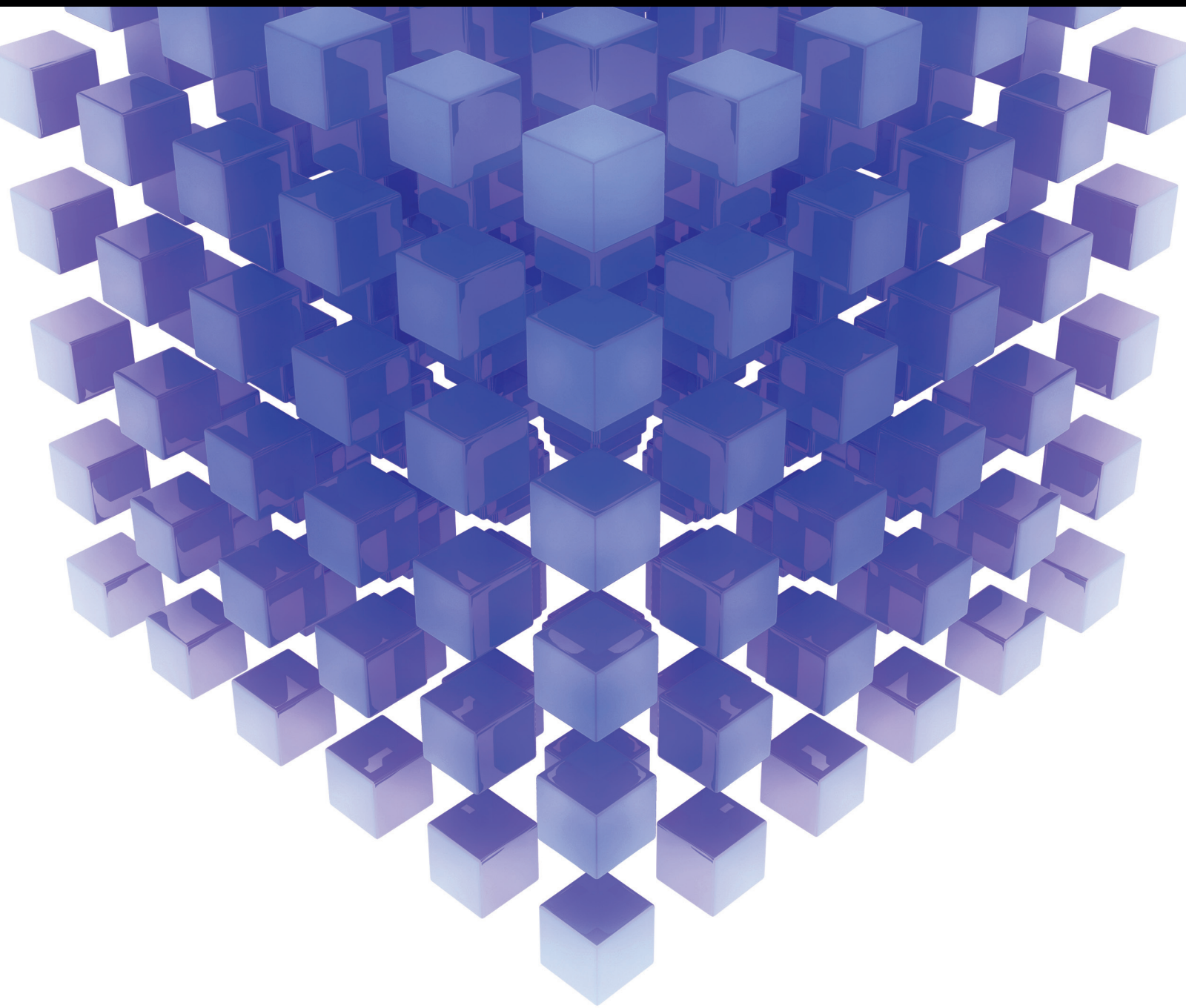
Mathematical Problems in Engineering

Control Problem of Nonlinear Systems with Applications

Guest Editors: Rongwei Guo, H. G. Enjieu Kadji, Xinguang Zhang, Uchechukwu E. Vincent, and Wenguang Yu



Control Problem of Nonlinear Systems with Applications

Mathematical Problems in Engineering

Control Problem of Nonlinear Systems with Applications

Guest Editors: Rongwei Guo, H. G. Enjieu Kadji, Xinguang Zhang, Uchechukwu E. Vincent, and Wenguang Yu

Editorial Board

Contents

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 3137609, 2 pages
http://dx.doi.org/10.1155/2016/3137609

Editorial

Control Problem of Nonlinear Systems with Applications

Rongwei Guo,[1] H. G. Enjieu Kadji,[2] Xinguang Zhang,[3] Uchechukwu E. Vincent,[4] and Wenguang Yu[5]

[1] *School of Science, Qilu University of Technology, Jinan 250353, China*
[2] *Center for Neural Engineering, The Pennsylvania State University, University Park, PA 16802, USA*
[3] *Department of Mathematics and Statistics, Curtin University of Technology, Perth, WA 6845, Australia*
[4] *Department of Physical Sciences, Redeemer's University, PMB 230, Ede, Nigeria*
[5] *School of Insurance, Shandong University of Finance and Economics, Jinan 250014, China*

Correspondence should be addressed to Rongwei Guo; rongwei_guo@163.com

Received 29 June 2016; Accepted 17 July 2016

The control problem of nonlinear systems with applications is general in the actual process and has attracted many scholars' attention owing to the wide applications in various fields such as physics, mathematics, finance, and engineering. Therefore, the analysis and synthesis of control problems play important roles in many practical systems. The aim of this special issue is to bring together the latest/innovative knowledge, analysis, and synthesis of control problems of nonlinear systems. We publish in this issue a number of state-of-the-art studies on the topic that span the control problems in economics, nonlinear control problems, nonlinear stability problems, and control problems in chaotic systems.

For control problems in economics, W. Yu and Y. Huang investigate a dependent insurance risk model with surrender and investment under the thinning process. By the martingale theory, the properties of the surplus process, adjustment coefficient equation, the upper bound of ruin probability, and explicit expression of ruin probability are obtained. X. Zhou et al. provide a practical optimal reinsurance scheme under particular conditions, with the goal of minimizing total insurer risk, and explore the optimization of limited stop-loss reinsurance under three risk measures: value at risk (VaR), tail value at risk (TVaR), and conditional tail expectation (CTE). Q. Liu et al. consider the decomposition and decoupling analysis of energy-related carbon emissions from China manufacturing. L. Xu et al. study the optimal investment and consumption for an insurer with high-watermark performance fee and the object of insurance company is to maximize the expected cumulated discount utility up to ruin time. C. Chen et al. establish a posteriori error estimate for finite volume element method of a second-order hyperbolic equation. Residual-type a posteriori error estimator is derived. The computable upper and lower bounds on the error in the H_1-norm are established. Numerical experiments are provided to illustrate the performance of the proposed estimator. H.-L. To et al. present a cascade probability control scheme using margin optimal method to address such challenges under different kinds of real-world TCP stacks. Simulation results guarantee the measured round trip time tracking to a low value of delay.

On nonlinear control problems, M.-D. Tran and H.-J. Kang present a high-performance nonsingular terminal sliding mode control method for uncertain second-order nonlinear systems. Y. Abe et al. propose a systematic numerical method for designing robust nonlinear H_∞ controllers without a priori low-dimensional approximation with respect to solutions of Hamilton-Jacobi equations. J. Yu et al. study the global asymptotic stabilization control problem for a class of nonlinear systems with input-to-state stable (ISS) dynamic uncertainties and uncertain time-varying control coefficients. Z. Wang and Z. Yan consider a three-stage tracking control for the LED wafer transporting robot and obtain a high order polynomial interpolation method to plan the motion process of the LED wafer transporting robot. D. I. R. Almeida et al. designed a strategy and implemented a robust controller for a class of underactuated mechanical systems, with two degrees of freedom, which solves the problems of regulation and trajectory tracking.

L. Li et al. apply an unsupervised learning algorithm to learn language-independent stroke feature and combine unsupervised stroke feature learning together with automatically multilayer feature extraction to improve the representational power of text feature and develop a novel nonlinear network based on traditional Convolutional Neural Network that enables detecting multilingual text regions in the images. Y.-C. Wang et al. present a unified approach to nonlinear dynamic inversion control algorithm with the parameters for desired dynamics determined by using an eigenvalue assignment method. F. A. G. Becerra et al. investigate the optimal controller and controller based on differential flatness in a linear guide system: a performance comparison of indexes. Z. Yao et al. studied the coordinated stability control of wind-thermal hybrid AC/DC power system. A coordinated control strategy for the wind-thermal hybrid AC/DC power system is proposed and an experimental prototype is made. Z.-R. Tsai and Y.-Z. Chang propose an adaptive control scheme for nonlinear systems with significant nonminimum phase dynamics. The scheme is composed of an inner-level adaptive fuzzy PD control law and an outer-level supervisory control law. C. He et al. address the development and application of novel Prognostics and Health Management (PHM) technologies to a prototype machining process (a screw tightening machine). The enabling technologies are built upon a series of tasks starting with failure analysis, testing, and data processing aimed at extracting useful features or condition indicators from raw data, a symbolic regression modeling framework, and a Bayesian estimation method called particle filtering to predict the feature state estimate accurately. H. Lou et al. propose a novel visual tracking algorithm based on multifeature selection and sparse representation in order to enhance the robustness of visual tracking algorithm in complex environment. E. Kolsi-Gdoura et al. consider the surface design as a case of virtual controller design using the back-stepping method. P. Jiang et al. study the multivariable fuzzy control based mobile robot odor source localization via semitensor product. Z. Wu et al. present an adaptive neural control for the longitudinal dynamics of a morphing aircraft. W. Sun et al. present H_∞ excitation control design problem for power systems with input time delay and disturbances by using nonlinear Hamiltonian system theory. S. ud Din et al. present a robust control design for the class of underactuated uncertain nonlinear systems. C. Gong et al. researched on submarine straight-line track control underwater base on nonlinear proportion differential.. M. C. Razali et al. studied the singularly perturbation method which is applied to multivariable PID controller design, found that the singularly perturbed system obtained by Naidu method can maintain the originality of the system characteristics, and then designed MPID controllers. It should be pointed out that the closed loop performance and process interactions were analyzed and compared to see the effectiveness of the singularly perturbed MPID control design.

On nonlinear stability problems, J. Chen investigates the generalized Degasperis-Procesi equation with variable coefficients and establishes the L_1 (R) stability of the strong solution for the equation under certain assumptions. J. A. Taborda and F. Angulo describe and prove a new method to compute and control the basins of attraction in multistability scenarios and guarantee monostability condition. In particular, the basins of attraction are computed only using a submap, and the coexistence of periodic solutions is controlled through fixed-point inducting control technique, which has been successfully used until now to stabilize unstable periodic orbits. D. Xu et al. discuss the global dynamics of a model involving an endemic equilibrium and a disease-free equilibrium, respectively. V. Nosov et al. analyze the stability of autonomous dynamical switched systems by means of multiple Lyapunov functions and give the stability theorems which have finite number of conditions to check. Y. Tian et al. address the problem of exponential stabilization of a class of time-varying delay systems with nonlinear perturbations. J. Liu and Y. Yang investigated the limiting distribution of the size of binary interval tree. M. De la Sen and A. Ibeas investigated some boundedness and convergence properties of sequences which are generated iteratively through switched mappings defined on probabilistic metric spaces as well as conditions of existence and uniqueness of fixed points. Such switching mappings are built from a set of primary self-mappings selected through switching laws. H. Xu et al. studied the L_2-gain analysis problem for a class of discrete-time switched systems with time-varying delays. A mode-dependent average dwell time (MDADT) approach is applied to analyze the L_2-gain performance for these discrete-time switched delay systems. H. Zhu et al. study the flow past a pair of cylinders in tandem at Reynolds numbers of 1000 by Domain Decomposition Method by applying a parallel computation.

Four papers on control problems in chaotic systems have been published. T. H. Cortés et al. studied the complete synchronization and the generalized synchronization problem of the discrete-time chaotic fuzzy systems by means of fuzzy output regulation using genetic algorithm. L. Ren and R. Guo investigated the synchronization and antisynchronization for a class of chaotic systems, and not only proposed a necessary and sufficient condition to synchronize and antisynchronize simultaneously the chaotic systems but also obtained two methods to realize coexistence of synchronization and antisynchronization in the chaotic systems, and give the corresponding adaptive controllers. Z. Li et al. studied chaotification problem for a class of delay difference equations by using the snap-back repeller theory and the feedback control approach. L. Huang et al. investigate the design and application in secure communication and image encryption of a new Lorenz-like system with varying parameter. L. Huang et al. proposed a new Lorenz-like chaotic system with varying parameter by adding a state feedback function.

We hope these papers will be of help to readers in furthering their exploratory research on control problem of nonlinear systems with applications and related topics.

Rongwei Guo
H. G. Enjieu Kadji
Xinguang Zhang
Uchechukwu E. Vincent
Wenguang Yu

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 8973583, 11 pages
http://dx.doi.org/10.1155/2016/8973583

Research Article

The Design and Its Application in Secure Communication and Image Encryption of a New Lorenz-Like System with Varying Parameter

Lilian Huang, Donghai Shi, and Jie Gao

Harbin Engineering University, Harbin, Heilongjiang 150001, China

Correspondence should be addressed to Lilian Huang; lilian_huang@163.com

Received 9 October 2015; Revised 8 December 2015; Accepted 24 February 2016

Academic Editor: Herve G. E. Kadji

A new Lorenz-like chaotic system with varying parameter is proposed by adding a state feedback function. The structure of the new designed system is simple and has more complex dynamic behaviors. The chaos behavior of the new system is studied by theoretical analysis and numerical simulation. And the bifurcation diagram shows a chaos-cycle-chaos evolution when the new parameter changes. Then a new synchronization scheme by a single state variable drive is given based on the new system and a chaotic parameter modulation digital secure communication system is also constructed. The results of simulation demonstrate that the new proposed system could be well applied in secure communication. Otherwise, based on the new system, the encryption and decryption of image could be achieved also.

1. Introduction

With the development of the mobile, PC, cloud computing, the Internet of things, and wearable devices, the data-intensive science such as big data [1] has become the main topic of the technological reform. The most prominent features of the big data are enormous volume of data, wide variety of data types, lower value density, and faster processing. But the big data has both advantages and disadvantages. It brings great convenience to individuals and enterprises; at the same time the data security is an urgent problem to be solved. From the view of information security, the traditional cryptography and secure communication model can be cracked easily, so that great security risks exist in the information system in every country. Therefore it is very urgent to improve the information security technology for the country and the enterprise in which the big data is main stream.

