\partial^\alpha u_2}{\partial x_3^\alpha} \right) \mathbf{e}_1^\alpha \\ &\quad + \left(\frac{\partial^\alpha u_1}{\partial x_3^\alpha} - \frac{\partial^\alpha u_3}{\partial x_1^\alpha} \right) \mathbf{e}_2^\alpha \\ &\quad + \left(\frac{\partial^\alpha u_2}{\partial x_1^\alpha} - \frac{\partial^\alpha u_1}{\partial x_2^\alpha} \right) \mathbf{e}_3^\alpha. \end{aligned} \quad (7)$$

- (3) The following equations are valid only in Cantor coordinates [27]:

$$\begin{aligned} \nabla^\alpha (\mathbf{u}\mathbf{v}) &= (\nabla^\alpha \mathbf{u})\mathbf{v} + \mathbf{u}\nabla^\alpha \mathbf{v}, \\ \nabla^{2\alpha} \mathbf{A} &= \nabla^\alpha (\nabla^\alpha \cdot \mathbf{A}) - \nabla^\alpha \times (\nabla^\alpha \times \mathbf{A}), \\ \mathbf{A} \cdot \nabla^\alpha \mathbf{A} &= \nabla^\alpha \left(\frac{\mathbf{A} \cdot \mathbf{A}}{2} \right) - (\nabla^\alpha \times \mathbf{A}) \times \mathbf{A}, \end{aligned} \quad (8)$$

where $\nabla^{2\alpha} = \nabla^\alpha \cdot \nabla^\alpha$ is local fractional Laplace operator [27].

The local fractional line integral of the function $\mathbf{u}(x_P, y_P, z_P)$ in the local fractional vector form

$$\begin{aligned} \mathbf{u}(x_P, y_P, z_P) &= u_1(x_P, y_P, z_P)\mathbf{e}_1^\alpha + u_2(x_P, y_P, z_P)\mathbf{e}_2^\alpha \\ &\quad + u_3(x_P, y_P, z_P)\mathbf{e}_3^\alpha \end{aligned} \quad (9)$$

along a fractal line l^α is written as [27]

$$\int_{l^\alpha} \mathbf{u}(x_P, y_P, z_P) \cdot d\mathbf{l}^\alpha = \lim_{N \rightarrow \infty} \sum_{P=1}^N \mathbf{u}(x_P, y_P, z_P) \cdot \Delta \mathbf{l}_P^\alpha, \quad (10)$$

where for $P = 1, 2, \dots, N$ and N elements of line $\Delta \mathbf{l}_P^\alpha$, it is required that all $|\Delta \mathbf{l}_P^\alpha| \rightarrow 0$ as $N \rightarrow \infty$.

The local fractional surface integral of the given function (9) across a surface $S^{(\beta)}$ is defined as [27]

$$\begin{aligned} \iint_{S^{(\beta)}} \mathbf{u}(x_P, y_P, z_P) \cdot d\mathbf{S}^{(\beta)} &= \lim_{N \rightarrow \infty} \sum_{P=1}^N \mathbf{u}(x_P, y_P, z_P) \cdot \mathbf{n}_P \Delta S_P^{(\beta)}, \end{aligned} \quad (11)$$

where for $P = 1, 2, \dots, N$ and N elements of area $\Delta S_P^{(\beta)}$ with a unit normal local fractional vector \mathbf{n}_P , it is required that all $\Delta S_P^{(\beta)} \rightarrow 0$ as $N \rightarrow \infty$.

The local fractional volume integral of the given function (9) in a fractal region $V^{(y)}$ is given by [27]

$$\iiint_{V^{(y)}} \mathbf{u}(x_P, y_P, z_P) dV^{(y)} = \lim_{N \rightarrow \infty} \sum_{P=1}^N \mathbf{u}(x_P, y_P, z_P) \Delta V_P^{(y)}, \quad (12)$$

where for $P = 1, 2, \dots, N$ and N elements of volume $\Delta V_P^{(y)}$, it is required that all $\Delta V_P^{(y)} \rightarrow 0$ as $N \rightarrow \infty$.

Let us consider a local fractional vector field $\mathbf{u} = u_1\mathbf{e}_1^\alpha + u_2\mathbf{e}_2^\alpha + u_3\mathbf{e}_3^\alpha$, the following results hold [27]:

- (a) Divergence Theorem of local fractional field states that

$$\iiint_{V^{(y)}} \nabla^\alpha \cdot \mathbf{u} dV^{(y)} = \iint_{S^{(\beta)}} \mathbf{u} \cdot d\mathbf{S}^{(\beta)}. \quad (13)$$

- (b) Stokes' theorem of local fractional field states that

$$\oint_{l^{(\alpha)}} \mathbf{u} \cdot d\mathbf{l}^\alpha = \iint_{S^{(\beta)}} (\nabla^\alpha \times \mathbf{u}) \cdot d\mathbf{S}^{(\beta)}. \quad (14)$$

- (c) Green's first theorem in fractal domain states that

$$\begin{aligned} \iint_{S^{(\beta)}} \phi \nabla^\alpha \varphi \cdot d\mathbf{S}^{(\beta)} &= \iiint_{V^{(y)}} (\phi \nabla^{2\alpha} \varphi + (\nabla^\alpha \phi) \cdot (\nabla^\alpha \varphi)) dV^{(y)}. \end{aligned} \quad (15)$$

(d) Green's second theorem in fractal domain states that

$$\begin{aligned} & \oint\!\oint_{S^{(\beta)}} (\phi \nabla^\alpha \varphi - \phi \nabla^\alpha \varphi) \cdot d\mathbf{S}^{(\beta)} \\ &= \iiint_{V^{(\gamma)}} (\phi \nabla^{2\alpha} \varphi - \varphi \nabla^{2\alpha} \phi) dV^{(\gamma)}, \end{aligned} \quad (16)$$

where the relationships of fractal dimensions of line, surface, and volume fields are

$$\beta = 2\alpha, \quad \gamma = 3\alpha. \quad (17)$$

3. Fractal Kinematics, Continuity Equations, and Constitutive Equations for Newtonian Fluid on Cantor Sets

In this section, we investigate the fractal kinematics, balance of mass on Cantor sets, and constitutive equations for Newtonian fluid on Cantor sets. We first start with fractal kinematics.

3.1. Fractal Kinematics. For any scalar property ψ associated with an abstract Cantor body B expressed as

$$\psi = \psi(P, t) = \hat{\psi}(X, t) = \tilde{\psi}(x, t), \quad (18)$$

we define the following Lagrangian and Eulerian temporal and spatial derivatives with local fractional differential operator:

$$\begin{aligned} \dot{\psi} &= \frac{D^\alpha \hat{\psi}}{Dt^\alpha}, \quad \psi_{,t}^\alpha = \frac{\partial^\alpha \tilde{\psi}}{\partial t^\alpha}, \quad \nabla_X^\alpha \psi = \frac{\partial^\alpha \hat{\psi}}{\partial X^\alpha}, \\ \nabla_x^\alpha \psi &= \nabla^\alpha \psi = \frac{\partial^\alpha \tilde{\psi}}{\partial x^\alpha}. \end{aligned} \quad (19)$$

The velocity \mathbf{v} and the acceleration \mathbf{a} with local fractional operator are defined through

$$\mathbf{v} = \frac{D^\alpha \chi}{Dt^\alpha}, \quad \mathbf{a} = \frac{D^{2\alpha} \chi}{Dt^{2\alpha}}. \quad (20)$$

It immediately follows that the velocity gradient Θ and its symmetric part Λ are expressed through

$$\Theta = \nabla^\alpha \cdot \mathbf{v}, \quad \Theta^T = \mathbf{v} \cdot \nabla^\alpha, \quad \Lambda = \frac{1}{2} (\Theta + \Theta^T). \quad (21)$$

The deformation gradient Σ is given by

$$\Sigma = \nabla^\alpha \cdot \chi, \quad (22)$$

which leads to

$$\Theta = \frac{\partial^\alpha \Sigma}{\partial t^\alpha} \Sigma^{-1}. \quad (23)$$

The fractal material derivative of the fluid density ρ is defined as [27]

$$\frac{D^\alpha \rho}{Dt^\alpha} = \frac{\partial^\alpha \rho}{\partial t^\alpha} + \mathbf{v} \cdot \nabla^\alpha \rho. \quad (24)$$

3.2. Balance of Mass on Cantor Sets. The mass of fractal fluid in $V^{(\gamma)}$ is defined through [27]

$$M = \iiint_{V^{(\gamma)}} \rho dV^{(\gamma)}, \quad (25)$$

which yields that the balance of mass on Cantor sets takes the form [27]

$$\frac{\partial^\alpha M}{\partial t^\alpha} = - \oint\!\oint_{S^{(\beta)}} \rho \mathbf{v} \cdot d\mathbf{S}^{(\beta)}. \quad (26)$$

Using Divergence Theorem of local fractional field, we deduce to

$$\begin{aligned} & \iiint_{V^{(\gamma)}} \frac{\partial^\alpha}{\partial t^\alpha} \rho dV^{(\gamma)} + \oint\!\oint_{S^{(\beta)}} \rho \mathbf{v} \cdot d\mathbf{S}^{(\beta)} \\ &= \iiint_{V^{(\gamma)}} \left[\frac{\partial^\alpha \rho}{\partial t^\alpha} + \nabla^\alpha \cdot (\rho \mathbf{v}) \right] dV^{(\gamma)} = 0 \end{aligned} \quad (27)$$

which implies that

$$\frac{\partial^\alpha \rho}{\partial t^\alpha} + \nabla^\alpha \cdot (\rho \mathbf{v}) = 0. \quad (28)$$

This is called continuity equation on Cantor sets.

It becomes

$$\begin{aligned} \frac{\partial^\alpha \rho}{\partial t^\alpha} + \nabla^\alpha \cdot (\rho \mathbf{v}) &= \frac{\partial^\alpha \rho}{\partial t^\alpha} + \mathbf{v} \cdot (\nabla^\alpha \rho) + \rho (\nabla^\alpha \cdot \mathbf{v}) \\ &= \frac{D^\alpha \rho}{Dt^\alpha} + \rho (\nabla^\alpha \cdot \mathbf{v}) = 0, \end{aligned} \quad (29)$$

where the fractal material derivative of the fluid density ρ is noted by

$$\frac{D^\alpha \rho}{Dt^\alpha} = \frac{\partial^\alpha \rho}{\partial t^\alpha} + \mathbf{v} \cdot (\nabla^\alpha \rho). \quad (30)$$

If the fractal fluid is incompressible, we deduce that

$$\frac{\partial^\alpha \rho}{\partial t^\alpha} + \mathbf{v} \cdot (\nabla^\alpha \rho) = 0 \iff \frac{D^\alpha \rho}{Dt^\alpha} = 0 \quad \text{or} \quad \nabla^\alpha \cdot \mathbf{v} = 0. \quad (31)$$

It follows that Reynolds transport theorem on Cantor sets states that

$$\begin{aligned} \frac{D^\alpha}{Dt^\alpha} \iiint_{V^{(\gamma)}} F(x, t) dV^{(\gamma)} &= \iiint_{V^{(\gamma)}} \frac{\partial^\alpha}{\partial t^\alpha} F(x, t) dV^{(\gamma)} \\ &+ \oint\!\oint_{S^{(\beta)}} F(x, t) \mathbf{v} \cdot d\mathbf{S}^{(\beta)}, \end{aligned} \quad (32)$$

where \mathbf{v} is the fractal fluid velocity, $V^{(\gamma)}$ is the fractal material volume, and $S^{(\beta)}$ is the fractal surfaces moving with the fractal fluid.

3.3. Constitutive Equations for Newtonian Fluid on Cantor Sets. We will assume a linear relation of the type of fractal Cauchy stress

$$\mathbf{J} = -p\mathbf{I} + \mathbf{K} : \mathbf{\Lambda}, \quad (33)$$

where p is the thermodynamic pressure, \mathbf{K} is a fourth-order fractal tensor, Λ is the fractal strain rate tensor, and \mathbf{I} is unit vector in local fractional field.

The fractal velocity gradient tensor can be decomposed into symmetric and antisymmetric parts [27]

$$\nabla^\alpha \cdot \mathbf{v} = \frac{1}{2} (\Theta + \Theta^T) + \frac{1}{2} (\Theta - \Theta^T) = \Lambda + \frac{1}{2} (\Theta - \Theta^T). \quad (34)$$

We can write

$$\mathbf{K} : \Lambda = 2\mu\Lambda + \lambda(\nabla^\alpha \cdot \mathbf{v})\mathbf{I} \quad (35)$$

which leads to

$$\mathbf{J} = -p\mathbf{I} + 2\mu\Lambda + \lambda(\nabla^\alpha \cdot \mathbf{v})\mathbf{I}, \quad (36)$$

where $(\nabla^\alpha \cdot \mathbf{v})$ is the fractal volumetric strain rate and $\Lambda = (1/2)(\nabla^\alpha \cdot \mathbf{v} + \mathbf{v} \cdot \nabla^\alpha)$ [27], λ and μ are the bulk and shear moduli of viscosity.

The constitutive equation of homogeneous compressible Euler fluid on Cantor sets reads

$$\mathbf{J} = -p(\rho)\mathbf{I}. \quad (37)$$

The constitutive equation of homogeneous incompressible Euler fluid on Cantor sets is

$$\mathbf{J} = -p\mathbf{I}, \quad \nabla^\alpha \cdot \mathbf{v} = 0. \quad (38)$$

The constitutive equation of homogeneous compressible Navier-Stokes fluid on Cantor sets is written in the form

$$\mathbf{J} = -p(\rho)\mathbf{I} + 2\mu(\rho)\Lambda + \lambda(\rho)(\nabla^\alpha \cdot \mathbf{v})\mathbf{I}. \quad (39)$$

The constitutive equation of homogeneous incompressible Navier-Stokes fluid on Cantor sets is expressed as

$$\mathbf{J} = -p\mathbf{I} + 2\mu\Lambda, \quad \nabla^\alpha \cdot \mathbf{v} = 0. \quad (40)$$

By using Stokes' assumption $\lambda(\rho) = -2\mu(\rho)/3$, the constitutive equation of homogeneous compressible Navier-Stokes fluid on Cantor sets becomes

$$\mathbf{J} = -p(\rho)\mathbf{I} + 2\mu(\rho)(\nabla^\alpha \cdot \mathbf{v} + \mathbf{v} \cdot \nabla^\alpha) - \frac{2}{3}\mu(\rho)(\nabla^\alpha \cdot \mathbf{v}), \quad (41)$$

while the constitutive equation of homogeneous incompressible Navier-Stokes fluid on Cantor sets can obtain that

$$\mathbf{J} = -p\mathbf{I} + \mu(\nabla^\alpha \cdot \mathbf{v} + \mathbf{v} \cdot \nabla^\alpha), \quad \nabla^\alpha \cdot \mathbf{v} = 0. \quad (42)$$

4. Cauchy's Equations of Motion and Mechanical Energy Equations on Cantor Sets

In this section, we consider balance of linear and angular momentums and energy on Cantor sets.

4.1. Balances of Linear and Angular Momentums on Cantor Sets. If the second law of Newton in fractal mechanics is valid, the balance of linear momentum for flows on Cantor sets takes the form [27]

$$\frac{D^\alpha}{Dt^\alpha} \iiint_{V^{(\gamma)}} \rho \mathbf{v} dV^{(\gamma)} = \iiint_{V^{(\gamma)}} \rho \mathbf{b} dV^{(\gamma)} + \oint_{S^{(\beta)}} \mathbf{J} \cdot d\mathbf{S}^{(\beta)}, \quad (43)$$

where \mathbf{J} denotes the fractal Cauchy stress tensor [27] and \mathbf{b} denotes the specific fractal body force [27].

By using (32), (43) becomes

$$\begin{aligned} \frac{D^\alpha}{Dt^\alpha} \iiint_{V^{(\gamma)}} \rho \mathbf{v} dV^{(\gamma)} &= \iiint_{V^{(\gamma)}} \frac{\partial^\alpha}{\partial t^\alpha} (\rho \mathbf{v}) dV^{(\gamma)} \\ &\quad + \oint_{S^{(\beta)}} \rho \mathbf{v} \mathbf{v} \cdot d\mathbf{S}^{(\beta)} \\ &= \iiint_{V^{(\gamma)}} \rho \mathbf{b} dV^{(\gamma)} + \oint_{S^{(\beta)}} \mathbf{J} \cdot d\mathbf{S}^{(\beta)} \end{aligned} \quad (44)$$

which is rewritten as

$$\frac{\partial^\alpha}{\partial t^\alpha} (\rho \mathbf{v}) + \nabla^\alpha \cdot [\rho \mathbf{v} \mathbf{v}] = \rho \mathbf{b} + \nabla^\alpha \cdot \mathbf{J}. \quad (45)$$

In view of (7), (45) yields

$$\frac{\partial^\alpha \rho}{\partial t^\alpha} \mathbf{v} + \rho \frac{\partial^\alpha \mathbf{v}}{\partial t^\alpha} + \mathbf{v} \mathbf{v} \cdot \nabla^\alpha \rho + \rho \mathbf{v} \nabla^\alpha \cdot \mathbf{v} + \rho \mathbf{v} \cdot \nabla^\alpha \mathbf{v} = \rho \mathbf{b} + \nabla^\alpha \cdot \mathbf{J} \quad (46)$$

which implies that

$$\mathbf{v} \left(\frac{\partial^\alpha \rho}{\partial t^\alpha} + \nabla^\alpha (\mathbf{v} \cdot \rho) \right) + \rho \left(\frac{\partial^\alpha \mathbf{v}}{\partial t^\alpha} + \mathbf{v} \cdot \nabla^\alpha \mathbf{v} \right) = \rho \mathbf{b} + \nabla^\alpha \cdot \mathbf{J} \quad (47)$$

or

$$\begin{aligned} \mathbf{v} \left(\frac{\partial^\alpha \rho}{\partial t^\alpha} + \mathbf{v} \cdot \nabla^\alpha \rho + \rho \nabla^\alpha \cdot \mathbf{v} \right) + \rho \left(\frac{\partial^\alpha \mathbf{v}}{\partial t^\alpha} + \mathbf{v} \cdot \nabla^\alpha \mathbf{v} \right) \\ = \rho \mathbf{b} + \nabla^\alpha \cdot \mathbf{J}. \end{aligned} \quad (48)$$

Therefore, (48) is reexpressed by

$$\mathbf{v} \left(\frac{\partial^\alpha \rho}{\partial t^\alpha} + \nabla^\alpha \cdot (\rho \mathbf{v}) \right) + \rho \left(\frac{\partial^\alpha \mathbf{v}}{\partial t^\alpha} + \mathbf{v} \cdot \nabla^\alpha \mathbf{v} \right) = \rho \mathbf{b} + \nabla^\alpha \cdot \mathbf{J}. \quad (49)$$

Taking continuity equation on Cantor sets and fractal material derivative yields

$$\rho \frac{D^\alpha \mathbf{v}}{Dt^\alpha} = \rho \mathbf{b} + \nabla^\alpha \cdot \mathbf{J}. \quad (50)$$

Hence, balance of linear momentum in its local Cantorian form (the Newton's law in its local Cantorian form) is rewritten as

$$\rho \frac{D^\alpha \mathbf{v}}{Dt^\alpha} = \nabla^\alpha \cdot \mathbf{J} + \rho \mathbf{b} \quad (51)$$

or

$$\rho \frac{\partial^\alpha \mathbf{v}}{\partial t^\alpha} = \nabla^\alpha \cdot \mathbf{J} + \rho \mathbf{b} - \rho \mathbf{v} (\nabla^\alpha \cdot \mathbf{v}). \quad (52)$$

It is called as the Cauchy's equation of motion of flows on Cantor sets.

In the absence of internal couples, the balance of angular momentum on Cantor sets implies that the fractal Cauchy stress is symmetric [27]; that is,

$$\rho \frac{D^\alpha \mathbf{v}}{Dt^\alpha} = \nabla^\alpha \cdot \mathbf{J}^T + \rho \mathbf{b} \quad \text{or} \quad \mathbf{J} = \mathbf{J}^T. \quad (53)$$

Taking fractal material derivative, (45) is expressed by

$$\rho \frac{\partial^\alpha \mathbf{v}}{\partial t^\alpha} + \rho \mathbf{v} (\nabla^\alpha \cdot \mathbf{v}) = \nabla^\alpha \cdot \mathbf{J}^T + \rho \mathbf{b} \quad (54)$$

or

$$\mathbf{J} = \mathbf{J}^T. \quad (55)$$

which yields the Newton's law in local fractional integration form given by

$$\iiint_{V^{(\gamma)}} \nabla^\alpha \cdot (\rho \mathbf{v} \cdot \mathbf{v} - \mathbf{J}^T) dV^{(\gamma)} = \oint_{S^{(\beta)}} [\rho \mathbf{v} \cdot \mathbf{v} - \mathbf{J}^T] d\mathbf{S}^{(\beta)} \quad (56)$$

if the compressible flow on Cantor sets is steady and body forces are absent. The results are different from [36–38] because of the fractional operators.

In local Eulerian form, Cauchy's equation of motion of flows on Cantor sets is given by using (37)

$$\frac{D^\alpha}{Dt^\alpha} \iiint_{V^{(\gamma)}} \rho \mathbf{v} dV^{(\gamma)} = \iiint_{V^{(\gamma)}} \rho \mathbf{b} dV^{(\gamma)} - \oint_{S^{(\beta)}} p \mathbf{I} \cdot d\mathbf{S}^{(\beta)} \quad (57)$$

which implies that

$$\rho \frac{D^\alpha \mathbf{v}}{Dt^\alpha} = -\nabla^\alpha p (\rho) + \rho \mathbf{b}. \quad (58)$$

For compressible fluid, Cauchy's equation of motion of flows on Cantor sets is obtained by using (39)

$$\begin{aligned} \frac{D^\alpha}{Dt^\alpha} \iiint_{V^{(\gamma)}} \rho \mathbf{v} dV^{(\gamma)} \\ = \iiint_{V^{(\gamma)}} \rho \mathbf{b} dV^{(\gamma)} \\ - \oint_{S^{(\beta)}} \left(p(\rho) \mathbf{I} + \mu(\rho) \left(\frac{1}{3} (\nabla^\alpha \cdot \mathbf{v}) + \mathbf{v} \cdot \nabla^\alpha \right) \right) \cdot d\mathbf{S}^{(\beta)} \end{aligned} \quad (59)$$

which is

$$\begin{aligned} \rho \frac{D^\alpha \mathbf{v}}{Dt^\alpha} = & -\nabla^\alpha p (\rho) + \frac{1}{3} \nabla^\alpha \cdot [\mu(\rho) (\nabla^\alpha \cdot \mathbf{v})] \\ & + \nabla^\alpha \cdot (\mu(\rho) \mathbf{v} \cdot \nabla^\alpha) + \rho \mathbf{b}. \end{aligned} \quad (60)$$

For incompressible fluid, Cauchy's equation of motion of flows on Cantor sets is obtained by using (40)

$$\begin{aligned} \frac{D^\alpha}{Dt^\alpha} \iiint_{V^{(\gamma)}} \rho \mathbf{v} dV^{(\gamma)} \\ = \iiint_{V^{(\gamma)}} \rho \mathbf{b} dV^{(\gamma)} \\ + \oint_{S^{(\beta)}} (-p \mathbf{I} + \mu (\nabla^\alpha \cdot \mathbf{v} + \mathbf{v} \cdot \nabla^\alpha)) \cdot d\mathbf{S}^{(\beta)} \end{aligned} \quad (61)$$

which leads to

$$\rho \frac{D^\alpha \mathbf{v}}{Dt^\alpha} = -\nabla^\alpha p + \mu \nabla^\alpha (\mathbf{v} \cdot \nabla^\alpha) + \rho \mathbf{b} \quad (62)$$

or

$$\rho \frac{D^\alpha \mathbf{v}}{Dt^\alpha} = -\nabla^\alpha p + \mu \nabla^\alpha (\mathbf{v} \cdot \nabla^\alpha). \quad (63)$$

4.2. Balance of Energy on Cantor Sets. The mechanical energy equation with viscous dissipation on Cantor sets is repeated by

$$\begin{aligned} \frac{D^\alpha}{Dt^\alpha} \iiint_{V^{(\gamma)}} (\rho \varphi) dV^{(\gamma)} + \oint_{S^{(\beta)}} (\rho \varphi \mathbf{v}) \cdot d\mathbf{S}^{(\beta)} \\ = \iiint_{V^{(\gamma)}} (\rho \mathbf{b} \cdot \mathbf{v}) dV^{(\gamma)} + \oint_{S^{(\beta)}} \mathbf{v} \cdot \mathbf{J} d\mathbf{S}^{(\beta)} \\ - \iiint_{V^{(\gamma)}} \nabla^\alpha \cdot (p \mathbf{v}) dV^{(\gamma)} - \iiint_{V^{(\gamma)}} \phi dV^{(\gamma)}, \end{aligned} \quad (64)$$

where the first term is rate of change of kinetic energy in fractal domain, the second term is rate of outflow across fractal boundary, the third term is rate of work by fractal body force, the fourth term is rate of work by surface force, the fifth term is rate of work by volume expansion, and the last term is rate of viscous dissipation.