Due to the characteristic of the long-term unpredictability and extreme sensitivity to initial values, the chaos system has been researched deeply in the secure communications and the cryptography. Since Pecora and Carroll [2] first proposed the master-slave synchronization method in 1990, many synchronization types were presented, such as complete synchronization [3], lag synchronization [4], generalized synchronization [5], modified projective synchronization [6], modified function projective synchronization [7], phase synchronization [8], and dislocation synchronization [9]. Now more and more people paid their attention to chaotic secure communication, and the research mainly focuses on two aspects: one is to find a safer secure communication scheme, such as chaotic masking [10, 11], chaotic modulation [12], chaos shift keying [13], and chaos spreading spectrum [14] and the other is to research chaotic systems with a better encryption performance, such as fractional-order chaotic systems [15, 16], time-delay chaotic systems [17], complex chaotic system [18], and multiscroll chaotic systems [19]. Kiani-B et al. [20] applied the fractional-order Kalman filter in secure communications system. Mei [21] proposed a new secure communication scheme based on uncertain time-delay chaos system. Mahmoud et al. [22] researched the projective synchronization for complex hyperchaotic system and achieved secure communications with four-order complex Lorenz system. In addition, other secure communication schemes based on fractional-order [23], time-delay [24], and

multiscroll [25] chaotic systems have been proposed also. In the field of cryptography, compared with the traditional password, the generated mechanisms for chaos passwords are different and have real-time, so it has a greater advantage in terms of image encryption and video and other multimedia data encryption, and therefore the research on chaos image encryption has attracted more and more people [26–28]. The quality of the chaotic password is closely related to the chaos systems. For the low-dimensional chaotic system, because of its simple form, small key space, and low chaos sequence complexity, its security is not high enough. So many scholars focus on the hyperchaotic systems and fractional-order chaotic systems. Zhu and Sun [29] analyzed the security of the hyperchaos image encryption (HIE) algorithm, improved hyperchaos image encryption (IHIE) algorithm, and proposed the enhanced hyperchaos image encryption algorithms. Zhao et al. [30] gave an image encryption scheme based on an improper fractional-order chaotic system.

So the more complicated structure of the chaotic systems, the better performance of the secure communications and cryptography. However, these complicated chaotic systems are usually not easy to design synchronous controller, which decreases the communication efficiency. Therefore it is necessary to seek a new chaotic system with simple structure and complex behaviors.

In this paper, we propose a new Lorenz-like system with varying parameter by adding a state feedback factor in Lorenz-like system [31]. By theoretical analysis and numerical simulation, the structure of the new system is simple and easy to construct. At the same time, it has more complicated behaviors. This paper is divided into three parts as follows. Firstly, the new Lorenz-like chaotic system with varying parameter is designed based on the Lorenz-like system and analyzes its chaos characteristics theoretically. Secondly, a synchronization scheme driven by a single state variable is achieved based on the new proposed system, and the chaotic parameter modulation digital secure communications system is constructed. Finally, the designed variable parameter chaotic system is applied to image encryption and a three-chaotic-image encryption algorithm is proposed.

2. The Lorenz-Like System with Varying Parameter

2.1. The New Chaotic System. The Lorenz-like system is given by [31]

$$\begin{aligned} \dot{x} &= a(y-x), \\ \dot{y} &= bx - xz, \\ \dot{z} &= xy - cz, \end{aligned} \tag{1}$$

where a and c are real constants and b is a bifurcation parameter. Compared with the traditional Lorenz system, y is not in the second equation. If we replace parameter b with a function of x, such as

$$b = \begin{cases} d_1 + d_2, & |x| \geq \theta \\ d_1 - d_2, & |x| < \theta \end{cases} \quad d_1, d_2, \theta \in R, \tag{2}$$

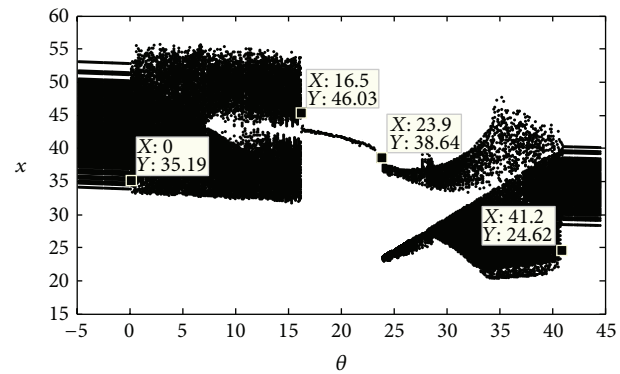


Figure 1: Bifurcation diagram of chaotic system (3) with the change of θ.

then a new system is generated and can be written as

$$\begin{aligned} \dot{x} &= a(y-x), \\ \dot{y} &= bx - xz, \\ \dot{z} &= xy - cz, \\ b &= d_1 + d_2 \operatorname{sgn}(|x| - \theta), \end{aligned} \tag{3}$$

where a, c, d_1, d_2, and θ are real constants, b is a state feedback control function, and θ is the threshold. From (2), we can know that b switches between $d_1 + d_2$ and $d_1 - d_2$ under the control of x; then the Lorenz-like system (3) shows the bifurcation under the control of state variable x.

When choosing $a = 20$, $c = 8$, $d_1 = 70$, and $d_2 = 15$, we can get a bifurcation diagram (Figure 1) of chaotic system (3) with the change of θ. For convenience, the Lorenz-like system when $b = 85$ is denoted as A chaotic system and when $b = 55$ as B chaotic system. From Figure 1, we can get that when $\theta < 0$, $b = 85$, the new Lorenz-like system is equivalent to the A chaotic system, while when $\theta > 41.2$, $b = 55$; it is equivalent to the B chaotic system. When $0 < \theta < 41.2$, b switches between 85 and 55 under the control of the state variable x; in other words, the new Lorenz-like system automatically switches between A and B chaotic systems (Figure 2). And when $16.5 < \theta < 23.9$, system appears periodic oscillation obviously.

As shown in Figure 2, the blue part denotes A chaotic system and the red part B chaotic system. With the change of parameter, the nonlinear dynamical behaviors change significantly. When $\theta = 10$ or $\theta = 30$, a strange attractor appears in Figures 2(b) and 2(d). In Figure 3, the three-dimensional phase diagram of chaotic system (3) is given with $\theta = 10$, 20, 30, and 37.

2.2. Chaotic Characters

2.2.1. Symmetry and Invariance. For system (3), let $(x, y, z) \rightarrow (-x, -y, z)$; the system equation remains the same. Then the system is symmetrical about the z-axis, and the symmetry is not associated with the system parameters. If we let $x(0) = 0$, $y(0) = 0$, and $z(0)$ be any value, the system equation can transform into $\dot{z} = -cz$; that

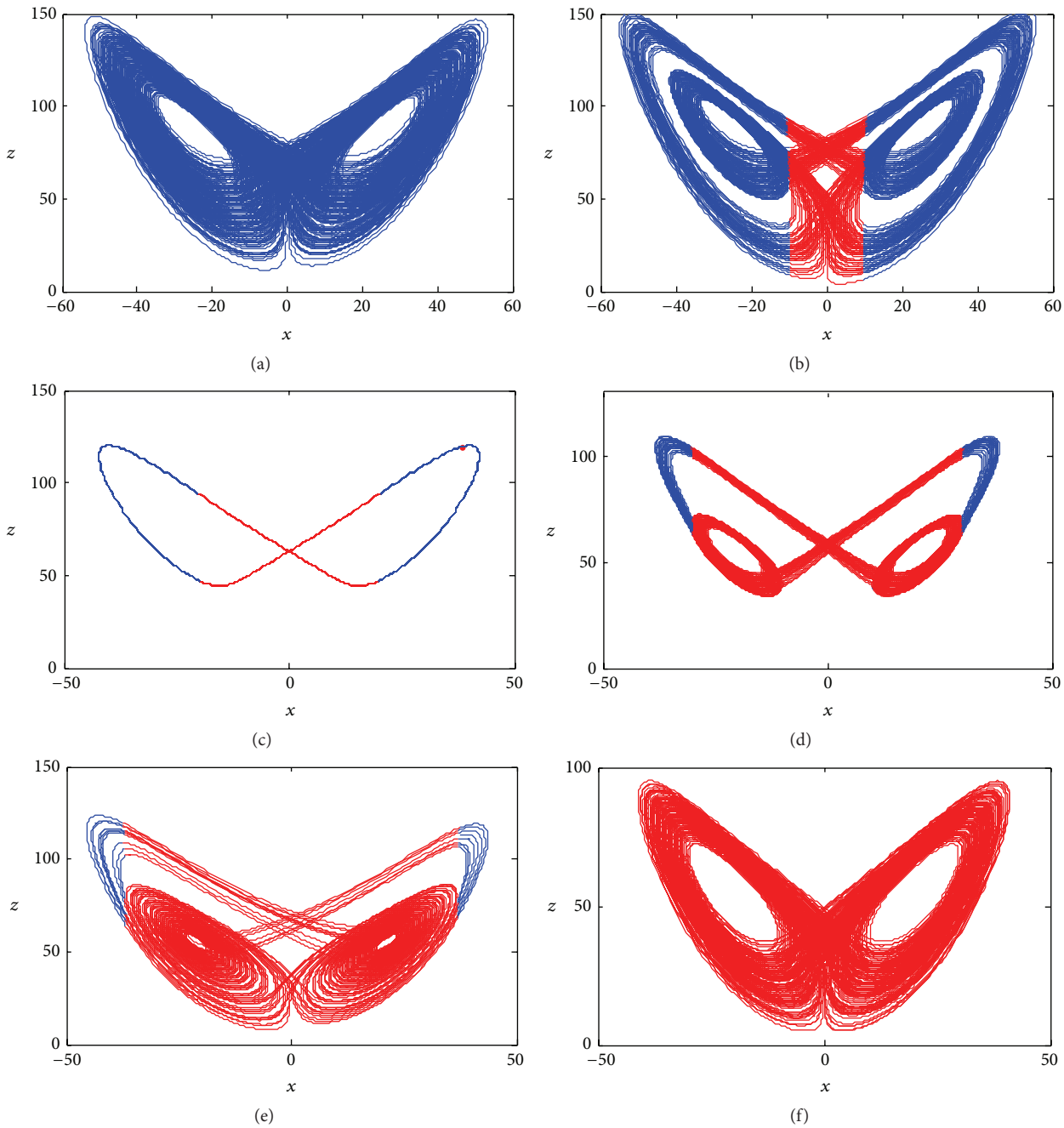


FIGURE 2: x-z plane phase diagram of chaotic system (3) (a) $\theta = -1$; (b) $\theta = 10$; (c) $\theta = 20$; (d) $\theta = 30$; (e) $\theta = 37$; and (f) $\theta = 50$.

is, the system will move on z-axis and will be stable at the origin.

2.2.2. Dissipation and the Existence of Attractor. For system (3),

$$\nabla \cdot f = \frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} + \frac{\partial \dot{z}}{\partial z} = -a - c < 0. \tag{4}$$

So, we can conclude that the system is dissipative and converges by exponential $dV/dt = e^{-(a+c)t}$; that is, the volume element with the initial volume $V(0)$ converges to $V(0)e^{-(a+c)t}$ at time t. When $t \to \infty$, each small volume that contains the system trajectories converges to zero at an exponential rate of $-(a+c)t$. All the trajectories of the system will eventually be limited to a subset of zero volume, and this limit subset is called attractor.

2.2.3. The Existence and Stability of Equilibrium Point. For system (3), the equilibrium points are

$$\begin{gathered} O(0,0,0), \\ P^{+}\left(\sqrt{cb}, \sqrt{cb}, b\right), \\ P^{-}\left(-\sqrt{cb}, -\sqrt{cb}, b\right). \end{gathered} \tag{5}$$

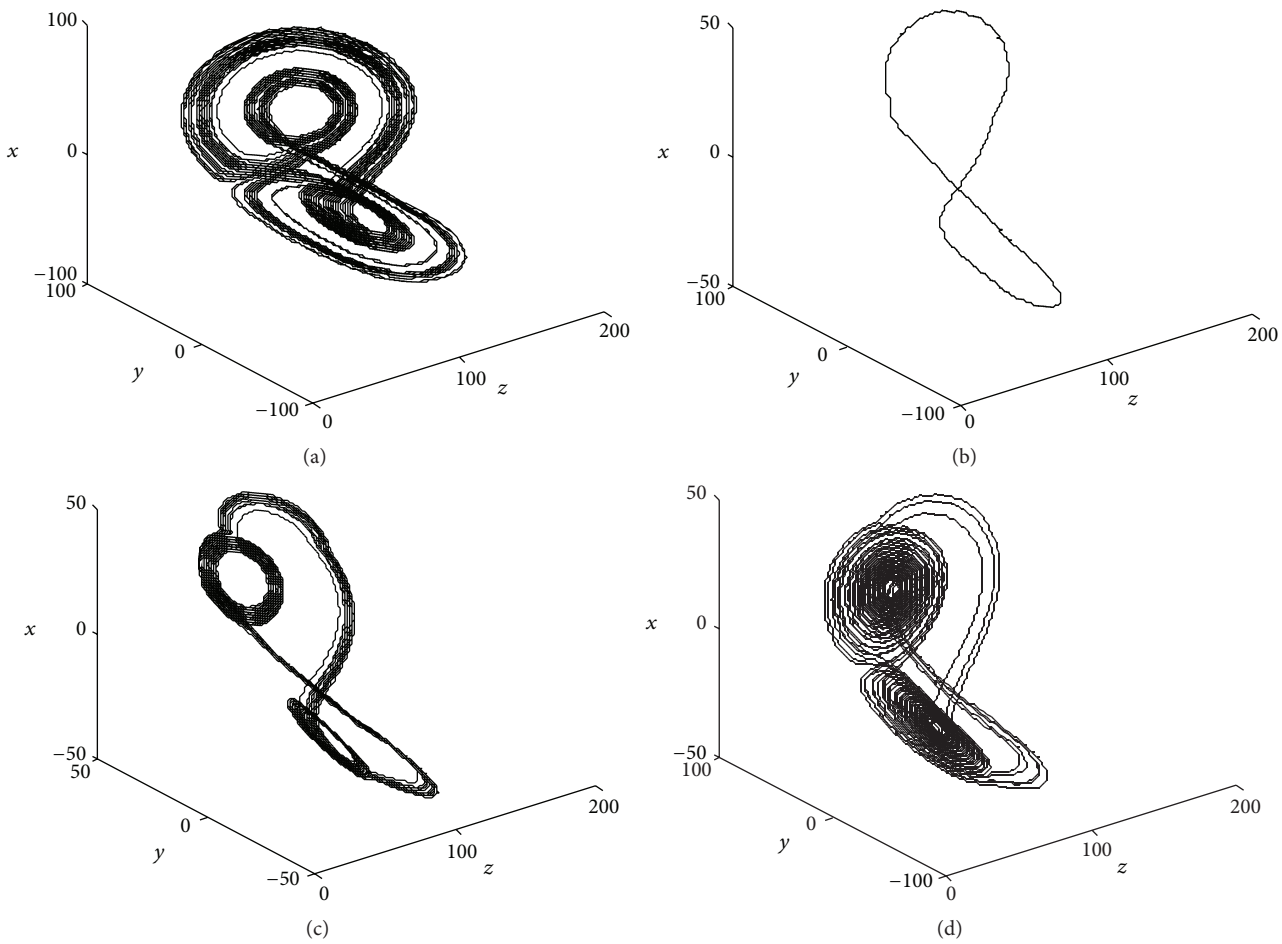


Figure 3: Three-dimensional phase diagram of chaotic system (3): (a) $\theta = 10$; (b) $\theta = 20$; (c) $\theta = 30$; and (d) $\theta = 37$.