Local fractional differential form of the mechanical energy equation with viscous dissipation on Cantor sets gives

$$\begin{aligned} \frac{D^\alpha}{Dt^\alpha} (\rho \varphi) + \nabla^\alpha \cdot (\rho \varphi \mathbf{v}) = \rho \mathbf{b} \cdot \mathbf{v} + \mathbf{v} \cdot (\nabla^\alpha \cdot \mathbf{J}) \\ + \nabla^\alpha \cdot (p \mathbf{v}) - \phi, \end{aligned} \quad (65)$$

which implies for fractal material derivative of $\rho \varphi$

$$\begin{aligned} \frac{\partial^\alpha}{\partial t^\alpha} (\rho \varphi) + \mathbf{v} \cdot \nabla^\alpha (\rho \varphi) + \nabla^\alpha \cdot (\rho \varphi \mathbf{v}) \\ = \rho \mathbf{b} \cdot \mathbf{v} + \mathbf{v} \cdot (\nabla^\alpha \cdot \mathbf{J}) - \nabla^\alpha \cdot (p \mathbf{v}) - \phi, \end{aligned} \quad (66)$$

where the rate of viscous dissipation of kinetic energy per unit volume is ϕ .

The integration form of the balance of energy on Cantor sets is expressed by

$$\begin{aligned} & \frac{D^\alpha}{Dt^\alpha} \iiint_{V^{(\gamma)}} \rho (\theta + \varphi) dV^{(\gamma)} \\ &= \oint\oint_{S^{(\beta)}} \mathbf{v} \cdot \mathbf{J} \cdot d\mathbf{S}^{(\beta)} - \oint\oint_{S^{(\beta)}} \rho (\theta + \varphi) \mathbf{v} \cdot d\mathbf{S}^{(\beta)} \\ &+ \iiint_{V^{(\gamma)}} (\rho \mathbf{b} \cdot \mathbf{v}) dV^{(\gamma)} + \oint\oint_{S^{(\beta)}} K^{2\alpha} \mathbf{q} \cdot d\mathbf{S}^{(\beta)} \\ &- \iiint_{V^{(\gamma)}} \phi dV^{(\gamma)} \end{aligned} \quad (67)$$

which leads to

$$\begin{aligned} \frac{D^\alpha}{Dt^\alpha} [\rho (\theta + \varphi)] &= \mathbf{v} \cdot (\nabla^\alpha \cdot \mathbf{J}) + \rho \mathbf{b} \cdot \mathbf{v} - \nabla^\alpha \cdot (p\mathbf{v}) \\ &- \nabla^\alpha \cdot [\rho (\theta + \varphi) \mathbf{v} - K^{2\alpha} \cdot \mathbf{q}] - \phi, \end{aligned} \quad (68)$$

where $K^{2\alpha}$ is denoted as fractal heat conduction coefficient and \mathbf{q} is fractal temperature flied.

The term

$$\oint\oint_{S^{(\beta)}} \rho (\theta + \varphi) \mathbf{v} \cdot d\mathbf{S}^{(\beta)} = \iiint_{V^{(\gamma)}} \nabla^\alpha \cdot [\rho (\theta + \varphi)] \mathbf{v} dV^{(\gamma)} \quad (69)$$

vanishes if the term $\rho\theta\mathbf{v}$ is zero at the boundaries and

$$\begin{aligned} \iiint_{V^{(\gamma)}} (p (\nabla^\alpha \cdot \mathbf{v})) dV^{(\gamma)} &= \oint\oint_{S^{(\beta)}} (p\mathbf{v}) d\mathbf{S}^{(\beta)} \\ &- \iiint_{V^{(\gamma)}} [\mathbf{v} \cdot (\nabla^\alpha p)] dV^{(\gamma)}. \end{aligned} \quad (70)$$

The mechanical energy equation with no viscous dissipation on Cantor sets is written in the local Cantorian form as

$$\rho \frac{D^\alpha}{Dt^\alpha} (\theta + \varphi) = -\nabla^\alpha p \cdot \mathbf{v} + \mathbf{v} \cdot (\nabla^\alpha \cdot \mathbf{J}) + \rho \mathbf{b} \cdot \mathbf{v} + K^{2\alpha} \nabla^\alpha \cdot \mathbf{q}, \quad (71)$$

where φ is the kinetic energy per unit of mass, \mathbf{b} is the external force per unit of mass, \mathbf{J} is the fractal Cauchy stress tensor, and θ is the internal energy per unit of mass.

5. Systems of Navier-Stokes Equation on Cantor Sets

In this section, we investigate the Navier-Stokes equation on Cantor sets and derive systems of Navier-Stokes equation on Cantor sets.

5.1. Navier-Stokes Equation on Cantor Sets. Substituting the constitutive equation into Cauchy's equation (41) yields

$$\begin{aligned} \rho \frac{D^\alpha \mathbf{v}}{Dt^\alpha} &= -\nabla^\alpha \cdot (p\mathbf{I}) \\ &+ \nabla^\alpha \left[2\mu (\nabla^\alpha \cdot \mathbf{v} + \mathbf{v} \cdot \nabla^\alpha) - \frac{2}{3}\mu (\nabla^\alpha \cdot \mathbf{v}) \mathbf{I} \right] + \rho \mathbf{b}, \end{aligned} \quad (72)$$

where viscosity μ in this equation can be a function of the thermodynamic state.

This is a general form of the Navier-Stokes equation on Cantor sets, which is the equation of motion for a Newtonian fluid on Cantor sets.

For incompressible fluids $\nabla^\alpha \cdot \mathbf{v} = 0$, we deduce to

$$\rho \frac{D^\alpha \mathbf{v}}{Dt^\alpha} = -\nabla^\alpha p + \mu \nabla^{2\alpha} \mathbf{v} + \rho \mathbf{b}. \quad (73)$$

Applying fractal material derivative of the fluid velocity,

$$\frac{D^\alpha \mathbf{v}}{Dt^\alpha} = \frac{\partial^\alpha \mathbf{v}}{\partial t^\alpha} + \mathbf{v} (\nabla^\alpha \cdot \mathbf{v}), \quad (74)$$

in compressible fluids, a general form of the Navier-Stokes equation on Cantor sets is stated as

$$\rho \frac{\partial^\alpha \mathbf{v}}{\partial t^\alpha} = -\nabla^\alpha p + \mu \nabla^{2\alpha} \mathbf{v} + \rho \mathbf{b} - \rho \mathbf{v} (\nabla^\alpha \cdot \mathbf{v}). \quad (75)$$

For compressible fluid, the Navier-Stokes equation on Cantor sets reads

$$\begin{aligned} \rho \frac{D^\alpha \mathbf{v}}{Dt^\alpha} &= -\nabla^\alpha p + \nabla^\alpha [\mu (\nabla^\alpha \cdot \mathbf{v} + \mathbf{v} \cdot \nabla^\alpha)] \\ &+ \nabla^\alpha (\lambda (\nabla^\alpha \cdot \mathbf{v})) + \rho \mathbf{b} \end{aligned} \quad (76)$$

which becomes for $\lambda = -(2/3)\mu$

$$\rho \frac{D^\alpha \mathbf{v}}{Dt^\alpha} = -\nabla^\alpha p + \frac{1}{3}\mu \nabla^\alpha ((\nabla^\alpha \cdot \mathbf{v})) + \mu \nabla^{2\alpha} \mathbf{v} + \rho \mathbf{b} \quad (77)$$

which yields by using (55)

$$\begin{aligned} \rho \frac{\partial^\alpha \mathbf{v}}{\partial t^\alpha} &= -\nabla^\alpha p + \frac{1}{3}\mu \nabla^\alpha ((\nabla^\alpha \cdot \mathbf{v})) \\ &+ \mu \nabla^{2\alpha} \mathbf{v} + \rho \mathbf{b} - \rho \mathbf{v} (\nabla^\alpha \cdot \mathbf{v}). \end{aligned} \quad (78)$$

5.2. Systems of Navier-Stokes Equation on Cantor Sets. In this section, we consider systems of Navier-Stokes fluid on Cantor sets, which states that the systems consists of the continuity equation, the motion equation, and the energy balance equation on Cantor sets.

By using (29), (71), and (77), systems of compressible Navier-Stokes equations on Cantor sets become as follows:

$$\begin{aligned} \frac{\partial^\alpha \rho}{\partial t^\alpha} + \nabla^\alpha \cdot (\rho \mathbf{v}) &= 0, \\ \rho \frac{D^\alpha \mathbf{v}}{Dt^\alpha} &= -\nabla^\alpha p + \frac{1}{3}\mu \nabla^\alpha ((\nabla^\alpha \cdot \mathbf{v})) + \mu \nabla^{2\alpha} \mathbf{v} + \rho \mathbf{b}, \\ \rho \frac{D^\alpha}{Dt^\alpha} (\theta + \varphi) &= -\nabla^\alpha \cdot (p\mathbf{v}) + \mathbf{v} \cdot (\nabla^\alpha \cdot \mathbf{J}) \\ &+ \rho \mathbf{b} \cdot \mathbf{v} + K^{2\alpha} \nabla^\alpha \cdot \mathbf{q}. \end{aligned} \quad (79)$$