It is easy to know that O is the shared equilibrium point for both A system and B system. When choosing $a = 20$, $c = 8$, $d_1 = 70$, and $d_2 = 15$, the other two equilibrium points of A system are $P_A^+(26.0768, 26.0768, 85)$ and $P_A^-(-26.0768, -26.0768, 85)$, and the other two equilibrium points of B system are $P_B^+(20.9762, 20.9762, 55)$ and $P_B^-(-20.9762, -20.9762, 55)$. The distribution of equilibrium points in the phase space can be seen in Table 1, in which D_{-1}, D_0, and D_{+1} denote three areas separated by θ as follows:

$$\begin{aligned} D_{+1} &= \{(x, y, z) \mid x > \theta\}, \\ D_0 &= \{(x, y, z) \mid -\theta \le x \le \theta\}, \\ D_{-1} &= \{(x, y, z) \mid x < -\theta\}. \end{aligned} \tag{6}$$

For A chaotic system, the corresponding eigenvalues for each equilibrium point can be calculated as follows:

$$\begin{aligned} \lambda_1^O &= -52.4264 \quad \lambda_1^{P_A^+} = \lambda_1^{P_A^-} = -30.1071, \\ \lambda_2^O &= +32.4264 \quad \lambda_2^{P_A^+} = \lambda_2^{P_A^-} = 1.0536 + j30.0388, \\ \lambda_3^O &= -8 \quad \lambda_3^{P_A^+} = \lambda_3^{P_A^-} = 1.0536 - j30.0388. \end{aligned} \tag{7}$$

Table 1: The distribution of equilibrium points in the phase space.

The scope of θ	Region		
	$D_{-1}{}^{\ddagger}$	$D_0{}^{\dagger}$	$D_{+1}{}^{\ddagger}$
$\theta \le 0$		$P_A^{-\ddagger}$ $O^{\ddagger}$ $P_A^{+\ddagger}$	
$0 < \theta < 20.9762$	$P_A^{-\ddagger}$ $P_B^{-\ddagger}$	$O^{\dagger}$	$P_B^{+\ddagger}$ $P_A^{+\ddagger}$
$20.9762 \le \theta \le 26.0768$	$P_A^{-\ddagger}$	$P_B^{-\dagger}$ $O^{\dagger}$ $P_B^{+\dagger}$	$P_A^{+\ddagger}$
$\theta > 26.0768$		$P_A^{-\dagger}$ $P_B^{-\dagger}$ $O^{\dagger}$ $P_B^{+\dagger}$ $P_A^{+\dagger}$	
$\theta \gg 26.0768$		$P_B^{-\dagger}$ $O^{\dagger}$ $P_B^{+\dagger}$	

Note: † represents D_0 and ‡ represents D_{-1} and D_{+1}.

Obviously, O is a saddle point, and P_A^+, P_A^- are saddle-focus equilibrium points. And all these three equilibrium points are unstable, which leads the orbits of system stretch in phase space. Under the interactive stretching and contractions, the chaotic motion is generated. In the same way, we also can draw a similar conclusion that B chaotic system also has three unstable equilibrium points and the chaotic condition is satisfied.

From Table 1, it is easy to know that the system has three equilibrium points when $\theta \le 0$ and it is equal to A chaotic

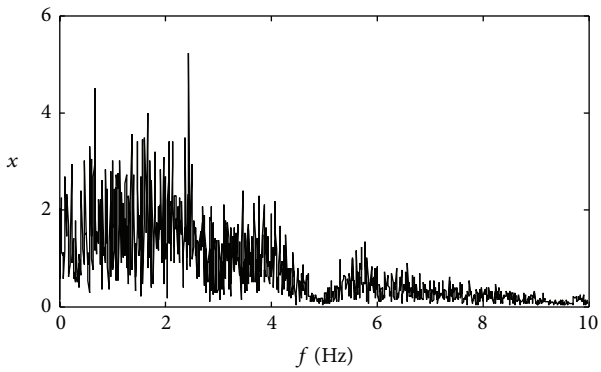


Figure 4: The spectrum of Lorenz-like system with varying parameter when $\theta = 4$.

Table 2: Lyapunov exponents and Lyapunov dimension.

θ	Lyapunov exponents			Lyapunov dimension (D_L)
	λ_{L1}	λ_{L2}	λ_{L3}	
$\theta \le 0$ (A system)	2.5110	−0.0029	−30.4572	2.0824
$\theta = 4$	2.2269	−1.6438	−28.5472	2.0204
$\theta \gg 26.0678$ (B system)	2.0533	0.0242	−30.0060	2.0684

system. The system also can be treated approximately such that it has three equilibrium points when $\theta \gg 26.0768$ and the system is equal to B chaotic system. But in addition to these two cases, the system has five equilibrium points. Of all the five points, O influences the trajectory in the whole region, while P_B^+ and P_B^- influence the trajectory in D_0, and P_A^+ and P_A^- influence the trajectory in D_{-1} and D_{+1}. With the increase of θ, D_0 gradually extends and $D_{\pm 1}$ is reduced; that is, the influence of P_B^+ and P_B^- gradually increases while P_A^+ and P_A^- are reduced until they nearly disappear. So the system shows a complex dynamic chaos-cycle-chaos when the new parameter θ changes.

2.2.4. Spectrum. In Figure 4, we can see that spectrum of the system is continuous, which shows that the new designed system has the chaotic characteristics.

2.2.5. Lyapunov Exponents and Lyapunov Dimension. Lyapunov exponent measures the exponential rates of divergence or convergence of nearby trajectories in phase space. A three-order nonlinear system has three Lyapunov exponents $(\lambda_{L1}, \lambda_{L2}, \lambda_{L3})$. All the Lyapunov exponents are listed in Table 2, and the curves with the change of θ are also given as in Figure 5. Obviously λ_{L2} of A and B systems is close to zero, while λ_{L2} (when $\theta = 4$) is a negative number for the reason of θ; this implies that a new chaotic attractor occurred in the new system.

For a n-order system, the Lyapunov dimension can be calculated as follows:

$$D_L = j + \frac{\lambda_{L1} + \lambda_{L2} + \cdots + \lambda_{Lj}}{\left|\lambda_{L(j+1)}\right|}, \tag{8}$$

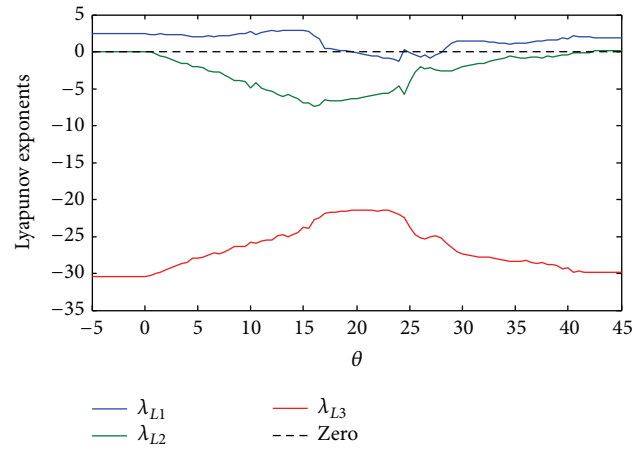


Figure 5: Lyapunov exponents curves of chaotic system (3) with the change of θ.

where $\lambda_{L1} > \lambda_{L2} > \cdots > \lambda_{Ln}$ and $\lambda_{L1} + \lambda_{L2} + \cdots + \lambda_{Lj} > 0$ while $\lambda_{L1} + \lambda_{L2} + \cdots + \lambda_{L(j+1)} < 0$.

From the results in Table 2, we can get that all the Lyapunov dimensions are fractions and $2 < D_L < 3$. Thus, it is another evidence of chaos. In addition, both Lyapunov exponents' curves and bifurcation diagram can show the effect of parameter, so the same conclusion can be obtained from Figure 5 as Figure 1.

2.2.6. A Brief Summary. This section shows a new Lorenz-like system (3) with varying parameter; several conclusions can be gotten as follows: (i) b is a constant in Lorenz-like system (1) while it switches between $d_1 + d_2$ and $d_1 - d_2$ in system (3), so the new system's structure has a slight difference with the Lorenz-like system (1) and it equals system (1) when $\theta \le 0$ or $\theta \gg 26.0678$; (ii) new chaotic behaviors occur when θ change (see Figures 1, 2, and 3); (iii) the equilibrium points are not fixed for the reason of θ as Table 1 shows; and (iv) the Lyapunov exponent λ_{L2} is apparently different (see Table 2 and Figure 5). All the conclusions imply that the new proposed system has more complicated behaviors with respect to the Lorenz-like system (1).

3. The Application in Secure Communication for the New Lorenz-Like System

3.1. Synchronization Design for Single Variable Drive. For a better description of the synchronization scheme, here we use notation (x_1, x_2, x_3) in place of (x, y, z) in (3); then the master system is

$$\begin{aligned} \dot{x}_1 &= a(x_2 - x_1), \\ \dot{x}_2 &= b_x x_1 - x_1 x_3, \\ \dot{x}_3 &= -c x_3 + x_1 x_2, \\ b_x &= d_1 + d_2 \operatorname{sgn}(|x_1| - \theta), \quad \theta \in R. \end{aligned} \tag{9}$$

And the slave system is

$$\begin{aligned} \dot{y}_1 &= a(y_2 - y_1), \\ \dot{y}_2 &= b_y y_1 - y_1 y_3 + u_1, \\ \dot{y}_3 &= -c y_3 + y_1 y_2 + u_2, \\ b_y &= d_1 + d_2 \operatorname{sgn}(|y_1| - \theta), \quad \theta \in R. \end{aligned} \tag{10}$$

So the controller is designed as follows:

$$\begin{aligned} u_1 &= -b_y y_1 + b_x x_1 + a x_1 + y_1 y_3 - x_1 y_3 - a y_1, \\ u_2 &= -y_1 y_2 + x_1 y_2. \end{aligned} \tag{11}$$

The designed controller only contains one state variable x_1 of the master system; thus it has simple structure and is driven by single variable. So it is easy to be achieved.

Let the error system be

$$\begin{aligned} e_1 &= y_1 - x_1, \\ e_2 &= y_2 - x_2, \\ e_3 &= y_3 - x_3. \end{aligned} \tag{12}$$

Then, the error dynamics equation is

$$\begin{aligned} \dot{e}_1 &= a(e_2 - e_1), \\ \dot{e}_2 &= -x_1 e_3 - a e_1, \\ \dot{e}_3 &= -c e_3 + x_1 e_2. \end{aligned} \tag{13}$$

Select the Lyapunov function as $(e) = (1/2)e_1^2 + (1/2)e_2^2 + (1/2)e_3^2$; then take the derivative of $V(e)$, so

$$\begin{aligned} \dot{V}(e) &= e_1 \dot{e}_1 + e_2 \dot{e}_2 + e_3 \dot{e}_3 \\ &= e_1 a(e_2 - e_1) + e_2(-x e_3 - a e_1) \\ &\quad + e_3(-c e_3 + x e_2) \\ &= -a e_1^2 + a e_1 e_2 - x_1 e_2 e_3 - a e_1 e_2 - c e_3^2 + x_1 e_2 e_3 \\ &= -a e_1^2 - c e_3^2. \end{aligned} \tag{14}$$

When $e = (e_1, e_2, e_3)^{\mathrm{T}} = (0,0,0)^{\mathrm{T}}$, $V(e) = 0$, and when $e \neq (0,0,0)^{\mathrm{T}}$, $V(e) > 0$ and $\dot{V}(e) < 0$. According to the Lyapunov stability theorem, $e_1 \to 0$, $e_2 \to 0$, and $e_3 \to 0$ when $t \to \infty$; that is, the synchronization between master and slave system has been achieved.

Figure 6 gives the synchronization error curves between the master system and slave system with $a = 20$, $c = 8$, $d_1 = 70$, $d_2 = 15$, and $\theta = 4$; the initial value $(x(0), y(0), z(0)) = (-10.5, -7, 8)$, $(x_1(0), y_1(0), z_1(0)) = (0, 10, 6)$. From Figure 6, the master system traces the slave system to achieve synchronization quickly. The advantage of this chaotic synchronization system is that the controller is simple and only one signal is to be transmitted to complete the synchronization between the drive system and response system, which improves communication efficiency and conserves resources.

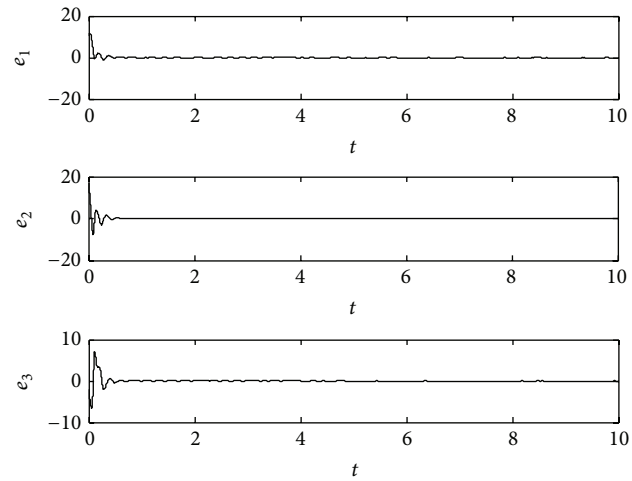


Figure 6: Synchronization error curve between master and slave systems when $\theta = 4$.