By applying (31), (71), and (73), systems of incompressible Navier-Stokes equations on Cantor sets are stated as

$$\begin{aligned} \nabla^\alpha \cdot \mathbf{v} &= 0, \\ \rho \frac{D^\alpha \mathbf{v}}{Dt^\alpha} &= -\nabla^\alpha p + \mu \nabla^{2\alpha} \mathbf{v} + \rho \mathbf{b}, \\ \rho \frac{D^\alpha}{Dt^\alpha} (\theta + \varphi) &= -\nabla^\alpha \cdot (p \mathbf{v}) + \mathbf{v} \cdot (\nabla^\alpha \cdot \mathbf{J}) \\ &\quad + \rho \mathbf{b} \cdot \mathbf{v} + K^{2\alpha} \nabla^\alpha \cdot \mathbf{q}, \end{aligned} \quad (80)$$

or in the 3D Cantorian coordinates, systems of Navier-Stokes equations on Cantor sets can be written for each component, x , y , and z as

$$\begin{aligned} \frac{\partial^\alpha v_x}{\partial x^\alpha} + \frac{\partial^\alpha v_y}{\partial y^\alpha} + \frac{\partial^\alpha v_z}{\partial z^\alpha} &= 0, \\ \rho \frac{D^\alpha v_x}{Dt^\alpha} &= -\frac{\partial^\alpha p}{\partial x^\alpha} + \mu \left(\frac{\partial^{2\alpha} v_x}{\partial x^{2\alpha}} + \frac{\partial^{2\alpha} v_x}{\partial y^{2\alpha}} + \frac{\partial^{2\alpha} v_x}{\partial z^{2\alpha}} \right) + \rho b_x, \\ \rho \frac{D^\alpha v_y}{Dt^\alpha} &= -\frac{\partial^\alpha p}{\partial y^\alpha} + \mu \left(\frac{\partial^{2\alpha} v_y}{\partial x^{2\alpha}} + \frac{\partial^{2\alpha} v_y}{\partial y^{2\alpha}} + \frac{\partial^{2\alpha} v_y}{\partial z^{2\alpha}} \right) + \rho b_y, \\ \rho \frac{D^\alpha v_z}{Dt^\alpha} &= -\frac{\partial^\alpha p}{\partial z^\alpha} + \mu \left(\frac{\partial^{2\alpha} v_z}{\partial x^{2\alpha}} + \frac{\partial^{2\alpha} v_z}{\partial y^{2\alpha}} + \frac{\partial^{2\alpha} v_z}{\partial z^{2\alpha}} \right) + \rho b_z, \\ \rho \frac{D^\alpha}{Dt^\alpha} (\theta + \varphi) &= -\left(\frac{\partial^\alpha (p v_x)}{\partial x^\alpha} + \frac{\partial^\alpha (p v_y)}{\partial y^\alpha} + \frac{\partial^\alpha (p v_z)}{\partial z^\alpha} \right) \\ &\quad + v_x \left(\frac{\partial^\alpha J_{xx}}{\partial x^\alpha} + \frac{\partial^\alpha J_{xy}}{\partial y^\alpha} + \frac{\partial^\alpha J_{xz}}{\partial z^\alpha} \right) \\ &\quad + v_y \left(\frac{\partial^\alpha J_{yx}}{\partial x^\alpha} + \frac{\partial^\alpha J_{yy}}{\partial y^\alpha} + \frac{\partial^\alpha J_{yz}}{\partial z^\alpha} \right) \\ &\quad + v_z \left(\frac{\partial^\alpha J_{zx}}{\partial x^\alpha} + \frac{\partial^\alpha J_{zy}}{\partial y^\alpha} + \frac{\partial^\alpha J_{zz}}{\partial z^\alpha} \right) \\ &\quad + \rho b_x v_x + \rho b_y v_y + \rho b_z v_z \\ &\quad + K^{2\alpha} \left(\frac{\partial^\alpha q_x}{\partial x^\alpha} + \frac{\partial^\alpha q_y}{\partial y^\alpha} + \frac{\partial^\alpha q_z}{\partial z^\alpha} \right). \end{aligned} \quad (81)$$

In the 2D Cantorian coordinates, systems of incompressible Navier-Stokes equations on Cantor sets are rewritten in the form

$$\begin{aligned} \frac{\partial^\alpha v_x}{\partial x^\alpha} + \frac{\partial^\alpha v_y}{\partial y^\alpha} &= 0, \\ \rho \frac{D^\alpha v_x}{Dt^\alpha} &= -\frac{\partial^\alpha p}{\partial x^\alpha} + \mu \left(\frac{\partial^{2\alpha} v_x}{\partial x^{2\alpha}} + \frac{\partial^{2\alpha} v_x}{\partial y^{2\alpha}} \right) + \rho b_x, \\ \rho \frac{D^\alpha v_y}{Dt^\alpha} &= -\frac{\partial^\alpha p}{\partial y^\alpha} + \mu \left(\frac{\partial^{2\alpha} v_y}{\partial x^{2\alpha}} + \frac{\partial^{2\alpha} v_y}{\partial y^{2\alpha}} \right) + \rho b_y, \end{aligned}$$

$$\begin{aligned} \rho \frac{D^\alpha}{Dt^\alpha} (\theta + \varphi) &= -\left(\frac{\partial^\alpha p v_x}{\partial x^\alpha} + \frac{\partial^\alpha p v_y}{\partial y^\alpha} \right) \\ &\quad + v_x \left(\frac{\partial^\alpha J_{xx}}{\partial x^\alpha} + \frac{\partial^\alpha J_{xy}}{\partial y^\alpha} \right) \\ &\quad + v_y \left(\frac{\partial^\alpha J_{yx}}{\partial x^\alpha} + \frac{\partial^\alpha J_{yy}}{\partial y^\alpha} \right) + \rho b_x v_x + \rho b_y v_y \\ &\quad + K^{2\alpha} \left(\frac{\partial^\alpha q_x}{\partial x^\alpha} + \frac{\partial^\alpha q_y}{\partial y^\alpha} \right). \end{aligned} \quad (82)$$

In the 1D Cantorian coordinates, systems of incompressible Navier-Stokes equations on Cantor sets are repeated by the expression

$$\begin{aligned} \frac{\partial^\alpha v}{\partial x^\alpha} &= 0, \\ \rho \frac{D^\alpha v}{Dt^\alpha} &= -\frac{\partial^\alpha p}{\partial x^\alpha} + \mu \frac{\partial^{2\alpha} v}{\partial x^{2\alpha}} + \rho b, \\ \rho \frac{D^\alpha}{Dt^\alpha} (\theta + \varphi) &= -\frac{\partial^\alpha (p v)}{\partial x^\alpha} + v \frac{\partial^\alpha J}{\partial x^\alpha} + \rho b v + K^{2\alpha} \frac{\partial^\alpha q}{\partial x^\alpha}. \end{aligned} \quad (83)$$

6. Conclusions

In the present work, we propose the systems of Navier-Stokes equations derived from local fractional vector calculus. These obtained Navier-Stokes equations in one-, two-, and three-dimension Cantorian coordinates are shown to describe the materials as being local fractional continuous and nondifferential functions, which are applied to describe fluid flow in fractal media. Comparing between the fractional result in Navier-Stokes equation in fractal media [39] and local fractional one, the former via fractional calculus is continuous and differential quantities as classical result, however, the latter is local fractional continuous and nondifferential quantities. The classical result is obtained in case of fractal space-time dimension, which is equal to 1.

References

- [1] C. Foias, O. Manley, R. Rosa, and R. Temam, *Navier-Stokes Equations and Turbulence*, vol. 83, Cambridge University Press, Cambridge, UK, 2001.
- [2] G. P. Galdi, *An Introduction to the Mathematical Theory of the Navier-Stokes Equations*, vol. 1-2, Springer, New York, NY, USA, 1994.
- [3] C. Y. Wang, "Exact solutions of the steady-state Navier-Stokes equations," in *Annual Review of Fluid Mechanics*, vol. 23, pp. 159–177, Annual Reviews, Palo Alto, CA, USA, 1991.
- [4] B. Dubrulle, J.-P. Laval, S. Nazarenko, and O. Zaboronski, "A model for rapid stochastic distortions of small-scale turbulence," *Journal of Fluid Mechanics*, vol. 520, pp. 1–21, 2004.
- [5] A. Alexakis, P. D. Mininni, and A. Pouquet, "Imprint of large-scale flows on turbulence," *Physical Review Letters*, vol. 95, no. 26, Article ID 264503, 2005.

- [6] P. D. Mininni, A. Alexakis, and A. Pouquet, "Large-scale flow effects, energy transfer, and self-similarity on turbulence," *Physical Review E*, vol. 74, no. 1, Article ID 016303, 2006.
- [7] R. Benzi, L. Biferale, S. Ciliberto, M. V. Struglia, and R. Tripiccione, "Scaling property of turbulent flows," *Physical Review E*, vol. 53, no. 4, pp. R3025–R3027, 1996.
- [8] Z.-S. She and E. Leveque, "Universal scaling laws in fully developed turbulence," *Physical Review Letters*, vol. 72, no. 3, pp. 336–339, 1994.
- [9] B. I. Shraiman and E. D. Siggia, "Scalar turbulence," *Nature*, vol. 405, no. 6787, pp. 639–646, 2000.
- [10] R. Benzi, S. Ciliberto, R. Tripiccione, C. Baudet, F. Massaioli, and S. Succi, "Extended self-similarity in turbulent flows," *Physical Review E*, vol. 48, no. 1, pp. R29–R32, 1993.
- [11] C. Meneveau and J. Katz, "Scale-invariance and turbulence models for large-eddy simulation," in *Annual Review of Fluid Mechanics*, vol. 32, pp. 1–32, Annual Reviews, Palo Alto, CA, USA, 2000.
- [12] P. Constantin, C. Foias, O. P. Manley, and R. Temam, "Determining modes and fractal dimension of turbulent flows," *Journal of Fluid Mechanics*, vol. 150, pp. 427–440, 1985.
- [13] V. V. Chepyzhov and A. A. Ilyin, "On the fractal dimension of invariant sets: applications to Navier-Stokes equations," *Discrete and Continuous Dynamical Systems A*, vol. 10, no. 1-2, pp. 117–135, 2004.
- [14] V. Scheffer, "Turbulence and Hausdorff dimension, in Turbulence and the Navier-Stokes equations," in *Lecture Notes in Mathematics*, vol. 565, pp. 94–112, Springer, Berlin, Germany, 1976.
- [15] V. Scheffer, "Hausdorff measure and the Navier-Stokes equations," *Communications in Mathematical Physics*, vol. 55, no. 2, pp. 97–112, 1977.
- [16] G. Boffetta, A. Mazzino, and A. Vulpiani, "Twenty-five years of multifractals in fully developed turbulence: a tribute to Giovanni Paladin," *Journal of Physics A*, vol. 41, 363001, 2008.
- [17] A. Cheskidov, D. D. Holm, E. Olson, and E. S. Titi, "On a Leray- α model of turbulence," *Proceedings of The Royal Society of London A*, vol. 461, no. 2055, pp. 629–649, 2005.
- [18] A. A. Ilyin, E. M. Lunasin, and E. S. Titi, "A modified-Leray- α subgrid scale model of turbulence," *Nonlinearity*, vol. 19, no. 4, pp. 879–897, 2006.
- [19] D. D. Holm, "Kármán-Howarth theorem for the Lagrangian-averaged Navier-Stokes-alpha model of turbulence," *Journal of Fluid Mechanics*, vol. 467, pp. 205–214, 2002.
- [20] R. Hilfer, *Applications of fractional calculus in physics*, World Scientific Publishing, River Edge, NJ, USA, 2000.
- [21] J. Sabatier, O. P. Agrawal, and J. A. T. Machado, *Advances in Fractional Calculus: Theoretical Developments and Applications in Physics and Engineering*, Springer, New York, NY, USA, 2007.
- [22] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, *Theory and Applications of Fractional Differential Equations*, vol. 204, Elsevier Science, Amsterdam, The Netherlands, 2006.
- [23] I. Podlubny, *Fractional Differential Equations*, vol. 198, Academic Press, New York, NY, USA, 1999.
- [24] M. El-Shahed and A. Salem, "On the generalized Navier-Stokes equations," *Applied Mathematics and Computation*, vol. 156, no. 1, pp. 287–293, 2004.
- [25] S. Momani and Z. Odibat, "Analytical solution of a time-fractional Navier-Stokes equation by Adomian decomposition method," *Applied Mathematics and Computation*, vol. 177, no. 2, pp. 488–494, 2006.
- [26] Z. Z. Ganji, D. D. Ganji, A. D. Ganji, and M. Rostamian, "Analytical solution of time-fractional Navier-Stokes equation in polar coordinate by homotopy perturbation method," *Numerical Methods for Partial Differential Equations*, vol. 26, no. 1, pp. 117–124, 2010.
- [27] X. J. Yang, *Advanced Local Fractional Calculus and Its Applications*, World Science Publisher, New York, NY, USA, 2012.
- [28] X. J. Yang, D. Baleanu, and J. A. T. Machado, "Mathematical aspects of Heisenberg uncertainty principle within local fractional Fourier analysis," *Boundary Value Problems*, no. 1, pp. 131–150, 2013.
- [29] X. J. Yang and D. Baleanu, "Fractal heat conduction problem solved by local fractional variation iteration method," *Thermal Science*, vol. 17, no. 2, pp. 625–628, 2013.
- [30] M.-S. Hu, R. P. Agarwal, and X.-J. Yang, "Local fractional Fourier series with application to wave equation in fractal vibrating string," *Abstract and Applied Analysis*, vol. 2012, Article ID 567401, 15 pages, 2012.
- [31] X. J. Yang, D. Baleanu, and W. P. Zhong, "Approximation solution to diffusion equation on Cantor time-space," *Proceedings of the Romanian Academy A*, vol. 14, no. 2, pp. 127–133, 2013.
- [32] X. J. Yang, *Local Fractional Functional Analysis and Its Applications*, Asian Academic publisher Limited, Hong Kong, China, 2011.
- [33] X. J. Yang, "Local fractional integral transforms," *Progress in Nonlinear Science*, vol. 4, pp. 1–225, 2011.
- [34] A. Carpinteri and A. Sapora, "Diffusion problems in fractal media defined on Cantor sets," *ZAMM Zeitschrift fur Angewandte Mathematik und Mechanik*, vol. 90, no. 3, pp. 203–210, 2010.
- [35] K. M. Kolwankar and A. D. Gangal, "Local fractional Fokker-Planck equation," *Physical Review Letters*, vol. 80, no. 2, pp. 214–217, 1998.
- [36] J. Li and M. Ostoja-Starzewski, "Micropolar continuum mechanics of fractal media," *International Journal of Engineering Science*, vol. 49, no. 12, pp. 1302–1310, 2011.
- [37] V. E. Tarasov, "Continuous medium model for fractal media," *Physics Letters, Section A*, vol. 336, no. 2-3, pp. 167–174, 2005.
- [38] C. S. Drapaca and S. Sivaloganathan, "A fractional model of continuum mechanics," *Journal of Elasticity*, vol. 107, no. 2, pp. 105–123, 2012.
- [39] V. E. Tarasov, "Fractional hydrodynamic equations for fractal media," *Annals of Physics*, vol. 318, no. 2, pp. 286–307, 2005.