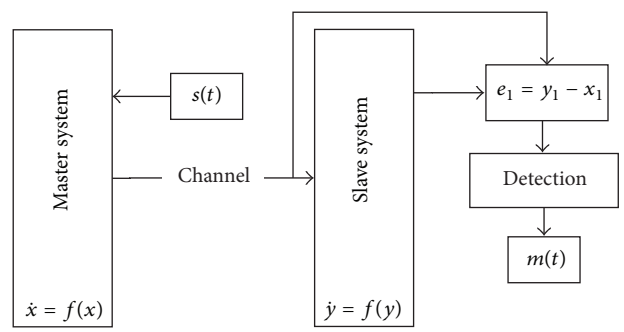


Figure 7: The schematic of digital secure communications.

3.2. Chaos Parameter Modulation Digital Secure Communications System. Based on the synchronization scheme designed in the previous section, a chaotic parameter modulation digital secure communication system is given in Figure 7.

The transmitter system is

$$\begin{aligned} \dot{x}_1 &= 20(x_2 - x_1), \\ \dot{x}_2 &= b_x x_1 - x_1 x_3, \\ \dot{x}_3 &= -8 x_3 + x_1 x_2, \\ b_x &= 70 + 15 \operatorname{sgn}(|x_1| - (4 + 2s(t))). \end{aligned} \tag{15}$$

The receiver system is

$$\begin{aligned} \dot{y}_1 &= 20(y_2 - y_1), \\ \dot{y}_2 &= (b_y + 20) x_1 - x_1 y_3 - a y_1, \\ \dot{y}_3 &= -8 y_3 + x_1 y_2, \\ b_y &= 70 + 15 \operatorname{sgn}(|x_1| - 4), \end{aligned} \tag{16}$$

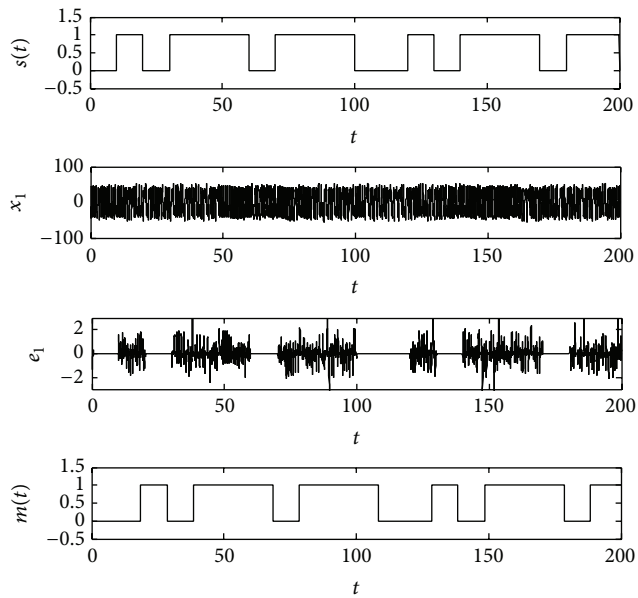


Figure 8: Chaotic parameter modulation digital secure communication.

where $s(t)$ is a digital signal to be transmitted. In order to encrypt $s(t)$, we select $\theta = 4$ to represent "0" and $\theta = 6$ to represent "1"; that is,

$$\theta = 4 + 2(s(t)) = \begin{cases} 4, & s(t) = 0 \\ 6, & s(t) = 1. \end{cases} \tag{17}$$

In this case, the topology of the phase diagram is similar, so it is difficult to crack encrypted signal and extract useful information by phase space reconstruction.

Based on the theory above, we simulate the digital secure communication system as in Figure 8. The digital signal $s(t)$ "0101110111001011101" will be transmitted and sent per symbol interval, that is, 10 seconds, where x_1 is not only the encrypted signal but also the driven signal. The synchronization between the master system and slave system only can be reached when $\theta = 4$, and there is a large error between the response system and the drive system when $\theta = 6$ just as e_1 in Figure 8. Finally the decrypted signal $m(t)$ can be gotten from e_1 after detection, and compared with $s(t)$, there is a nearly 10-second delay.

4. Image Encryption Algorithm Based on Lorenz-Like System with Varying Parameter

4.1. A Three-Order Image Encryption Algorithm. Based on system (3), a three-order image encryption algorithm is given and the diagram of the image encryption and decryption is shown in Figure 9.

The chaotic image encryption is to disrupt the original image (plaintext) by chaotic sequence. The process of the algorithm is as follows: first, set the initial value of the system as the key; then iterate system (3) for 5000 times to make it fully chaotic. Later, continue to iterate system (3) to obtain $x(n)$, $y(n)$, and $z(n)$. Then the chaotic sequences key$x(n)$, key$y(n)$, and key$z(n)$ can be gotten as below:

$$\begin{aligned} &\text{key}x(n) \\ &\quad = \operatorname{mod}\left((|x(n)| - \operatorname{floor}(|x(n)|)) \times 10^{14}, 256\right), \\ &\text{key}y(n) \\ &\quad = \operatorname{mod}\left((|y(n)| - \operatorname{floor}(|y(n)|)) \times 10^{14}, 256\right), \\ &\text{key}z(n) \\ &\quad = \operatorname{mod}\left((|z(n)| - \operatorname{floor}(|z(n)|)) \times 10^{14}, 256\right), \end{aligned} \tag{18}$$

where "floor" is the MATLAB function and floor(A) rounds the element A to the nearest integers less than or equal to A.

Through the above process in (18), we get the chaotic sequences key$x(n)$, key$y(n)$, and key$z(n)$ which range between 0 and 255. Original image matrix (plaintext) is P and the ciphertext matrices are C_1, C_2, and C_3, which are obtained after three-order encryption, respectively, where C_3 is the final ciphertext image matrix. The encryption formula is as follows:

$$\begin{aligned} C_1(n) &= P(n) \oplus \text{key}x(n), \\ C_2(n) &= C_1(n) \oplus \text{key}y(n), \\ C_3(n) &= C_2(n) \oplus \text{key}z(n). \end{aligned} \tag{19}$$

"$\oplus$" in (19) means "XOR", and the same is in (20). Just as Figure 9, the decryption process is the opposite of the encryption process. First, we should set the correct key; then for system (3), the decryption process is the same with the encryption to obtain the same chaotic sequences to decrypt correctly. Decryption formula is as follows:

$$\begin{aligned} C_2(n) &= C_3(n) \oplus \text{key}z(n), \\ C_1(n) &= C_2(n) \oplus \text{key}y(n), \\ P(n) &= C_1(n) \oplus \text{key}x(n). \end{aligned} \tag{20}$$

4.2. Simulation and Analysis. In this section, an image encryption experiment was given and a 512 × 512 color image "Lena" is chosen as the plaintext. In simulation, the step is selected as 0.01, and $a = 20$, $c = 8$, $d_1 = 70$, $d_2 = 15$, and $\theta = 4$. The encryption key is initial value $(x(0), y(0), z(0)) = (0.2, 0.7, 1.6)$. Based on the above algorithm, the image encryption system is designed to achieve the Lena encryption. The simulation results shown in Figure 10, and several tests have been carried out to demonstrate the effectiveness and efficiency of the proposed encryption algorithm.

4.2.1. Key Space Analysis. Key space size is the total number of different keys which can be used in an encryption process; it should be large enough to preclude the eavesdropping by brute-force attack. A single precision floating point format number has 2^{32} kinds of possibilities; then the key space

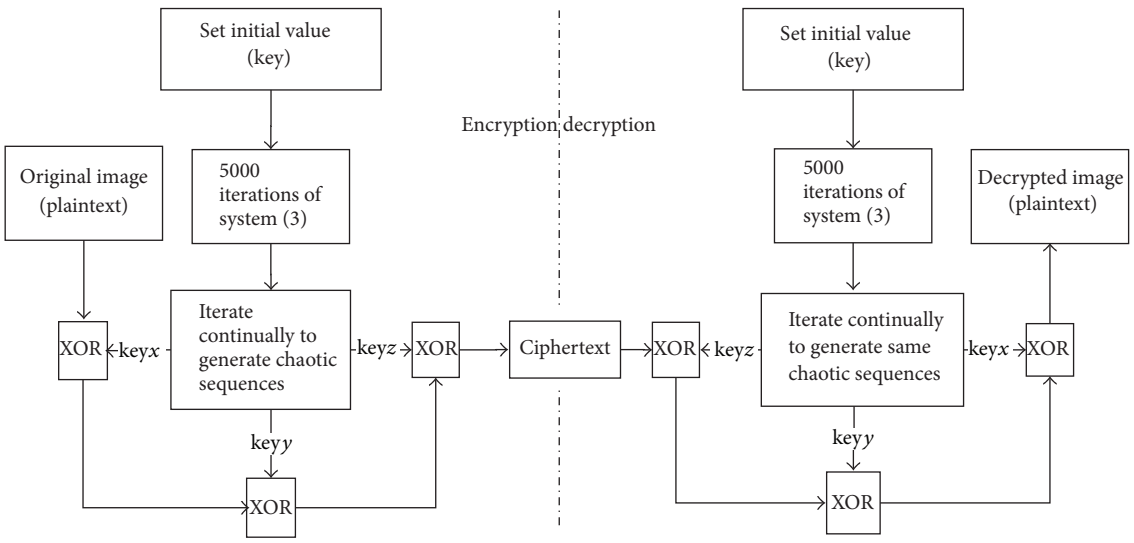


Figure 9: The diagram of image encryption and decryption.

(a) (b) (c) (d)

Figure 10: The simulation results of image encryption and decryption. (a) Original image (plaintext); (b) encrypted image (ciphertext); (c) decrypted image (plaintext); and (d) illegal decrypted image.

in this paper will be up to 2^{96}. Otherwise, if the system parameters are chosen as keys, it will get much larger key space, which greatly increases the security of the system.

4.2.2. Histogram Analysis. From the histogram of a digital image, the distribution of the pixel values can be gotten; if the encrypted image is well encrypted, the histogram will be uniform, so the histogram attack can be prevented effectively. Figure 11 gives the histograms of both original image and encrypted image; it is obvious that the histograms of red, green, and blue for original image are steep and not flat enough; the histograms for encrypted image are all uniform and quite different from that of the original image when using the proposed algorithm.

4.2.3. Key Sensitivity Analysis. A good cryptosystem must be highly sensitive at small changes in secret key in encryption and decryption process, so a full test contains two aspects: (i) slightly different keys to encrypt the same image are used and the difference between the corresponding encrypted images is computed; (ii) for an encrypted image only one correct key can decrypt it, so decrypt the encrypted image by an incorrect key which is similar to the correct one and observe whether it can be correctly decrypted.

Table 3 gives some special cases to evaluate the sensitivity in encryption process, and the encrypted images were also shown in Figure 12; the difference ratio is really high, which means a good key sensitivity in encryption process. The test result in decryption process also can be seen in Figure 12; Figure 12(f) is the correct decrypted image; Figure 12(g) is the incorrect one with only a slight change 10^{-14} for the key's first value.

4.2.4. Correlation Coefficients and Efficiency Analysis. A good image encryption algorithm should have two characteristics: (i) high security, which is partly analyzed in key space and key sensitivity, will be analyzed by correlation coefficients complementally in this section; (ii) high efficiency, which means low time consumption in encryption and decryption process, will be analyzed in this section also.

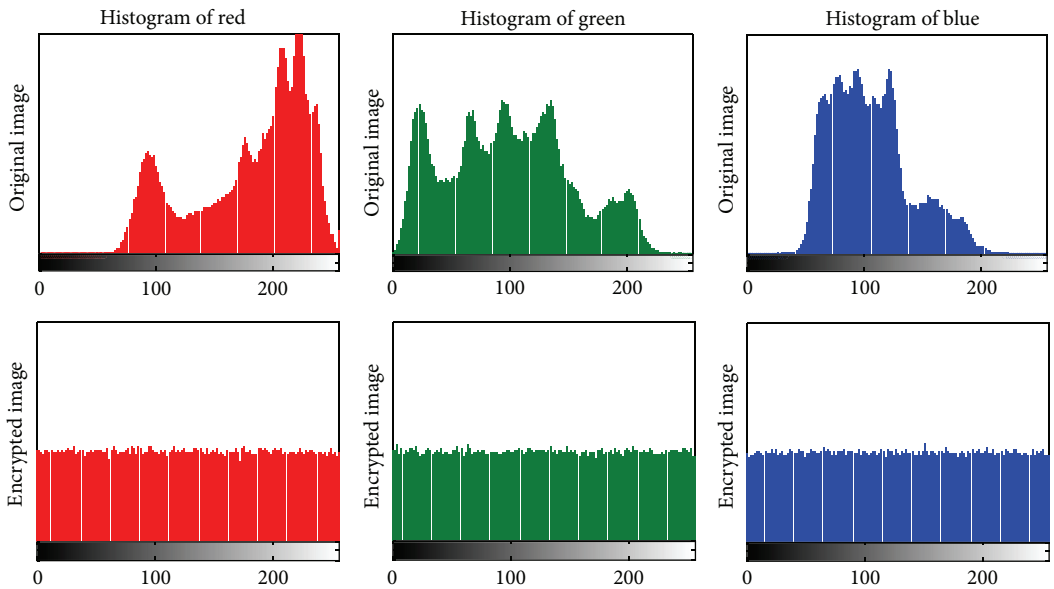


Figure 11: Histogram of original image and encrypted image.

Table 3: Differences between encrypted images produced by slightly different keys.

Encryption keys (I) Figure 12(b)			Encryption keys (II)				Difference ratio between (I) and (II) (%)
$x(0)$	$y(0)$	$z(0)$	$x(0)$	$y(0)$	$z(0)$		
			$2.5 + 1 \times 10^{-14}$	-3.7	9.3	Figure 12(c)	99.61
2.5	-3.7	9.3	2.5	$-3.7 + 1 \times 10^{-14}$	9.3	Figure 12(d)	99.60
			2.5	-3.7	$9.3 + 1 \times 10^{-14}$	Figure 12(e)	99.60

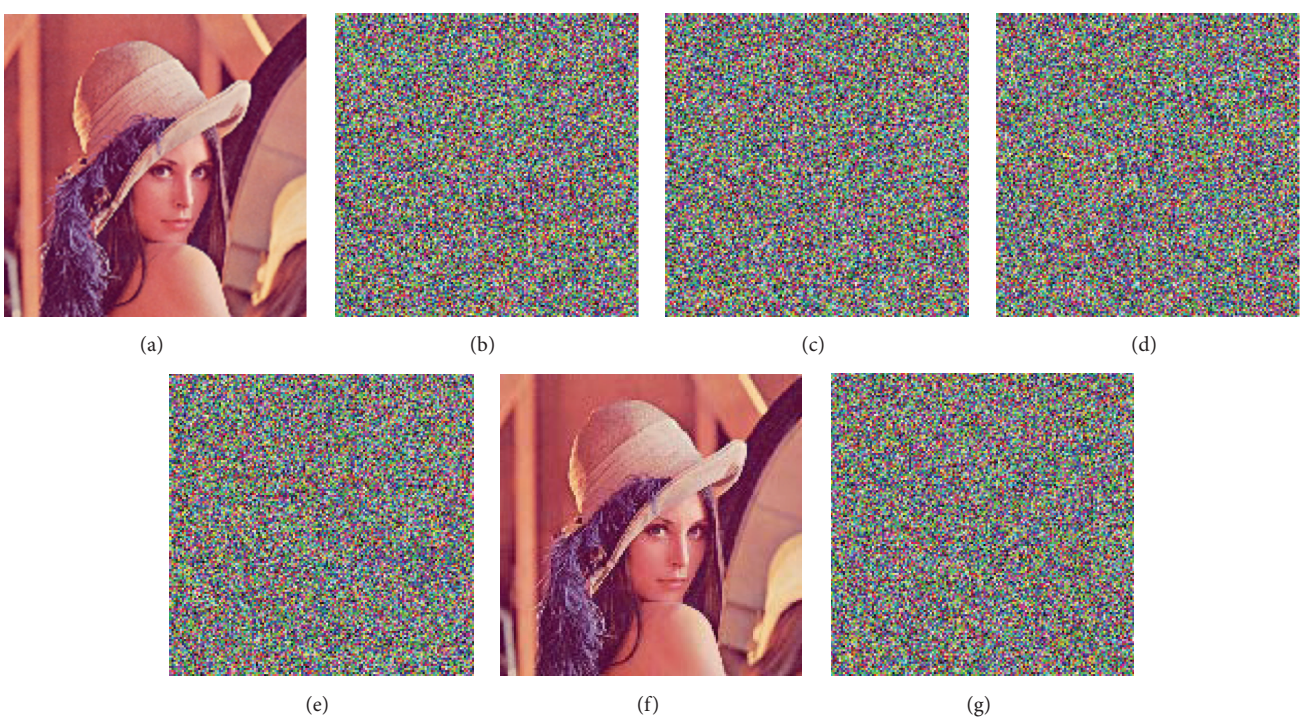


Figure 12: Key sensitivity test. (a) Original image (plaintext); (b), (c), (d), (e) the encrypted images with different keys as Table 3; (f) decrypt (b) with key (2.5, −3.7, 9.3); and (g) decrypt (b) with key ($2.5–1 \times 10^{-14}$, −3.7, 9.3).

Table 4: Correlation coefficients and cost comparisons.

Chaotic systems used	Correlation coefficients						Cost (s)
	Original image (plaintext)			Encrypted image (ciphertext)			
	Horizontal	Vertical	Diagonal	Horizontal	Vertical	Diagonal	
Proposed system	0.9788	0.9677	0.9752	0.0086	0.0307	−0.0326	10.9689
Reference [29]	0.9788	0.9677	0.9752	0.0167	−0.0004	0.0371	13.9933
Reference [30]	0.9788	0.9677	0.9752	−0.0750	−0.1078	0.0036	14.0401

The correlation coefficient of an image can be measured as follows:

$$C_r = \frac{N\sum_{i=0}^{N} x_i y_i - \left(\sum_{i=0}^{N} x_i\right)\left(\sum_{i=0}^{N} y_i\right)}{\sqrt{\left(N\sum_{i=0}^{N} x_i^2 - \left(\sum_{i=0}^{N} x_i\right)^2\right)\left(N\sum_{i=0}^{N} y_i^2 - \left(\sum_{i=0}^{N} y_i\right)^2\right)}}, \quad (21)$$

where N is the number of pair of pixels and x and y are values of two adjacent pixels in grey scale. The correlation coefficients are calculated out based on 3000 random pixels, and all correlation coefficients of plaintext are greater than 0.96 while those of ciphertext are all near "zero," which implies a good information hiding for plaintext.

Recently many chaotic systems with complex structure were found or applied to image encryption, such as systems in [29, 30]. Here we realize image encryption with different chaotic systems under the same simulation environment; then comparisons with the proposed algorithm were done and the results of correlation coefficients and cost are listed in Table 4. Compared with algorithms based on other chaotic systems in [29, 30], we can know that the correlation coefficients of ciphertext are nearly, but the time consumption is less for the proposed system. This comparison demonstrates that the proposed image encryption algorithm based on the new Lorenz-like system shows a good performance as well as algorithms based on other systems, and the efficiency is high for the reason of its simple structure.

5. Conclusion

A new Lorenz-like system with varying parameter is proposed by adding a state feedback function in this paper. Firstly, we analyze the influence of the threshold θ to the chaotic behavior of the new system and found that the system shows a chaos-cycle-chaos evolution when θ changes. Then we analyze the new system's chaotic characteristics. After that a new synchronization scheme using a single state variable drive based on the new system is designed. Finally, a chaotic parameter modulation digital secure communication system and image encryption based on the new system proposed is designed. The simulation results show that the new system has a good performance in application. Otherwise, according to the new system designed, we can modify many other systems to get more chaotic systems which have simple structure and complex dynamics. This will enrich the amount of chaotic signal sources, simplify designing, and improve the security of communication and image encryption.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This research is funded by the National Natural Science Foundation of China (nos. 61203004 and 61306142) and the Natural Science Foundation of Heilongjiang Province (Grant no. F201220).

References

[1] W. Feng, "The opportunities and challenges of information security in big data era," *China Venture Capital*, vol. 34, pp. 49–53, 2013.

[2] L. M. Pecora and T. L. Carroll, "Synchronization in chaotic systems," *Physical Review Letters*, vol. 64, no. 8, pp. 821–824, 1990.

[3] J. Q. Lu and J. D. Cao, "Adaptive complete synchronization of two identical or different chaotic (hyperchaotic) systems with fully unknown parameters," *Chaos*, vol. 15, no. 4, Article ID 043901, 10 pages, 2005.

[4] M. G. Rosenblum, A. S. Pikovsky, and J. Kurths, "From phase to lag synchronization in coupled chaotic oscillators," *Physical Review Letters*, vol. 78, no. 22, pp. 4193–4196, 1997.

[5] N. F. Rulkov, M. M. Sushchik, L. S. Tsimring, and H. D. I. Abarbanel, "Generalized synchronization of chaos in directionally coupled chaotic systems," *Physical Review E*, vol. 51, no. 2, pp. 980–994, 1995.

[6] G.-H. Li, "Modified projective synchronization of chaotic system," *Chaos, Solitons and Fractals*, vol. 32, no. 5, pp. 1786–1790, 2007.

[7] K. S. Sudheer and M. Sabir, "Adaptive modified function projective synchronization between hyperchaotic Lorenz system and hyperchaotic LU system with uncertain parameters," *Physics Letters A*, vol. 373, no. 41, pp. 3743–3748, 2009.

[8] B. Blasius, A. Huppert, and L. Stone, "Complex dynamics and phase synchronization in spatially extended ecological systems," *Nature*, vol. 399, no. 6734, pp. 354–359, 1999.

[9] L.-L. Huang, S.-S. Shi, and J. Zhang, "Dislocation synchronization of the different complex value chaotic systems based on single adaptive sliding mode controller," *Mathematical Problems in Engineering*, vol. 2015, Article ID 240586, 8 pages, 2015.

[10] K. M. Cuomo, A. V. Oppenheim, and S. H. Strogatz, "Synchronization of Lorenz-based chaotic circuits with applications to communications," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 40, no. 10, pp. 626–633, 1993.

[11] L. Kocarev, K. S. Halle, K. Eckert, L. O. Chua, and U. Parlitz, "Experimental demonstration of secure communications via

chaotic synchronization," *International Journal of Bifurcation and Chaos*, vol. 2, no. 3, pp. 709–713, 1992.

[12] T. Yang and L. O. Chua, "Secure communication via chaotic parameter modulation," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 43, no. 9, pp. 817–819, 1996.

[13] G. Kolumban, M. P. Kennedy, G. Kis, and Z. Jako, "FM-DCSK: a novel method for chaotic communications," in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS '98)*, pp. 477–480, Monterey, Calif, USA, June 1998.

[14] M. Itoh, "Spread spectrum communication via chaos," *International Journal of Bifurcation and Chaos*, vol. 9, no. 1, pp. 155–213, 1999.

[15] C. Li and G. Chen, "Chaos in the fractional order Chen system and its control," *Chaos, Solitons and Fractals*, vol. 22, no. 3, pp. 549–554, 2004.

[16] D. Chen, C. Liu, C. Wu, Y. Liu, X. Ma, and Y. You, "A new fractional-order chaotic system and its synchronization with circuit simulation," *Circuits, Systems, and Signal Processing*, vol. 31, no. 5, pp. 1599–1613, 2012.

[17] J. Tang, "Synchronization of different fractional order time-delay chaotic systems using active control," *Mathematical Problems in Engineering*, vol. 2014, Article ID 262151, 11 pages, 2014.

[18] P. Liu and S. Liu, "Anti-synchronization between different chaotic complex systems," *Physica Scripta*, vol. 83, no. 6, Article ID 065006, 9 pages, 2011.

[19] C.-X. Zhang and S.-M. Yu, "Design and implementation of a novel multi-scroll chaotic system," *Chinese Physics B*, vol. 18, no. 1, pp. 119–129, 2009.

[20] 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.

[21] R. Mei, "Secure communication scheme using uncertain delayed chaotic system synchronization based on disturbance observers," in *Proceedings of the International Workshop on Chaos-Fractals Theories and Applications (IWCFTA '09)*, pp. 177–181, IEEE, Shenyang, China, November 2009.

[22] G. M. Mahmoud, E. E. Mahmoud, and A. A. Arafa, "On projective synchronization of hyperchaotic complex nonlinear systems based on passive theory for secure communications," *Physica Scripta*, vol. 87, no. 5, Article ID 055002, 10 pages, 2013.

[23] H.-F. Cao and R.-X. Zhang, "Parameter modulation digital communication and its circuit implementation using fractional-order chaotic system via a single driving variable," *Acta Physica Sinica*, vol. 61, no. 2, pp. 123–130, 2012.

[24] M. J. Wang and X. Y. Wang, "A secure communication scheme based on parameter identification of first order time-delay chaotic system," *Acta Physica Sinica*, vol. 58, no. 3, pp. 1467–1472, 2009.

[25] L. Gámez-Guzmán, C. Cruz-Hernández, R. M. López-Gutiérrez, and E. E. García-Guerrero, "Synchronization of Chua's circuits with multi-scroll attractors: application to communication," *Communications in Nonlinear Science and Numerical Simulation*, vol. 14, no. 6, pp. 2765–2775, 2009.

[26] S. Li and X. Zheng, "Cryptanalysis of a chaotic image encryption method," in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS '02)*, vol. 2, pp. 708–711, Phoenix-Scottsdale, Ariz, USA, May 2002.

[27] A. Kanso and M. Ghebleh, "A novel image encryption algorithm based on a 3D chaotic map," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 7, pp. 2943–2959, 2012.

[28] L. Y. Zhang, X. B. Hu, Y. S. Liu, K.-W. Wong, and J. Gan, "A chaotic image encryption scheme owning temp-value feedback," *Communications in Nonlinear Science and Numerical Simulation*, vol. 19, no. 10, pp. 3653–3659, 2014.

[29] C.-X. Zhu and K.-H. Sun, "Cryptanalysis and improvement of a class of hyperchaos based image encryption algorithms," *Acta Physica Sinica*, vol. 61, no. 12, p. 120503, 2012.

[30] J. F. Zhao, S. H. Wang, Y. X. Chang, and X. F. Li, "A novel image encryption scheme based on an improper fractional-order chaotic system," *Nonlinear Dynamics*, vol. 80, no. 4, pp. 1721–1729, 2015.

[31] L. L. Huang, X. Y. Wang, and G. H. Sun, "Design and circuit simulation of the new Lorenz chaotic system," in *Proceedings of the 3rd International Symposium on Systems and Control in Aeronautics and Astronautics (ISSCAA '10)*, pp. 1443–1447, Harbin, China, June 2010.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 8432764, 7 pages
http://dx.doi.org/10.1155/2016/8432764

Research Article

Research on Submarine Straight-Line Track Control Underwater Based on Nonlinear Proportion Differential

Chenggang Gong,[1] Xia Zhu,[1] Fei Wang,[2] Pan Jiang,[3] and Wei Yu[4]

[1] *School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan 430073, China*
[2] *Business School, Sias International University, Xinzheng 451150, China*
[3] *School of Energy and Power Engineering, Wuhan University of Technology, Wuhan 430070, China*
[4] *State University of New York at Binghamton, Binghamton, NY 13902, USA*

Correspondence should be addressed to Xia Zhu; challengezhuzhu@163.com

Received 17 October 2015; Accepted 7 February 2016

Academic Editor: Herve G. E. Kadji

With the development of deep submergence technology, submarine is widely used in many aspects as marine analysis and detection of marine resources. For the reason of strong nonlinearity and coupling in submarine exercise, it is difficult to get satisfactory control effect by conventional control method. In order to control objective of stable straight-line suspension underwater, how to control the change of rudder angle to stabilize attitude and improve the control performance is researched from feature analysis to submarine. Aiming at improving the global stability, kinetic character of straight-line suspension movement underwater is analyzed and modeled firstly, and model of nonlinear relationship about change of rudder angle and attitude is built then. Based on the conditions of global stability asymptotically of submarine tracking control underwater and the physical significance of tracking control by nonlinear proportion differential, a controller is designed for controlling horizontal rudder angle and vertical rudder angle by dynamic feedback, which achieve the balance of tracking controlling both in local and global and guarantee global stable convergence asymptotically. At last, the stability, effectiveness, and global convergence of controller are proved by the simulation experiment.