Research Article

The Stochastic Dynamics for Ecological Tourism System with Visitor Educational Intervention

Dongping Wei,^{1,2} Shouwen Wen,³ Yizeng Chen,¹ Yong Xu,⁴ and Bing Liang²

¹ School of Management, Shanghai University, Shanghai 200444, Shanghai, China

² Math and Physics Department, Shenzhen Polytechnic, Shenzhen 518055, Guangdong, China

³ School of Management, Shenzhen Polytechnic, Shenzhen 518055, Guangdong, China

⁴ School of Mathematical Sciences and Computing Technology, Central South University, Changsha 410075, Hunan, China

Correspondence should be addressed to Yizeng Chen; shucyz@163.com

Received 28 March 2013; Accepted 16 April 2013

Academic Editor: J. A. Tenreiro Machado

Copyright © 2013 Dongping Wei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The ever-increasing visitation in parks and protected areas continues to present a considerable challenge for worldwide land managers with allowing recreational use while preserving natural conditions. In China, the fast expanding visitation in protected areas is quickly damaging the natural resources and precious culture without effective visitor education, while regulation and site management are also gaining very limited efficacy. We propose a differential equation to describe the ecological tourism system. Shown by the theoretical proof and numerical simulation, the ecological tourism system is unstable without any perturbed factors, especially visitor educational intervention, because the solution of the dynamic system explodes in a finite time given any initial value. Supposing that the intrinsic increasing rate of stakeholders in the systems stochastically perturbed by the visitor educational intervention, we discover that the stochastic dynamic model can effectively suppress the explosion of the solution. As such, we demonstrate that the tourism system can develop steadily and safely even under a large amount of visitors in public vacation, when employing continuous visitor education intervention programmes.

1. Introduction

Currently, the Chinese tourism market is the primary domestic tourism market owning about 1.7 billion visitors per year and the fourth inbound tourism market and the primary outbound tourism market in the Asian-Pacific region owning about 45.84 million visitors per year [1]. Many scholars proclaim in surprise that the era of recreation is approaching ordinary Chinese daily life. In the public holidays such as golden week of Labor's day and National day, it is crowded and chaotic in every popular scene site such as Huangshan Mountain, Huashan Mountain, and the Great Wall. In these crowded parks or natural protected areas, the negative impacts, such as litter, tree damage, noise, and the rude behaviors of visitors, shock the environmentalists. In the United States, scholars concluded that the undesirable impacts on the tourism destination with fast expanding visitation include resource degradation and social impacts [2, 3]. The resource impacts associated with large amount

of visitors include the trampled vegetation and soil damage, soil compaction and erosion, litter, human waste, and wildlife disturbance. The social impacts include the diminished satisfaction of visitors and recreation conflicts between visitors, visitors and park managers, visitors and neighbors of tourism destinations, and the culture conflicts [4]. Ferreira and Rosso [5] discovered that the Sao Paulo coast is highly depredated by the increased tourist activity. Manning and Aderson [6] also found that the dramatically grown visitors' use of Arches National Park has had several important impacts on the park including trampling of fragile soils and vegetation (impacts on soil and vegetation) and crowding on trails and at attraction sites (crowding, impacts on trails and attraction sites). Left unmanaged, these impacts can lead to unacceptable changes in resource and social conditions such as the loss of sensitive or rare plants and animals, or declines in visitor satisfaction [7]. In China, many visitors feel awful when they arrive the tourism destination because of crowded people and the rude behaviors.

Facing the increasingly serious negative tourism impacts, many scholars propose different strategies to reduce the impacts. Ferreira and Rosso [5] proposed that the management strategies should include isolation of sensitive areas, construction of boardwalks, visitor education, and monitoring programmes to reduce the degradation of rocky shore fauna on the Sao Paulo coast caused by the development of tourism. In United States, many park managers commonly employ regulations, site management, and visitor education to address the resource and social impacts. The regulations directly curtail tourism freedom and antagonize visitors by their enforcement. But the enforcements of regulations are often costly and lack efficacy. Peterson and Lime [8], Hendee and Dawson [3], and Stankey and Schreyer [9] noted that site management actions were less direct but they were also costly and permanently altered the natural setting and the nature of recreational experiences. The visitor education is a more appropriate, light-handed and indirect management response to reduce resource and social impacts [4, 10]. Docutte and Cole [11] presented that visitor education has been widely touted as the most appropriate approach to managing recreation in wilderness. Visitor education is also easier to win the support of visitors [3]. Hence, in the United States, the federal land management agencies employed a Leave No Trace (LNT) programme for visitor education in order to reduce the resource and social impacts. And many other international organizations such as Members of the World Tourism Organization developed a Global Code of Ethics for Tourism in 1999 to help minimize the negative impact of tourism on the environment and the cultural heritage (WTO, 2006). The Ecotourism society (1993) has also put forward “Ecotourism Guidelines for Nature Operators” to guide the visitor educational efforts to minimize the social and environment impacts. Hendee and Dawson [3] concluded that effective educational programmes can (1) reduce per capita and cumulative resource and experiential impacts; (2) aid in keeping resource and social conditions within acceptable limits; (3) reduce the need for site management on regulatory actions such as limits; (4) provide a key component of sustainable tourism programmes.

There are a lot of references to study the visitor educational message content, message delivery approaches, audience characteristics and theoretical grounding including moral development theory and decisionmaking theory, reasoned actions and planned behavior theory. However, there is little research to study the relationship between the stakeholders and the human ecological system of tourist destinations. In fact, the stakeholders in ecotourism destinations cannot survive without the support from one another, and they must unite to a population in a certain space [12]. In 2012, Shou-wen started to analyze the dynamic property of visitors and the ecological tourism system under the visitor educational intervention. Shou-wen et al. [13] set up the stochastic increasing model of stakeholders perturbed by the continuous visitor education. In this paper, we make efforts to the further study to better understand the relationship between the stakeholders and ecological tourism system intervened by visitor educational programmes. Based on stochastic differential equation theory [14], we explore

whether the ecological tourism system can develop sustainably when perturbed by the continual visitor educational intervention.

In order to clearly present the mathematical models in the next three sections, we introduce the key definitions and notations of variables in the following.

- $x(t)$: stakeholders in an ecological tourism system at time t .
- x_0 : the initial value of differential equation at time t_0 .
- α : the intrinsic increasing rate of stakeholders within the ecological tourism system.
- β : the immigration rate of stakeholders from outside the system.
- γ : the emigration rate of stakeholders from the inner system.
- K : the maximum carrying capacity of the ecological tourism system.
- $w(t)$: the Brownian movement.
- ε : the intensity of visitor education intervention.
- τ_e : the explosion time of stochastic differential equation.
- τ_k : the stopping time of stochastic differential equation.
- k_0 : a sufficiently large value.
- k : any nonnegative integer.
- $V(x)$: a C^2 -function, $V : R^+ \rightarrow R^+$.
- M : a sufficient large positive value.
- T : a sufficient large positive constant.

2. The Logical Differential Equation Model of Ecological Tourism System

In the literature, the ecological tourism system is defined as all the natural resources including air, rocks, mountains, soils, trees, wildlife, and all the stakeholders in visitation destination. And the stakeholders [12] in ecotourism include the seven factors: (1) the tourism resource suppliers' group such as tourism resource managers or park managers; (2) the visitation product sellers group including tourism company, and tourism website managers; (3) the group of relative products or services suppliers who are providing transportation and hotels, catering, other tourism memories products; (4) the group of local communities including the staff of tourism companies and workers in other company and local people; (5) the group of governments including super government organizations and central and local governments; (6) the special interest group including nongovernment organizations like environmental organizations, human rights and labor rights organizations, or trust and environmental charity institutes, academics, and social medias; and (7) the core tourism products and services consumers like visitors.

We supposed that there are $x(t)$ stakeholders in an ecological tourism system at time t , considering that there are some stakeholders entering the system from outside the ecological

system and some others leave the system when they finish their activities or business. We denote α as the intrinsic increasing rate of stakeholders within the ecological tourism system, β as the immigration rate of stakeholders from outside the system, and γ as the emigration rate of stakeholders from the inner system. K stands for the maximum carrying capacity of the ecological tourism system. And then $1/K$ represents the recreation impact factors per capita. So $x(t)/K$ means the recreation impact factor at time t . So mapping the logistical differential equation on the ecological system, we have

$$\frac{dx}{dt} = x \left(\alpha + \beta - \gamma - \frac{1}{K} x \right). \quad (1)$$

It is the original differential equation to describe the relationship between the stakeholders and the ecological system.

Theorem 1. *For any given initial value x_0 , the solution of (1) $x(t)$ cannot avoid an explosion in a finite time.*

Proof. Since the variable $x(t)$ denotes the population of stakeholder, $x(t)$ must be a positive value at time $t \geq 0$. When the parameters of (1) $\alpha, \beta, \gamma \in (0, 1)$, $K \in R^+$, (1) only has a local solution as follows

$$x(t) = \frac{\alpha + \beta - \gamma}{-(1/K) + e^{-(\alpha+\beta-\gamma)t} (\alpha + \beta - \gamma - (1/K)x_0) / x_0}. \quad (2)$$

Given that $x_0 > 0$, it is not difficult to conclude that $x(t)$ explodes to infinity when $t \rightarrow \mathcal{T}$,

$$\mathcal{T} = -\frac{1}{\alpha + \beta - \gamma} \log \frac{-(1/K)x_0}{(\alpha + \beta - \gamma - (1/K)x_0)}. \quad (3)$$

Hence, the solution of (1) $x(t)$ will explode to infinity in a finite time. \square

In other words, the original differential equation model just has the local solution. The simulation data in Table 1 and Figure 1 show that the dynamic model (1) explodes in a finite time for any initial value $x_0 = 30, 40, 50, 70, 80, 100$, when $\alpha = 0.25$, $\beta = 0.12$, $\gamma = 0.08$, and $K = 1000$. For instance, according to simulation data in Table 1 and Figure 1, we can easily find that the number of stakeholders $x(10) = 1081$ is a little larger than the maximum capacity of the system $K = 1000$, but at time $t = 11$, $x(11) = 2990$ is much larger than $K = 1000$, when the initial value of stakeholder $x_0 = 40$ without any control on the system. Hence, the system is chaotic and unsustainable after $t = 10$. Eventually, the system is collapsed at time $t = 12$, because $x(12) = -5679 < 0$. When $x_0 = 30, 40, 50, 70, 80, 100$, the system is collapsed at time $t = 14, 12, 10, 7, 6, 4$, see Figure 1; therefore, it can be concluded that the system is collapsed faster with larger initial value x_0 .