1. Introduction

During the stable operation under the water, the submarine would move horizontally between two points on purpose for some time. In order to improve the stability and to reduce the energy consumption and cost, the submarine maintains horizontal and straight navigation through operation facilities' control. The autopilot is the indispensable crucial equipment in the system operated by submarine. The traditional course control cannot meet the requirement of straight navigation, so it is quite meaningful to research on the horizontal and straight navigation via autocontrolling. For the feature of inertance, time-lag, and nonlinearity during the process of motion under the water, as well as the impact of environment interference including model parameters and storm disturbing, it is very difficult to control the submarine tracing. Therefore, the research on submarine tracing control has attracted extensive concern from both academia and industries.

The straight track control of the surface ships is a common manner of wake control, which has been widely researched on [1]. At present, the study focuses on the stability of track control via nonlinear feedback control [2], back-stepping technology [3], and feedback linearization [4] to ensure the control effect of global stability of the track. Aiming at straight unstable characteristics of large surface warships, [5] has researched the global asymptotic stability of direction PD control and presented the condition for stability.

Due to the vertical sideway and pitching, as well as inability to correct through GPS under the water, it is more

difficult to control the track of the submarine. The controlling equipment of submarine is more complex, and the system variables such as motion and power are nonlinear of each other. So only the nonlinear system control can accurately solve the aforementioned issues. For now, on account of nonlinear controlling system we usually utilize phase-plane technique, Lyapunov method, input/output stable method, approximation linearized method, and describing function method [6]. The research on differential geometry control method overcomes the limits from local linearization and small range of motion and achieves large range of analysis and integration for dynamic system control [7]. Based on the differential geometry theory, the nonlinear control systems theory implements the linearization of the nonlinear system via static state feedback and transformation of coordinates under certain conditions. The nonlinear system is decoupled and linear in the new state space through suitable diffeomorphism, as well as corresponding static or dynamic feedback [8]. The limitation of the research on submarine provides the development ideas of this paper.

Aiming at assurance for the global stability of the control system on submarine, we study the issues of how to stabilize the track control performance via rudder angles transformation during the project practice based on the features of the submarine. This paper primarily analyzes the dynamics features about the linear suspension motion of the submarine, models the nonlinear relationship between the rudder angle and the posture change, and designs the controller to dynamic feedback for the plane angle and the vertical rudder angle according to the conditions of global asymptotic stability controlled by the submarine track and the physical significance controlled by the differential on the nonlinear proportion of the track. Our methods effectively achieve the balance control both locally and globally to ensure the global asymptotic stable convergence.

2. Modeling the Submarine Motion

Since digital computers and devices are used to control ships, it is natural to model ships and their controllers as nonlinear sampled-data systems. In this brief, we extend straight-line trajectory tracking control of continuous-time underactuated ships with state feedback controllers [9] to that of sampled-data underactuated ships with both state and output feedback controllers. We introduce a straight line as a reference trajectory and a reference nonzero forward speed for the ship. On the basis of the Euler approximate models, we design both state and output feedback controllers.

The suspending motion of the submarine under the water is effected by gravity, impetus, and current force, which is complicated and capricious. If the motion of the submarine is to be controlled, it has to analyze the dynamical model of the submarine under the water. For the relative motion of submarine and the fluid, the water power changes constantly as the interaction between its motion state and the marine environment. The main gesture parameters are direction, heading, depth, rolling, and pitching. In the case of low speed and diminutive longitudinal trim, the submarine is disintegrated into the motions of four degrees of freedom in 2 planes.

(1) The motion above the water: we utilize the projection above the water of the submarine to analyze the direction, heading, and the remaining and the change of the horizontal speed without considering the change of floating, diving, or rolling.

(2) The motion of verticality: we study the depth, longitudinal section, and the remaining and change of the snorkeling speed from the view of longitudinal axis.

In the three-dimensional coordinate system under the water, we regard the submarine as a rigid body. The submarine motion with six degrees of freedom (the lateral displacement y, the longitudinal displacement x, the vertical displacement z, the heading angle α, the roll angle β, and the pitching angle γ) is modeled as

$$\begin{aligned}
\widehat{x} &= u\cos\alpha\cos\gamma - v\sin\alpha\cos\gamma,\\
\widehat{y} &= u\sin\alpha\cos\gamma + v\cos\alpha\cos\gamma,\\
\widehat{z} &= u\sin\gamma + v\sin\gamma,\\
\widehat{\alpha} &= \varepsilon,\\
\widehat{\beta} &= \xi,\\
\widehat{\gamma} &= \delta,
\end{aligned} \tag{1}$$

where u denotes the forward speed along the axle wire of rigid body, v denotes the sideway speed perpendicular to the axle wire of rigid body, ε denotes the degree of heading angle, ξ denotes the speed of roll rotation angle, and δ denotes the angular velocity of pitching angle. Considering that transverse speed approaches 0 in general case and the rolling of the submarine does not affect the track, we simplify formula (1) as

$$\begin{aligned}
\widehat{x} &= u\cos\alpha\cos\gamma,\\
\widehat{y} &= u\sin\alpha\cos\gamma,\\
\widehat{z} &= u\sin\gamma,\\
\widehat{\alpha} &= \varepsilon,\\
\widehat{\gamma} &= \delta.
\end{aligned} \tag{2}$$

Considering the effect of nonlinear items in the submarine model, we utilize second-order nonlinear equation of motion to describe the submarine motion and the effect of rudder. The motion equation in the horizontal plane is denoted as

$$T_1\widehat{\alpha} + H_A(\widehat{\alpha}) = K_1\theta_1, \tag{3}$$

where θ_1 denotes the level control rudder angle, T_1 denotes the operation performance parameter of navigation, K_1 denotes the coefficient of hydroplane steering in submarine, and both T_1 and K_1 are greater than 0. H_A reflects the

mechanical control and hysteresis feature of the state, which is indicated as

$$H_A(\widehat{\alpha}) = m_1\widehat{\alpha}^3 + n_1\widehat{\alpha}, \tag{4}$$

where m_1 is the coefficient. When $n_1 = 1$, the submarine performs the stability in straight line; when $n_1 = -1$, the submarine performs instability in straight line, and the motion under the water presents instable state in straight line. Therefore we choose $n_1 = -1$.

The motion equation in the vertical plane is denoted as

$$T_2\widehat{\gamma} + H_B(\widehat{\gamma}) = K_2\theta_2, \tag{5}$$

where θ_2 denotes the vertical control rudder angle, T_2 denotes the operation performance parameter of snorkeling and diving, K_2 denotes the coefficient of steering effect in vertical rudder, and both T_2 and K_2 are greater than 0. H_B reflects the mechanical control and hysteresis feature of the state, which is indicated as

$$H_B(\widehat{\gamma}) = m_2\widehat{\gamma}^3 + n_2\widehat{\gamma}. \tag{6}$$

We choose $n_2 = -1$ as well.

Since the longitudinal displacement along the forward direction does not affect the straight navigation, we ignore the analysis of longitudinal displacement x. Integrating (2) into (6), we nonlinearly model the straight instable submarine under the water as follows:

$$\begin{aligned}
\widehat{y} &= u\sin\alpha\cos\gamma,\\
\widehat{z} &= u\sin\gamma,\\
\widehat{\alpha} &= \varepsilon,\\
\widehat{\gamma} &= \delta,\\
\widehat{\varepsilon} &= -\frac{n_1}{T_1}\varepsilon - \frac{m_1}{T_1}\varepsilon^3 + \frac{K_1}{T_1}\theta_1,\\
\widehat{\delta} &= -\frac{n_2}{T_2}\delta - \frac{m_2}{T_2}\delta^3 + \frac{K_2}{T_2}\theta_2,
\end{aligned} \tag{7}$$

where the target of the controller in the automatic pilot in regard to straight navigation under the water is to assure that the offset displacement y, the vertical displacement z, the bow declination α, and the pitching angle γ are all approaching zero. Then we achieve controlling the straight navigation.

On account of the nonlinear model of (7), we design the nonlinear state feedback control rate as follows:

$$\begin{aligned}
\theta_1 &= -k_D'(\alpha + f(y)) - k_D'\cdot\varepsilon,\\
\theta_2 &= -k_P''(\gamma + g(z)) - k_D''\cdot\delta.
\end{aligned} \tag{8}$$

On account of the closed-loop system consisted by (7) and (8), the necessary and sufficient conditions of global asymptotic stability for the submarine under the water are as follows:

(1) when $|y| \to \infty$, $\int_0^y \sin(f(y))dy \to \infty$;

(2) when $y \neq 0$, $\sin(f(y))\cos(g(z))y > 0$;

(3) $f'(y) > 0$, $g'(z) > 0$;

(4) $k_P' > 0$, $k_D' > -n_1/K_1$, $m_1 \geq 0$;

(5) as to arbitrary $y \in R$, we have $(Kk_D'/T_1 + n/T_1) > \sup(uf'(y))$.

3. Local Optimal Control Analytical Solution in Zero Equilibrium Point of Submarine Track

Formula (7) decomposes to a linear expansion in zero equilibrium point and retains one equation term; new formula can be deduced as follows:

$$\begin{aligned}
\widehat{y} &= u\cos(\alpha)\alpha\cos(\gamma),\\
\widehat{z} &= u\cos(\gamma)\gamma,\\
\widehat{\alpha} &= \varepsilon,\\
\widehat{\gamma} &= \delta,\\
\widehat{\varepsilon} &= -\frac{n_1}{T_1}\varepsilon - 3\frac{m_1}{T_1}(0)^2\varepsilon + \frac{K_1}{T_1}\theta_1,\\
\widehat{\delta} &= -\frac{n_2}{T_2}\delta - 3\frac{m_2}{T_2}(0)^2\delta + \frac{K_2}{T_2}\theta_2.
\end{aligned} \tag{9}$$

Various deviation angles are very small when submarine is sailing in straight line, and we can consider $\alpha \approx 0$, $\gamma \approx 0$, $\varepsilon \approx 0$, and $\delta \approx 0$, so formula (9) can be simplified as follows:

$$\begin{aligned}
\widehat{y} &= u\alpha,\\
\widehat{z} &= u\gamma,\\
\widehat{\alpha} &= \varepsilon,\\
\widehat{\gamma} &= \delta,\\
\widehat{\varepsilon} &= -\frac{n_1}{T_1}\varepsilon + \frac{K_1}{T_1}\theta_1,\\
\widehat{\delta} &= -\frac{n_2}{T_2}\delta + \frac{K_2}{T_2}\theta_2.
\end{aligned} \tag{10}$$

Let $y_1 = y/u$ and $z_1 = z/u$; assume $x = [y_1, y_2, \alpha, \gamma, \varepsilon, \delta]^T$ and $\theta = [\theta_1, \theta_2]^T$. It can be gotten from formula (10) as follows:

$$\widehat{x} = Ax + B\theta, \tag{11}$$

where

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -\frac{n}{T_1} & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{n}{T_2} \end{bmatrix},$$
$$B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \frac{K_1}{T_1} & 0 \\ 0 & \frac{K_2}{T_2} \end{bmatrix}. \tag{12}$$

According to mechanical property and unit of different rudder angle, in order to guarantee value invariability of weight coefficient φ in performance criterion function, it can describe relation between state variables y_1 and y_2 and control angles θ_1 and θ_2 by quadratic performance criterion function as follows:

$$J = \int_0^{\infty} \left(y_1^2 + z_1^2 + \frac{\varphi_1 180^2}{\pi^2}\theta_1^2 + \frac{\varphi_2 180^2}{\pi^2}\theta_2^2 \right) dt, \tag{13}$$

where the unit of rudder angles θ_1 and θ_2 is rad, φ_1 is the coefficient of horizontal rudder, and φ_2 is the coefficient of vertical rudder, which is set by experimenter. The smaller φ is, the higher the precision of tracking is, and the range of control angle is larger by this time. The larger φ is, the lower the precision of tracking is, the range of control angle is smaller by this time, and loss of steering engine reduces less. Value range of φ is in the convergence of [0.1, 10] usually.

Algebraic Riccati equation of formula (11) can be translated as follows:

$$A^T P + PA - PBR^{-1}B^T P + Q = 0. \tag{14}$$

So the feedback control rate of optimal state in zero equilibrium point of submarine track can be calculated as follows:

$$\begin{aligned} \theta_1 &= -k'_y y_1 - k'_\alpha \alpha - k'_\varepsilon \varepsilon, \\ \theta_2 &= -k''_z z_1 - k''_\gamma \gamma - k''_\delta \delta. \end{aligned} \tag{15}$$

4. Controller Design

According to eventually uniform boundedness principle, we have the following conclusion: only $\Psi(y_1, z_1) = (u(\sqrt{y_1^2 + z_1^2})/\sqrt{1 + y_1^2 + z_1^2}) - 1/4(\varphi_1^2 + \varphi_2^2) > 0$ is satisfied; the system is uniform and stable boundedness ultimately.

Define $x_1 = y_1$, $x_2 = \sqrt{y_1^2 + z_1^2}$; make

$$\begin{aligned} V_1 &= \frac{1}{2}x_1^2 + \frac{1}{2}\theta^T \alpha \theta + \frac{1}{2}\alpha^{-1}\varphi^2, \\ V_2 &= V_1 + \frac{1}{2}x_2^2 + \frac{1}{2}\theta^T \gamma \theta + \frac{1}{2}\gamma^{-1}\varphi^2. \end{aligned} \tag{16}$$

According to approximation of Mamdani fuzzy system structured by multiple fuzzy inference engine, singleton fuzzifier, centre average defuzzifier, and specific membership functions, define

$$\begin{aligned} F(x_1) &= f'', \\ x_2 &= \left[y, \frac{\partial \alpha}{\partial y}, \frac{\partial \gamma}{\partial z}, \theta \right]^T. \end{aligned} \tag{17}$$

Control parameters of straight-line track underwater can be as follows:

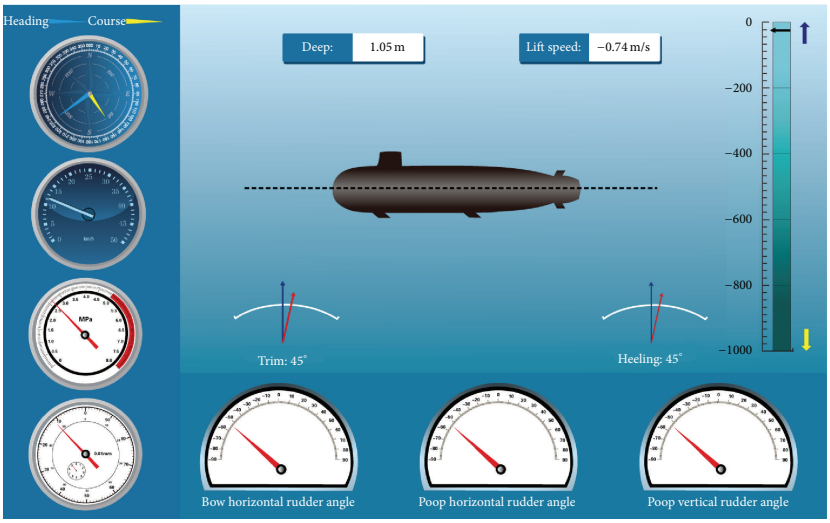
$$\begin{aligned} \theta_1 &= -k'_y \left(\alpha + \arctan\left(\frac{x_1}{k'_\alpha} \right) \right) - k'_\varepsilon \varepsilon, \\ \theta_2 &= -k''_z z_1 \left(\gamma + \arctan\left(\frac{x_2}{k'_\gamma} \right) \right) - k''_\delta \delta. \end{aligned} \tag{18}$$

5. Simulative Experiment

In this section we present an experimental study of our straight-line track control algorithms for detecting effectiveness and correctness of submarine underwater in simulation cabin. The experiment simulates the process of snorkeling and sailing through a virtual submarine system developed by *Dalian JXD Soft Ltd.*, which can realistically simulate 6-DOF motion and influence of ocean current environment of submarine underwater and water surface. In this experiment, the initial conditions of submarine are as follows: length is 149.5 meters, wide is 12.8 meters, draught is 12 meters, max rudder angle is 45°, the inertia link of motion model is used by a time constant as 5 s, the whole structure appears capsule shape, and the initial position is as follows: depth is 30 meters, heading angle is due north, pitching angle is 0.03, and speed is 12 knot. The operation interface of simulator is shown in Figure 1.

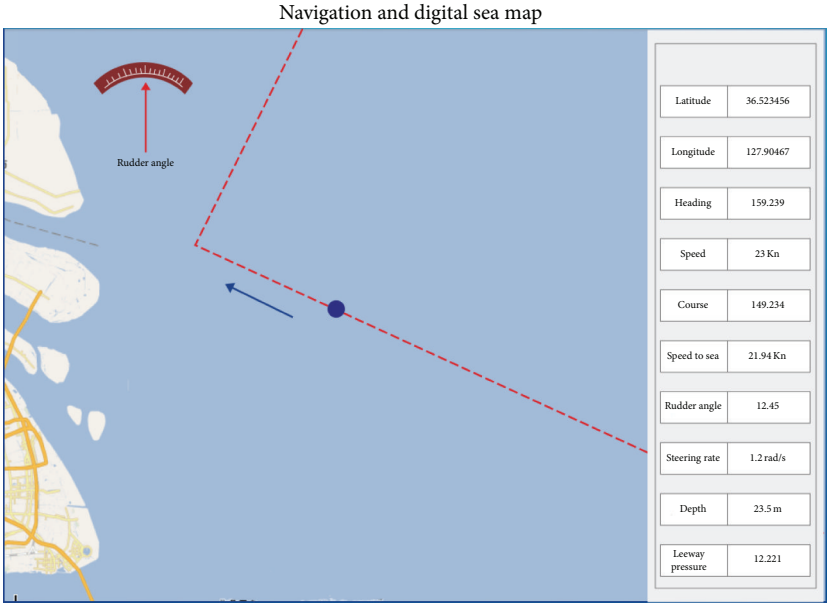
In this experiment, the horizontal plane and vertical rudder are controlled real-timely and dynamically by straight-line track automatic rudder controller designed in Section 4. The all track information as transverse offset, vertical offset, heading angle, heading angular velocity, pitching angle, and pitching angular velocity is recorded anytime real-timely. The sailing time lasts 90 minutes, and the result of records is shown in Figure 2.

The consult is shown in Figure 2, in the process of underwater sailing; the track offset of submarine is convergence in a small neighborhood by straight-line track automatic rudder controller as in Section 4. During the 90 minutes, the mean and variance of transverse offset are −0.02 and 0.45; the mean and variance of vertical offset are 0.07 and 0.72; the mean and variance of heading angle are 0.17 and 0.50; the mean and



(a) Sagittal operation interface of snorkeling

Navigation and digital sea map



(b) Horizontal operation interface of surface

Figure 1: Operation interface of simulator.

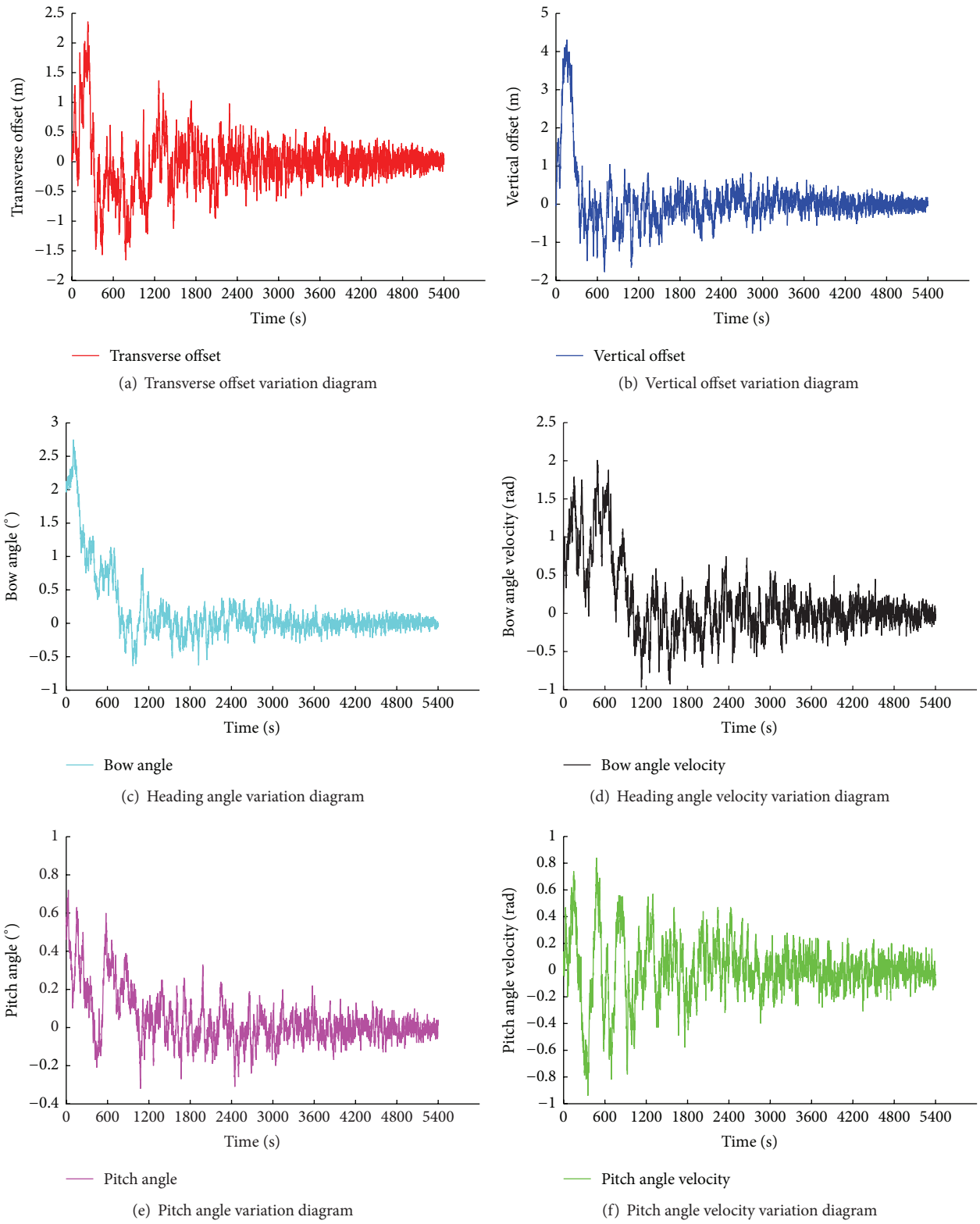
variance of heading angle velocity are 0.13 and 0.46, the mean and variance of pitch angle are −0.16 and 0.64; the mean and variance of pitch angle velocity are −0.15 and 1.14. The result shows that all the parameters are fluctuated in less offset and error precision requirement is satisfied.

In the result of simulation experiment, even if the offset to planned sea route in initial state is large, the submarine can be global asymptotic stability to planned sea route and sailing scheduled by straight-line track automatic rudder controller in this paper.

6. Conclusions

Affected by multiple factors such as gravity, power, ocean current, and external environment, the underwater trajectory of submarine shows such complicated characteristics of movement path and controlled feedback. Hence, the control problems of submarine are nonlinear in delayed and gradual conditions. As an essential ingredient of the submarine underwater operating system, the submarine control system achieved linear-suspended automatic navigation control in complicated surroundings, which enables the efficient and stable motion of submarine.

Against the calculating problem of rudder angle variation per instantaneous time interval under the global convergence of control varying, this paper demonstrates a nonlinear proportional differential based underwater controller to control the movement of submarine. By the dynamic analysis of the internal relations between controlling parameters, movement characteristics, and trajectory performance according to the



(a) Transverse offset variation diagram

(b) Vertical offset variation diagram

(c) Heading angle variation diagram

(d) Heading angle velocity variation diagram

(e) Pitch angle variation diagram

(f) Pitch angle velocity variation diagram

Figure 2: Track parameters during voyage period.

historical and real-time attitude data, the controller can solve the rudder angle control instruction per instantaneous time interval and adjust the instruction in real time, which enables the self-adaption control of the straight-line navigation of submarine.

Our experiment shows that the nonlinear proportional differential method achieved the equilibrium and stability of vertical and horizontal surfaces and enables fast global stabilization in some time, which reached the aim of underwater straight-line navigation. Our research provides the theoretical base for underwater automatic driving of submarine, which has certain theory value and practical significance.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This paper is supported by the National Natural Science Foundation of China (no. 61272109) and the Soft Science Research Program of Henan Province of China (no. 142400410890). The authors would like to thank the support by State Key Laboratory of Software Engineering in Wuhan University.

References

[1] T. I. Fossen, "Recent developments in ship control systems design," in *World Superyacht Review*, pp. 115–116, Sterling Publications Limited, London, UK, 2000.

[2] F. Mazenc, K. Pettersen, and H. Nijmeijer, "Global uniform asymptotic stabilization of an underactuated surface vessel," *IEEE Transactions on Automatic Control*, vol. 47, no. 10, pp. 1759–1762, 2002.

[3] Z.-P. Jiang, "Global tracking control of underactuated ships by Lyapunov's direct method," *Automatica*, vol. 38, no. 2, pp. 301–309, 2002.

[4] K. D. Do, Z. P. Jiang, and J. Pan, "Robust global stabilization of underactuated ships on a linear course: state and output feedback," *International Journal of Control*, vol. 76, no. 1, pp. 1–17, 2003.

[5] D. Neuffer and D. H. Owens, "Global stabilization of unstable ship dynamics using PD control," in *Proceedings of the 30th IEEE Conference on Decision and Control*, pp. 519–520, IEEE, December 1991.

[6] X. Cai and M. Krstic, "Nonlinear control under wave actuator dynamics with time- and state-dependent moving boundary," *International Journal of Robust and Nonlinear Control*, vol. 25, no. 2, pp. 222–251, 2015.

[7] X. Wei and J. E. Mottershead, "Block-decoupling vibration control using eigenstructure assignment," *Mechanical Systems and Signal Processing*, vol. 74, pp. 11–28, 2016.

[8] K. Baibeche and C. H. Moog, "Input-state feedback linearization of single-input nonlinear time-delay systems," in *Proceedings of the IEEE European Control Conference (ECC '14)*, pp. 460–465, Strasbourg, France, June 2014.

[9] E. Børhaug, A. Pavlov, E. Panteley, and K. Y. Pettersen, "Straight line path following for formations of underactuated marine surface vessels," *IEEE Transactions on Control Systems Technology*, vol. 19, no. 3, pp. 493–506, 2011.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 4126123, 8 pages
http://dx.doi.org/10.1155/2016/4126123

Research Article

Large-Scale Computations of Flow around Two Cylinders by a Domain Decomposition Method

Hongkun Zhu,[1] Qinghe Yao,[1] and Hiroshi Kanayama[2]

[1] *School of Engineering, Sun Yat-sen University, Guangzhou 510275, China*
[2] *Department of Mathematical and Physical Sciences, Japan Women's University, Tokyo 112-8681, Japan*

Correspondence should be addressed to Qinghe Yao; yaoqhe@mail.sysu.edu.cn

Received 25 September 2015; Revised 31 December 2015; Accepted 28 February 2016

Academic Editor: Uchechukwu E. Vincent

A parallel computation is applied to study the flow past a pair of cylinders in tandem at Reynolds numbers of 1000 by Domain Decomposition Method. The computations were carried out for different sets of arrangements at large scale. The modeling by domain decomposition was validated by comparing available well-recognized results. Two cylinders with different diameters were further investigated; for different diameter ratios, the wake width ratio and some properties of the critical space ratio that dominates the flow regime were discovered. This result has important implications on future industrial application efforts as well as codes and standards related to the two-cylinder structure.

1. Introduction

Cylinder is one of the most common structures in engineering, such as piers and chimney, the struts of offshore platform, the pipe of condenser, and more. When fluid flows through the cylinders, the shedding vortex may cause interference effect. This effect may lead to the vibration of the cylinder structures or fatigue of the materials and as a result, the structure may be destroyed. Due to its importance in engineering applications, flow past two cylinders has been studied by experimental and numerical investigations for several decades [1]. As early as 1977, Zdravkovich classified the characteristic of flow past two cylinders in tandem into three regimes at low Reynolds number (Re). For more complicated applications, Wu et al. [2] studied two cylinders in tandem with wind tunnel and water tunnel at the Re of 1000. The flow visualization in the water tunnel showed the existence of streamwise vortices in spanwise direction. To make out the relationship between different arrangements of two cylinders and Re, Mittal et al. [3] studied incompressible flows past two cylinders in tandem and staggered arrangements (Re = 100 and Re = 1000) by a stabilized finite element method (FEM). Jester and Kallinderis [4] investigated the incompressible flow about fixed cylinder pairs numerically and cylinder arrangements include tandem, side-by-side, and staggered at Reynolds numbers of 80 and 1000. Hysteresis effects and bistable biased gap flow in tandem arrangements were reproduced. Different combinations of Re and arrangements of cylinders have been reported by many [5–7], as is reported in [8–17] and summarized by Sumner [1]; most of the early researches on this problem contribute to the flow structures induced by different spacing ratios (L/D), Re and variant cylinder shapes.