Similarly, Figure 2 demonstrates that the explosion of the dynamic model (1) cannot avoid any initial value $x_0 = 30, 40, 45, 60, 70, 80$, given that $\alpha = 0.15$, $\beta = 0.08$, $\gamma = 0.06$, and $K = 1000$. For instance, Table 2 shows that the number of stakeholders $x(16) = 1698$ is much larger than the maximum

TABLE 1: The simulation of $x(t)$ given the different initial value x_0 .

t	$x(t)$						
	$\alpha = 0.25, \beta = 0.12, \gamma = 0.08, K = 1000$	$x_0 = 30$	$x_0 = 40$	$x_0 = 50$	$x_0 = 70$	$x_0 = 80$	$x_0 = 100$
1		53	79	112	215	301	688
2		63	95	138	280	415	1262
3		75	116	172	382	617	4545
4		91	143	219	555	1069	-3635
5		110	179	287	916	2907	-1422
6		135	229	392	2087	-5956	-931
7		169	302	575	-19651	-1638	-717
8		215	417	961	-1963	-1006	-598
9		281	622	2303	-1104	-755	-523
10		382	1081	-11137	-800	-621	-472
11		556	2990	-1841	-647	-538	-435
12		918	-5679	-1069	-555	-482	-408
13		2098	-1619	-785	-494	-443	-386
14		-18836	-1000	-638	-451	-413	-370
15		-1956	-752	-549	-420	-391	-357
16		-1102	-619	-490	-396	-373	-346
17		-800	-537	-448	-377	-359	-337
18		-646	-481	-417	-362	-348	-330
19		-554	-442	-394	-351	-339	-324

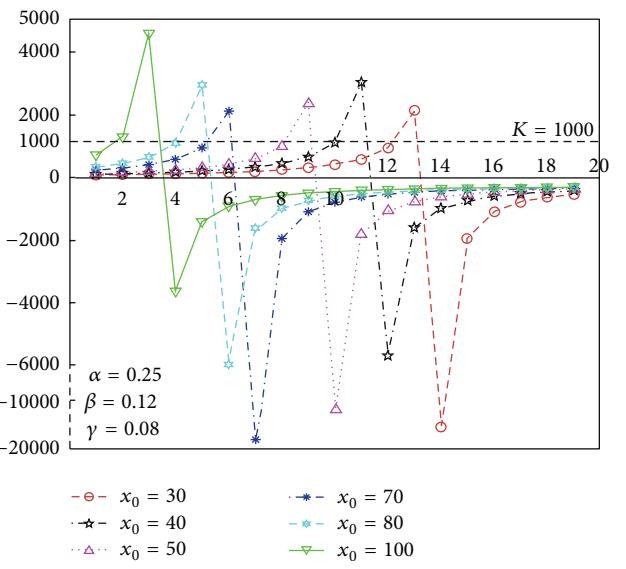


FIGURE 1: The simulation of $x(t)$ given the different initial values of x_0 .

capacity of $K = 1000$, given the initial value of stakeholder $x_0 = 30$ without any control on the system. Hence, the system is chaotic and unsustainable after time $t = 16$. Eventually, the system will be collapsed at time $t = 18$. When $x_0 = 30, 40, 45, 60, 70, 80$, the system is collapsed at time $t = 18, 13, 12, 10, 6, 3$, see Figure 1. Therefore, it can be concluded that the system is collapsed faster with larger initial value x_0 .

TABLE 2: The simulation of $x(t)$ given the different initial value x_0 and α , β , and γ .

t	$x(t)$						
	$\alpha = 0.15, \beta = 0.08, \gamma = 0.06, K = 1000$	$x_0 = 30$	$x_0 = 40$	$x_0 = 45$	$x_0 = 60$	$x_0 = 70$	$x_0 = 80$
1	58	98	127	166	311	828	
2	65	112	148	198	404	1588	
3	73	129	174	240	558	10110	
4	83	151	208	299	856	-2574	
5	94	179	254	385	1685	-1196	
6	108	214	319	525	15181	-802	
7	125	263	417	787	-2388	-615	
8	145	332	581	1456	-1157	-507	
9	171	437	908	6590	-786	-437	
10	204	616	1878	-2942	-607	-387	
11	249	989	102870	-1263	-502	-351	
12	312	2224	-2126	-829	-433	-323	
13	405	-15070	-1097	-630	-384	-301	
14	559	-1851	-760	-516	-349	-283	
15	860	-1025	-592	-443	-321	-268	
16	1698	-727	-493	-392	-299	-256	
17	16191	-574	-427	-354	-282	-246	
18	-2366	-481	-380	-325	-268	-237	
19	-1153	-418	-345	-303	-256	-230	

Comparing with Figure 1, Figure 2 indicates that the system is collapsed slower with smaller parameters α , β , and γ .

The visitors x_0 are very large in some protected areas especially on the public vacations in China. So the solution of the differential dynamic equation will inevitably explode without any interventional programme. The tourism system of some famous protected areas or precious fragile natural history park is currently in the great danger of large visitations.

3. The Stochastically Dynamics Model Perturbed by Visitor Educational Intervention

As demonstrated in the simulations of Section 2, with the high and disordered consumption, the ecological tourism system especially precious fragile protected areas or parks will eventually collapse without any visitor education or other visitor interventions. The booming China tourism market is extremely crowded in many famous tourism destinations on the public vacations. Although there are a lot of visitor regulations and rules in the tourism destinations, the lack of the effective visitors educational intervention caused the natural resource and culture resource to be badly degraded. The natural resource impacts and social impacts are extremely serious. So it is critical importance to develop effective and systematical visitor educational programme such as the Chinese version of LNT Programme to avoid systematical

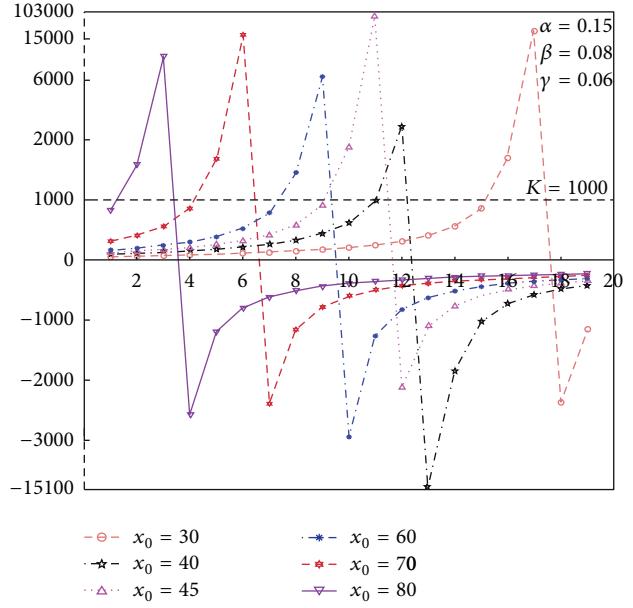


FIGURE 2: The simulation of $x(t)$ given the different initial values of x_0 and α , β , and γ .

explosion of the ecological tourism system. Increasing tourist activity in protected regions demands management strategies including visitor education to reduce the impacts on the ecological tourism system and avoid the explosion of solution of the dynamic equation describing the system. In this paper, we suppose that the intrinsic increasing rate of stakeholders is intervened by visitor education, but the efficacy of visitor education is different from diverse visitor characteristics, the educational message content, message delivery approaches, and so forth. Hence, the intrinsic increasing rate of stakeholders behaves stochastically when perturbed by visitor intervention. Naturally, we can suppose that the intrinsic rate of increase is stochastically perturbed with

$$\alpha \longrightarrow \alpha + \varepsilon \omega(t), \quad (4)$$

where $\omega(t)$ and ε represent the Brownian movement and the intensity of visitor education, respectively. Substituting (4) into the original differential equation (1), we can obtain the stochastic dynamic model of ecological tourism system perturbed by visitor education intervention programme,

$$dx = x \left(\alpha + \beta - \gamma - \frac{1}{K}x \right) dt + \varepsilon x^2 d\omega(t). \quad (5)$$

The critical purpose of the paper is to study whether the stochastic parameter of intrinsic increasing rate of stakeholders can effectively suppress the explosion of the solution of (2) in a finite time with any initial value x_0 and stop the collapse of the ecological tourism system eventually. In this paper, we should prove that the solution of (5) can no longer explode in a finite time with a probability one when $\alpha + \beta < \gamma + (1/K)x_0$. And we will then see that for arbitrary parameters $\alpha, \beta, \gamma \in [0, 1]$ and $K \in R^+$, the dynamic system will not explode in a finite time with a probability one provided any intensity of

the perturbed visitor educational intervention ε . The results demonstrate that effective visitor education programmes will support the sustainable development of ecological tourism system and suppress the explosion of the stochastic dynamic model.

4. The Existence of Nonexplosion Solutions of the Ecological Tourism Dynamic System with Visitor Education of Intervention

In the paper, we denote $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathcal{P})$ to be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ satisfying the usual condition (i.e., right continuous, and \mathcal{F}_0 contains all p -null sets) [14]. Let $\omega(t)$ be one-dimensional Brownian motion defined on the probability space.

The definition $x(t)$ should be nonnegative because it represents the stakeholders. For any given initial value x_0 , if the stochastic differential equation (5) has a unique global solution, then the coefficient of the equation should be satisfy the linear growth condition and local Lipschitz condition [14–17]. Unfortunately, the coefficients of (5) only satisfy the locally Lipschitz continuous condition, so the solution of (5) only has local solution which may explode at a finite time [18].

Now, we will show that the visitor education intervention can suppress the explosion of the solution as proved by the following theorem.

Theorem 2. Given any system parameters $\alpha, \beta, \gamma \in [0, 1]$, $K \in R^+$ and any initial value $x_0 \in R^+$ if intensity of visitor education intervention $\varepsilon \neq 0$, there is a unique solution $x(t)$ of (3) at $t \geq 0$, and the solution will be positive with probability one, namely, $x(t) > 0$ when $t \geq 0$ a.s.

Proof. For the coefficients of (5) are the locally Lipschitz continuous, for any given initial value $x_0 \in R^+$, there is a unique solution $x(t)$ on the $t \in (0, \tau_e)$ where τ_e is the explosion time [18, 19]. To prove the solution of (5) $x(t)$ is global and positive, we need to show that $\tau_e = \infty$ almost surely. Let $k_0 > 0$ be sufficiently large value and the initial value of the equation x_0 belong to the interval $[1/k_0, k_0]$. For any nonnegative integer $k \geq k_0$, we define the stopping time as follows

$$\tau_k = \inf \left\{ t \in [0, \tau_e) : x(t) \notin \left(\frac{1}{k}, k \right) \right\}, \quad (6)$$

where setting $\inf \emptyset = \infty$ throughout the whole paper, as usual \emptyset denotes the empty set. Obviously, τ_k is increasing when $k \rightarrow \infty$. Set $\tau_\infty = \lim_{k \rightarrow \infty} \tau_k$ hence $\tau_\infty \leq \tau_e$ a.s. If it can be proved that $\tau_\infty = \infty$ a.s., then it is natural that $\tau_e = \infty$ a.s. and $x(t) \in R^+$ for any $t \geq 0$. In other words, in order to show that $x(t) \in R^+$ will not explode, we should show that $\tau_\infty = \infty$ a.s. first. Employing contradiction methodology, if this statement $\tau_\infty = \infty$ a.s. is false, then there is a pair of constant $T \geq 0$ and $\varepsilon \in (0, 1)$, such that

$$P \{ \tau_\infty \leq T \} > \varepsilon. \quad (7)$$

Therefore, an integer $k_1 \geq k_0$ can be found, such that

$$P \{ \tau_\infty \leq T \} > \varepsilon \quad \forall k \geq k_1. \quad (8)$$

Define a C^2 -function $V: R^+ \rightarrow R^+$,

$$V(x) = \sqrt{x} - 1 - \log x^{0.5}. \quad (9)$$

It is easily proved that C^2 -function $V(x)$ is nonnegative on $x > 0$.

If $x(t) \in R^+$, it can be calculated from the Ito formula as follows

$$\begin{aligned} dV(x) &= \left\{ 0.5 \left(x^{-0.5} - x^{-1} \right) x \left[\left(\alpha + \beta - \gamma - \frac{1}{K} x \right) dt + \varepsilon x d\omega(t) \right] \right\} \\ &\quad + 0.5 \left(0.5x^{-2} - 0.25x^{-1.5} \right) \varepsilon^2 x^4 dt \\ &= \left\{ 0.5 \left(x^{0.5} - 1 \right) \left[\left(\alpha + \beta - \gamma - \frac{1}{K} x \right) \right. \right. \\ &\quad \left. \left. + \left(0.25 - 0.125x^{0.5} \right) \varepsilon^2 x^2 \right] \right\} dt \\ &\quad + 0.5 \left(x^{0.5} - 1 \right) \varepsilon x d\omega(t), \end{aligned} \quad (10)$$

where we denote $x(t) = x$ for simplification.

Then (10) can be simplified as follows

$$\begin{aligned} &0.5 \left(x^{0.5} - 1 \right) \left(\alpha + \beta - \gamma - \frac{1}{K} x \right) \\ &= 0.5 (\alpha + \beta - \gamma) x^{0.5} - 0.5 \frac{1}{K} x^{1.5} \\ &\quad - 0.5 (\alpha + \beta - \gamma) + 0.5 \frac{1}{K} x \\ &= -0.5 \frac{1}{K} x^{1.5} + 0.5 \frac{1}{K} x \\ &\quad + 0.5 (\alpha + \beta - \gamma) x^{0.5} - 0.5 (\alpha + \beta - \gamma), \\ &(0.25 - 0.125x^{0.5}) \varepsilon^2 x^2 = 0.25\varepsilon^2 x^2 - 0.125\varepsilon^2 x^{2.5}. \end{aligned} \quad (11)$$

We note that

$$\begin{aligned} F(x) &= -0.125\varepsilon^2 x^{2.5} + \frac{1}{4} \varepsilon^2 x^2 - 0.5 \frac{1}{K} x^{1.5} \\ &\quad + 0.5 \frac{1}{K} x + 0.5 (\alpha + \beta - \gamma) x^{0.5} - 0.5 (\alpha + \beta - \gamma) \\ &\leq -0.125\varepsilon^2 x^{2.5} + \frac{1}{4} \varepsilon^2 x^2 + 0.5 \frac{1}{K} x^{1.5} \\ &\quad + 0.5 \frac{1}{K} x + 0.5 (\alpha + \beta - \gamma) x^{0.5} + 0.5 (\alpha + \beta - \gamma). \end{aligned} \quad (12)$$

Then we have $dV(x) = F(x)dt + 0.5(x^{0.5} - 1)\varepsilon x d\omega(t)$.

If $\varepsilon \neq 0$, then it is not difficult to yield that $F(x)$ is bounded by sufficiently large positive value M , so that $dV(x) \leq M + 0.5(x^{1.5} - x)\varepsilon d\omega^T(t)$.

It is clear to obtain the following expression by integrating the both sides of the proceeding inequality from 0 to $T \wedge \tau_k$ as follows

$$\int_0^{T \wedge \tau_k} dV(x) \leq \int_0^{T \wedge \tau_k} Mdt + \int_0^{T \wedge \tau_k} 0.5(x^{0.5} - 1)\varepsilon xd\omega(t). \quad (13)$$

With $(T \wedge \tau_k) \in R^+$, taking expectations on both sides of inequality (8), we can yield

$$EV(x(T \wedge \tau_k)) \leq V(x_0) + ME(T \wedge \tau_k) \leq V(x_0) + MT. \quad (14)$$

Let $\Omega_{\varepsilon} = \{\tau_{\varepsilon} \leq T\}$ for $\varepsilon \geq k$ and the inequality (7) so that $P(\Omega_{\varepsilon}) > \varepsilon$. Taking any $w \in \Omega_{\varepsilon}$, with the definition of stopping time (6), $x(\tau_k, w)$ equals either $1/k$ or k , and then it is easy to conclude that $V(x(\tau_k, w))$ is not less than either $\sqrt{k} - 1 - 0.5 \log k$ or $\sqrt{1/k} - 1 - 0.5 \log(1/k)$. So $x(t) \geq k$ or $x(t) \leq 1/k$ when $t \geq \tau_k$. Therefore, it is not difficult to get the following inequality:

$$\begin{aligned} V(x(\tau_k, w)) \\ \geq [\sqrt{k} - 1 - 0.5 \log k] \wedge \sqrt{\frac{1}{k}} - 1 - 0.5 \log\left(\frac{1}{k}\right). \end{aligned} \quad (15)$$

Taking expectation on both sides of (15), we get

$$\begin{aligned} EV(x(\tau_k, w)) \\ \geq [\sqrt{k} - 1 - 0.5 \log k] \wedge \sqrt{\frac{1}{k}} - 1 - 0.5 \log\left(\frac{1}{k}\right). \end{aligned} \quad (16)$$

It is easy to show that

$$\begin{aligned} V(x_0) + MT \\ \geq E[1_{\Omega_{\varepsilon}(w)}V(x(T \wedge \tau_k))] \\ \geq \varepsilon[\sqrt{k} - 1 - 0.5 \log k] \wedge \sqrt{\frac{1}{k}} - 1 - 0.5 \log\left(\frac{1}{k}\right), \end{aligned} \quad (17)$$

where $1_{\Omega_{\varepsilon}(w)}$ is the indicator function of Ω_{ε} . Letting $k \rightarrow \infty$ in (17), then

$$\infty > V(x_0) + MT = \infty, \quad (18)$$

where T is a constant denoted in (7). Obviously, (18) is contradicted. So we can conclude that the statement of (7) is false. Hence, there are constants $T > 0$ and $\varepsilon \in (0, 1)$ such that $P\{\tau_{\infty} \leq T\} > \varepsilon$. Consequently, the original statement is true in which $P\{\tau_k \leq \tau_{\infty}\} < \varepsilon$. In other words, the statement $\tau_{\infty} = \infty$ is true almost surely. \square

It is clear to derive from the Theorem 2 that there is a positive solution $x(t)$ of (7) and $x(t)$ will never explode in a finite time. Following the condition of Theorem 2, we can

deduce that the visitor educational intervention brings the stochastic intrinsic increasing rate of stakeholders into the tourism ecological system. When the intensity of intervention $\varepsilon \neq 0$, the stochastic intrinsic increasing rate can effectively suppress the explosion of the solution $x(t)$ for any given x_0 and avoid the collapse mentioned in Section 2. It means that the slight or tiny stochastical change of intrinsic increasing rate of stakeholder introduced by visitor educational intervention can keep the tourism ecological system developing safely and steadily. In fact, early in 1970s, Fishbein, Fazio [20], and Bradley [21] already noticed that the negative recreation impacts on wildlands and wildlife, and the ecological system or environment could not keep sustainable development without recreation planning and management. According to Theorem 2, the visitor educational intervention is a necessary condition for the safe and stable developing of the tourism ecological system. Especially in China, the booming recreation market is in urgent need of the visitor educational intervention to stop the chaotic tourism activities.

5. The Boundedness of the Solution of the Dynamic Model

With the intensities of the visitor educational intervention, the solution of the dynamic model (5) cannot explode in a finite time following by the Theorem 2. However, it is not enough to show that the solution exists and does not explode. We consider how the solution of the dynamic model varies in a finite interval in this section. In other words, we will show that the solution of (5) is ultimately bounded.

Denote $x(t, x_0)$ as the unique global solution of (5) for any given initial value $x_0 > 0$. We define stochastically ultimate bounded property of the solution of the dynamic model as follows.

Definition 3. For any initial value $x_0 > 0$, if the solution of (3) $x(t, x_0)$ has the property

$$\liminf_{t \rightarrow \infty} x(t, x_0) > 0, \quad (19)$$

then (5) is defined to be stochastically ultimately bounded with probability one. The following theorem shows (5) is stochastically ultimately bounded.

Theorem 4. If the coefficient of (5) satisfies the following inequality:

$$\begin{aligned} \frac{1}{2}\varepsilon^2 - \frac{1}{K} &< 0, \\ \left(\alpha + \beta - \gamma + \frac{1}{K}\right)^2 + 4\left(\frac{1}{2}\varepsilon^2 - \frac{1}{K}\right)(\alpha + \beta - \gamma) &\leq 0. \end{aligned} \quad (20)$$

Then the solution of (5) $x(t, x_0)$ is stochastically ultimately bounded with a probability one. Given any initial value $x_0 \in R^+$, the solution of (3) has the property

$$\limsup_{t \rightarrow \infty} x(t, x_0) < \infty. \quad (21)$$

So the solution of (5) $x(t, x_0)$ belongs to a finite interval

$$\left[\liminf_{t \rightarrow \infty} x(t), \limsup_{t \rightarrow \infty} x(t) \right]. \quad (22)$$

Proof. It is necessary to define a C^2 -function $V : R^+ \rightarrow R^+$ by

$$V(x) = x - 1 - \log x. \quad (23)$$

Applying the Ito formula, we derive from (23) that

$$dV(x) = LV(x) dt + (1 - x^{-1}) x^2 \varepsilon d\omega(t), \quad (24)$$

where

$$\begin{aligned} LV(x) &= (1 - x^{-1}) x \left(\alpha + \beta - \gamma - \frac{1}{K} x \right) + \frac{1}{2} x^{-2} \varepsilon^2 x^4 \\ &= (x - 1) \left(\alpha + \beta - \gamma - \frac{1}{K} x \right) + \frac{1}{2} \varepsilon^2 x^2 \\ &= \left(\frac{1}{2} \varepsilon^2 - \frac{1}{K} \right) x^2 + \left(\alpha + \beta - \gamma + \frac{1}{K} \right) x - (\alpha + \beta - \gamma). \end{aligned} \quad (25)$$

It is obvious to see that $LV(x(t))$ is a quadratic function. Applying the property of quadratic function, when

$$\frac{1}{2} \varepsilon^2 - \frac{1}{K} < 0, \quad (26)$$

$$\Delta = \left(\alpha + \beta - \gamma + \frac{1}{K} \right)^2 + 4 \left(\frac{1}{2} \varepsilon^2 - \frac{1}{K} \right) (\alpha + \beta - \gamma) \leq 0. \quad (27)$$

If (20) is satisfied, the quadratic function $LV(x(t))$ must be less than zero. Therefore, we can get that

$$dV(x(t)) \leq 0 dt + (x^2 - x) \varepsilon d\omega(t). \quad (28)$$

Integrating both sides of inequality (28) from 0 to t , we can get

$$\int_0^t dV(x(t)) \leq \int_0^t 0 dt + \int_0^t (x^2 - x) \varepsilon d\omega(t). \quad (29)$$

Thus,

$$V(x(t)) \leq V(x_0) + \int_0^t (x^2 - x) \varepsilon d\omega(t). \quad (30)$$

For convenience, we define

$$X(t) := V(x_0) + \int_0^t (x^2 - x) \varepsilon d\omega(t). \quad (31)$$

Since $V(x)$ is a nonnegative function, applying Doob's martingale convergence theorem, we can yield

$$\lim_{t \rightarrow \infty} x(t) < \infty \quad \text{a.s.} \quad (32)$$

Thus, it is natural to yield $\lim_{t \rightarrow \infty} \sup x(t) < \infty$ a.s.

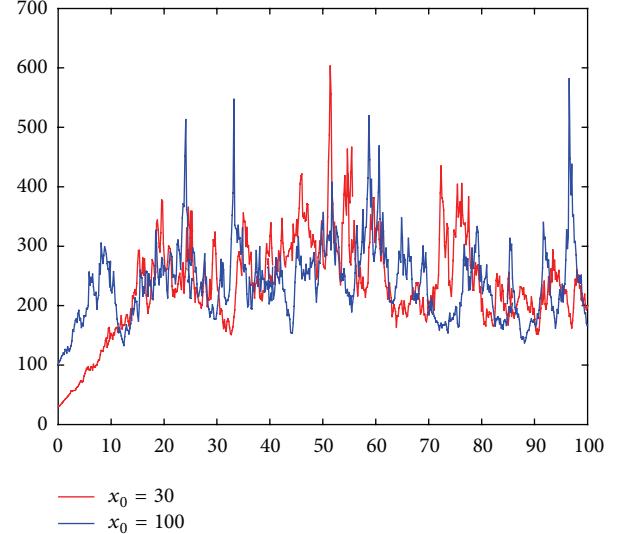


FIGURE 3: The simulation of $x(t)$ with $\alpha = 0.25$, $\beta = 0.12$, $\gamma = 0.08$, and $K = 1000$ and different initial values x_0 .

Recalling the definition of C^2 -function $V : R^+ \rightarrow R^+$, if and only if $x \rightarrow \infty$ or $x \rightarrow 0$, the limit of $V(x) = x - 1 - \log x$ is ∞ . Hence $\lim_{t \rightarrow \infty} \sup V(x(t)) = \lim_{t \rightarrow \infty} \sup(x(t)) - 1 - \log(x(t)) < \infty$ a.s.

Then it is not difficult to deduce that

$$0 < \liminf_{t \rightarrow \infty} x(t) < \limsup_{t \rightarrow \infty} x(t) < \infty \quad \text{a.s.} \quad (33)$$

In conclusion, given any initial value x_0 , the solution of (3) $x(t)$ will vary in the interval of $[\liminf_{t \rightarrow \infty} x(t), \limsup_{t \rightarrow \infty} x(t)]$. \square

From Theorem 4, the stakeholders of the ecological tourism system with visitor educational intervention can develop within an interval sustainably and steadily. Different kinds of stakeholders retained in a safety interval can immigrate from outside the system or emigrate from inside the system with regular rate. Hence, visitor education programmes are important elements of management efforts to ensure the protection of protected areas including Natural History Park and precious culture resource areas, and so forth.

6. The Simulations of the Stochastic Dynamic Model of Ecological Tourism System

Following the Euler-Maruyama method [22] and using Matlab program, we can obtain the numerical solutions and their figures of the stochastic dynamic model of ecological tourism system (5) with different parameters α , β , γ and different initial values x_0 .

Given $\alpha = 0.25$, $\beta = 0.12$, $\gamma = 0.08$, and $K = 1000$, it is easy to gain the simulation data of $x(t)$ and draw Figure 3 when $x_0 = 30$ and $x_0 = 100$. Figure 3 demonstrates that the stochastic dynamic model is varying far away from the maximum carrying capacity of the ecological system

$K = 1000$. And the peak value of $x(t)$ is 625 much smaller than $K = 1000$. Comparing Figure 1 with Figure 3, it can be concluded that the stochastic dynamic model (5) can effectively suppress the explosion and avoid the chaotic and unsustainable situation with the same parameters α, β, γ .

Given $\alpha = 0.15, \beta = 0.08, \gamma = 0.06, K = 1000$, it is not difficult to get the simulation data of $x(t)$ and draw Figure 4 when $x_0 = 30$ and $x_0 = 80$. Figure 2 shows that the original logistical differential Equation (1) collapses in a finite time. But Figure 4 shows that the stochastic dynamic model (5) does not explode with the same parameters α, β, γ . The peak value of $x(t)$ is 298 with smaller parameters α, β, γ , see Figure 4.

Shown by Figure 5, the ecological tourism system perturbed by visitor educational intervention does not collapse even with larger initial value $x_0 = 900$ which is very close to $K = 1000$. When there is a large number of visitors in the ecological tourism system, the effective visitor education programme will also keep resource and social conditions within acceptable limits by employing low impact behaviors. Figure 5 demonstrates that the system restores to normal condition quickly. Following Figures 3 to 5, it can be concluded that the ecological tourism system with visitor educational intervention can develop safely and steadily.

7. Conclusions

The ever-increasing demand for outdoor recreation in parks and protected areas worldwide continues to present a considerable challenge for managers [22]. Allowing recreational use while preserving natural conditions, it also presents a challenge for park managers in China because of fast expanding visitation. The protected areas including some natural history parks will be quickly damaged without visitor education intervention, while regulation and site management are also gaining very limited efficacy.

In Section 2, supposing that the different kinds of stakeholders of ecological tourism system grow with the specified intrinsic increasing rate and the stakeholders from outside competitors can immigrate into the system and some insider of the system can also go out freely when they finish their activities or business, we set up a differential equation to describe the ecological tourism system. Without any perturbed factors especially visitor educational intervention shown by the theoretical proof and numerical simulation, the solution of the dynamic system will explode in a finite time given any initial value.

In order to avoid the explosion of the solution of the dynamic model and the unstable situation of the ecological tourism system, we add visitor educational intervention into the system in Section 3. Supposing that the intrinsic increasing rate of stakeholders is stochastically perturbed by the visitor educational intervention, we explore a stochastic dynamic model which can effectively suppress the explosion of the solution shown in Section 4. And in Section 5, we also prove that the solution of the stochastic dynamic model will vary in a finite interval. In Section 6, following the Euler-Maruyama method, we obtain the numerical solutions and their figures of the stochastic dynamic model of ecological

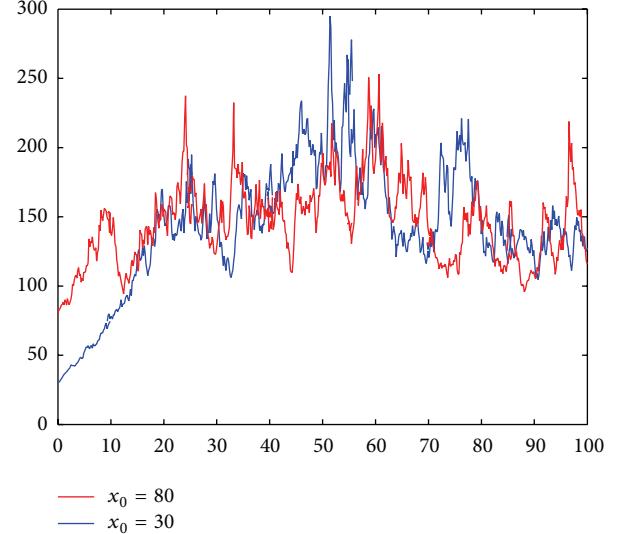


FIGURE 4: The simulation of $x(t)$ with $\alpha = 0.15, \beta = 0.08, \gamma = 0.06, K = 1000$ and different initial value x_0 .

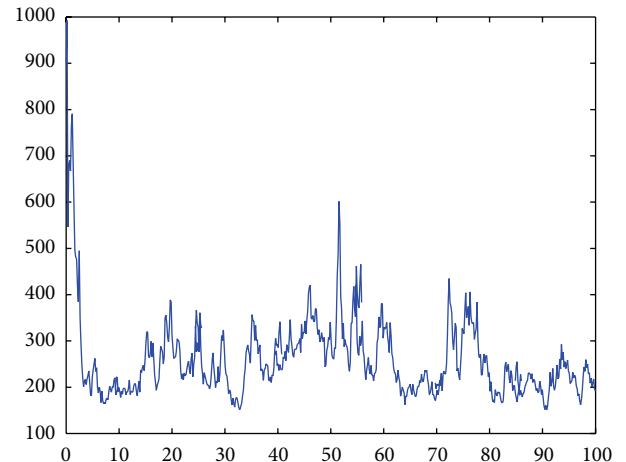


FIGURE 5: The simulation of $x(t)$ with $\alpha = 0.25, \beta = 0.12, \gamma = 0.08, K = 100$ and $x_0 = 900$.

tourism system. So we demonstrate that the tourism system will develop steadily and safely even under a large amount of visitors in public vocation, when employing continuous visitor education programmes.

Generally, the findings of the stochastic dynamic model support that the continuous visitor education intervention including ecological knowledge, minimum impact behavior or technique, and regulations message is the necessary response to the challenge of recreation impact issues which try to manage resources sustainably and provide quality visitor experiences.

Authors' Contribution

Dongping Wei and Shouwen Wen equally contributed to this work.

Acknowledgments

This work is supported by National Social Science Foundation of China (Grant no. 11BGL053), National Science Fund of China (Grant no. 71272177/G020902, Grant No. 11101434), Shenzhen Education Society Fund (Grant no. XH021). And the authors would like to thank the editor and anonymous referees for their helpful comments on the paper.

References

- [1] China National Tourism Administration, *The Yearbook of China Tourism Statistics*, Travel & Tourism Press, Beijing, China, 2012.
- [2] W. E. Hammitt and D. N. Cole, *Wildland Recreation: Ecology and Management*, John Wiley, New York, NY, USA, 2nd edition, 1998.
- [3] J. C. Hendee and C. P. Dawson, "Wilderness visitor management: stewardship for quality experiences," in *Wilderness Management: Stewardship and Protection of Resources and Values*, Fulcrum, Golden, Colo, USA, 2002.
- [4] R. E. Manning, *Studies in Outdoor Recreation: Search and Research for Satisfaction*, Oregon State University Press, Corvallis, Ore, USA, 1999.
- [5] M. N. Ferreira and S. Rosso, "Effects of human trampling on a rocky shore fauna on the Sao Paulo coast, southeastern Brazil," *Brazilian Journal of Biology*, vol. 69, no. 4, pp. 993–999, 2009.
- [6] R. E. Manning and L. E. Andrson, "how many visitors is too many at Arches?" in *Managing Outdoor Recreation: Case Studies in the National Parks*, pp. 71–75, 2012.
- [7] J. L. Marion and T. A. Farrell, "Management practices that concentrate visitor activities: camping impact management at Isle Royale National Park, USA," *Journal of Environmental Management*, vol. 66, no. 2, pp. 201–212, 2002.
- [8] G. L. Peterson and D. W. Lime, "People and their behavior: a challenge for recreation management," *Journal of Forestry*, vol. 77, no. 6, pp. 343–347, 1979.
- [9] G. H. Stankey and R. Schreyer, "Attitudes towards wilderness and factors affecting visitor behavior:A state of knowledge review," in *Proceedings of the National Wilderness Research Conference. Issues, State of Knowledge, Future. Directions*, R. C. Lucas, Ed., General Technical Report INT-220, pp. 246–293, USDA Forest Service, Intermountain Research Station, Ogden, Utah, USA, 1987.
- [10] J. L. Marion and S. E. Reid, "Minimising visitor impacts to protected areas: The efficacy of low impact education programmes," *Journal of Sustainable Tourism*, vol. 15, no. 1, pp. 5–27, 2007.
- [11] J. E. Doucette and D. N. Cole, "The wilderness visitor education: Information about alternative techniques," General Technical Report-Intermountain Research Station, USDA Forest Service No.INT-2995, 1993.
- [12] W. Shou-wen, "On the growth of stakeholders in ecotourism destination," *Journal of Wuhu Vocational Institute of Technology*, vol. 11, no. 1, pp. 1–3, 2009 (Chinese).
- [13] W. Shou-wen, X. Yong, and W. Dongping, "Stochastic ecological tourism system dynamic model with visitor educational intervention," Wroking Paper, 2012.
- [14] 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.
- [15] G. S. Ladde and V. Lakshmikantham, *Random Differential Inequalities*, vol. 150, Academic Press, New York, NY, USA, 1980.
- [16] R. Liptser and A. N. Shirayev, *Theory of Martingales*, Kluwer Academic, Dordrecht, The Netherlands, 1989, Translation of the Russian Edition, Nauka, Moscow.
- [17] X. Mao, *Stochastic Differential Equations and Applications*, Horwood, New York, NY, USA, 1997.
- [18] X. Mao, *Stochastic Differential Equations and Applications*, Harwood, Chichester, UK, 2007.
- [19] A. Friedman, *Stochastic Differential Equations and their Applications*, Academic Press, New York, NY, USA, 1976.
- [20] J. R. Fazio, *Communicating with the Wilderness User. Bulletin Number 28*, University of Idaho, College of Forestry, Wildlife and Range Experiment Station, Moscow, Idaho, USA, 1979.
- [21] J. A. Bradley, "A human approach to reducing wildland impacts. Proceedings-Recreational Impacts on Wildlands (P. 222-6)," Tech. Rep. R-6-001-1970, USDA Forest Service, Pacific Northwest Region, Portland, Ore, USA, 1979.
- [22] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin, Germany, 1999.