Recently, there has been a lot of interest on the diameter ratio (d/D) of two cylinders in tandem. Zhao et al. studied turbulent flow past two cylinders with different diameters numerically. The hydrodynamic force and vortex shedding characteristics were proved to depend on the relative position of small cylinders around the main cylinder [18]. Mahbub Alam and Zhou investigated the Strouhal number, hydrodynamic forces and flow structures, and vortex shedding frequency of flow past two cylinders in tandem with different diameters in wind tunnel [19]. The diameter of upstream cylinder varied from 0.24 to 1 of the downstream cylinder diameter and the distance between two cylinders remains 5.5 times of the diameter of the upstream cylinder. Ding-Yong et al. conducted the simulation of different spacing ratios and different diameter ratios at Re = 200 using Fluent [20]. Besides two cylinders, Zhang et al. determined the influence of the diameter ratio on the flow past three

cylinders in two dimensions by applied finite element method [21].

As far as we know, the flow structure of two cylinders in tandem affected by the changing of diameter ratio and spacing ratio is still uncertain, and the computation scale of numerical experiments published research is very limited. To investigate the flow regime more concretely and more subtly, large-scale simulations by Domain Decomposition Method (DDM), which is considered to have better accuracy and less time cost when comparing with conventional methods, are implemented [22, 23]. The large-scale modeling was validated by comparing with others' reports for two cylinders of the same diameter with different spacing ratios. Moreover, to study its influence to the flow past two cylinders in tandem more comprehensively, the flow structures at specified diameter ratios ($d/D = 0.5$ and $d/D = 1$) with different spacing ratios (L/D: 3–6) are investigated. The critical spacing ratio that affects the flow regime is expected to be determined for different diameter ratios.

This paper is organized as follows: Section 2 introduces the governing equations to be solved as well as the domain decomposition method (DDM). In Section 3, the models and the boundary conditions are described in detail. Numerical results for different spacing ratio and different diameter ratios ($d/D = 1$ and $d/D = 0.5$) are present and compared with others' work in Section 4. At last, Section 5 draws the conclusions of this work.

2. Formulations

2.1. Governing Equations. Let $\partial\Omega$ be the boundary of a three-dimensional polyhedral domain Ω. $H^1(\Omega)$ is the first order Sobolev space and $L^2(\Omega)$ is the space of 2nd power summable functions on Ω. Under the assumption that the flow field is incompressible, viscous, and laminar, the solving of the model can be summarized as finding $(u, p) \in H^1(\Omega)^3 \times L^2(\Omega)$ such that for any $t \in (0, T)$, the following set of equations hold [24]:

$$\begin{gathered}\partial_t u + (u \cdot \nabla) u - \frac{1}{\rho} \nabla \cdot \sigma(u, p) = \frac{1}{\rho} f, \\ \nabla \cdot u = 0, \\ \text{in } \Omega \times (0, T),\end{gathered} \tag{1}$$

where u is the velocity [m/s]; p is the pressure [Pa]; ρ is the density (const.) [kg/m^3]; f is the body force [N/m^3]; $\sigma(u, p)$ is the stress tensor [N/m^2] defined by

$$\begin{gathered}\sigma_{ij}(u, p) \equiv -p\delta_{ij} + 2\mu D_{ij}(u), \\ D_{ij}(u) \equiv \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \\ i, j = 1, 2, 3,\end{gathered} \tag{2}$$

with the Kronecker delta δ_{ij} and the viscosity μ [kg/(m·s)].

An initial velocity u_0 is applied in Ω at $t = 0$. Dirichlet boundary conditions

$$u = \hat{u} \quad \text{on } \Gamma \times (0, T) \tag{3}$$

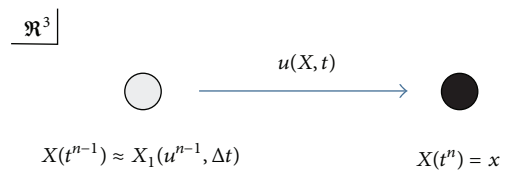


Figure 1: A characteristic finite element scheme.

and Neumann boundary conditions

$$\sum_{j=0}^{3} \sigma_{ij} n_j = 0 \quad \text{on } \partial\Omega \setminus \Gamma \times (0, T) \tag{4}$$

are also applied, where $\Gamma \subset \partial\Omega$ and n is the outward normal direction to $\partial\Omega$.

2.2. Characteristic Finite Element Scheme. Using the definition

$$\begin{gathered}t^n \equiv n\Delta t, \\ N_T \equiv \left[\frac{T}{\Delta t}\right],\end{gathered} \tag{5}$$

a characteristic finite element scheme approximates the material derivative in (1) at t^n as follows: [23]

$$\partial_t u + (u \cdot \nabla) u \approx \frac{u^n - u^{n-1}\left(X_1\left(u^{n-1}, \Delta t\right)\right)}{\Delta t}, \tag{6}$$

where $X_1(\cdot, \cdot)$ is a position function; see Figure 1.

Let $\mathfrak{T}_h \equiv \{K\}$ denote a triangulation of Ω consisting of tetrahedral elements and the subscript h denotes the representative length of the triangulation. The finite element spaces are defined as follows:

$$\begin{gathered}X_h \equiv \left\{v_h \in C^0(\Omega)^3;\ \ v_h|_K \in P_1(K)^3,\ \forall K \in \mathfrak{T}_h\right\}, \\ M_h \equiv \left\{q_h \in C^0(\Omega);\ \ q_h|_K \in P_1(K),\ \forall K \in \mathfrak{T}_h\right\}, \\ V_h(\hat{u}) \equiv \left\{v_h \in X_h;\ v_h(P) = \hat{u}(P),\ \forall P \in \Gamma\right\}, \\ V_h \equiv V_h(0), \\ Q_h = M_h,\end{gathered} \tag{7}$$

where the $C^0(\Omega)^3$ in (7) denotes the continuous function on Ω in three dimensions. $P_1(K)^3$ denotes the first order polynomial defined by K in three dimensions.

Note that piecewise linear interpolations are employed for velocity and pressure (see (7) and Figure 2), which does not provide a sufficient condition to link the velocity and pressure space; therefore, the so-called inf-sup condition [25] should be satisfied. In previous work [26], a penalty Galerkin least-squares (GLS) stabilization method for pressure [27]

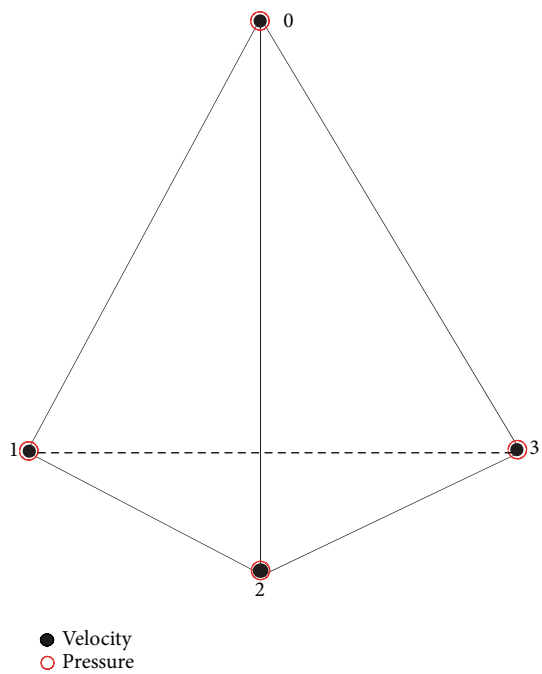


Figure 2: A tetrahedron element.

was employed and it was found difficult to be applied for the simulation of complex flows; especially when the flow is very turbulent, the scheme becomes easy to diverge. A stabilization technique for the saddle point problem is employed and the finite element scheme for (1) read as follows [22].

Find $\{(u_h^n, p_h^n) \in V_h(g) \times Q_h\}_{n=1}^{N_T}$ such that

$$\begin{aligned}\left(\frac{u_h^n - u_h^{n-1}\left(X_1\left(u_h^{n-1}, \Delta t\right)\right)}{\Delta t}, v_h\right) &+ a\left(u_h^n, v_h\right) \\ &+ b\left(v_h, p_h^n\right) + b\left(u_h^n, q_h\right) \\ &+ \frac{1}{\rho^2}\sum_{K\in\mathfrak{T}_h} \tau_K \left(\nabla p_h^n, -\nabla q_h\right)_K \\ &= \frac{1}{\rho}\left(f, v_h\right) + \frac{1}{\rho^2}\sum_{K\in\mathfrak{T}_h} \tau_K \left(f, -\nabla q_h\right)_K .\end{aligned} \tag{8}$$

Let $(\cdot,\cdot)$ defines the L_2 inner product; the continuous bilinear forms a and b in (8) are introduced by

$$a(u, v) \equiv \frac{2\mu}{\rho}\left(D(u), D(v)\right), \tag{9}$$

$$b(v, q) \equiv -\frac{1}{\rho^2}\left(\nabla \cdot v, q\right), \tag{10}$$

respectively. The following stabilization parameter is employed:

$$\tau_K = \min\left\{\frac{\Delta t}{2}, \frac{h_K}{2\left\|u_h^{n-1}\right\|_\infty}, \frac{\rho h_k^2}{24\mu}\right\}, \tag{11}$$

where h_K denotes the maximum diameter of an element K and $\|\cdot\|_\infty$ is the maximum norm.

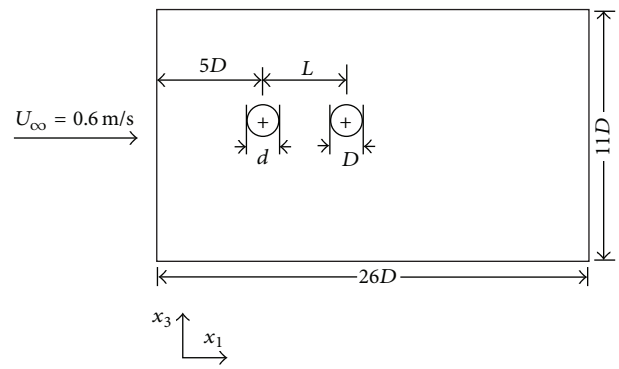


Figure 3: Two cylinders in tandem.

A weak coupling of finite element scheme in (8) is applied and the element searching algorithm for Lagrange-Galerkin method only needs to be implemented once in each nonsteady loop [23].

2.3. Domain Decomposition Method. The calculating area is a cuboid. The left side of the model is the entrance and the right side is the free outlet. The inlet is placed $5D$ to the upstream cylinder and total length is $26D$. The nonslip boundaries are placed $5.5D$ above and below the cylinders. The height is $4D$; see Figure 3 for a two-dimensional projection of the three-dimensional model.

In order to investigate the relationship between them by large-scale simulation, a parallel computation by DDM [22, 23, 28, 29], which is considered to have better accuracy and less time cost, is implemented in this work. In the domain decomposition system, the whole computation domain is split into N nonoverlapping parts, where N is the number of threads; in each smaller part, an FEM process is preceded. For the comparison of efficiency of DDM and FEM, please see [23].

A Linux cluster (Intel Xeon E5606@2.13 GHz × 44, LV RDIMM 12 GB@1333 MHz × 44) in Sun Yat-sen University was used for this computation. For each computation model, 176 threads were created, as can be seen in Figure 4.

For each computational model, nonsteady time step is set to 0.01 s and the number of total simulation time is 10 s. The maximum number of Degrees Of Freedom (DOF) is up to 16.6 million and it takes the cluster about 5.8 hours. Details of the computations for each model can be found in Table 1.

3. Modeling and Boundary Conditions

Air is assumed to be the fluid in this work. L is the distance between the two centers of the cylinders. The diameter of upstream cylinder is d and the diameter of downstream cylinder is D ($D = 0.025$ m). The time step is 0.01 s and total computation time for each model is 10 s.

The kinematic viscosity of the fluid is $0.000015\,\mathrm{m}^2/\mathrm{s}$ at constant temperature and Re = 1000. The velocity of the coming flow is 0.6 m/s. The air flows into the model with a constant horizontal velocity from the left to the right. The vertical velocity is prescribed to zero.

Table 1: Computation information.

Name	Spacing ratio (L/D)	Diameter ratio (d/D)	Number of elements	Number of DOF	Nonsteady loops	Computation time (h)
Model 1–1	2	1	9,420,393	14,130,590	1000	5.15
Model 1–2	2.15	1	9,466,412	14,211,123	1000	5.00
Model 1–3	2.5	1	9,566,559	14,390,891	1000	4.35
Model 1–4	3	1	9,710,157	14,649,874	1000	4.85
Model 2–1	3	0.5	6,908,469	9,309,739	1000	3.10
Model 2–2	3	1	9,710,157	14,855,135	1000	4.68
Model 2–3	4	0.5	10,146,357	15,654,964	1000	5.32
Model 2–4	4	1	9,997,736	15,312,582	1000	4.74
Model 2–5	5	0.5	10,428,101	16,114,294	1000	4.97
Model 2–6	5	1	10,283,886	15,785,861	1000	5.45
Model 2–7	6	0.5	10,715,636	16,597,006	1000	5.75
Model 2–8	6	1	10,565,696	16,259,744	1000	5.53

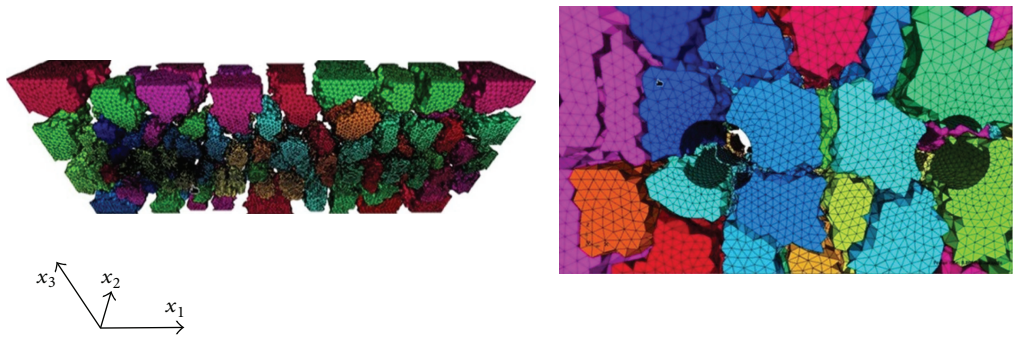


Figure 4: The domain decomposition of a model: the 176 parts and local details.

The model of two cylinders in tandem with different diameter is shown in Figure 5. The diameter of the upstream cylinder decreases to $0.5D$. The diameter of the downstream cylinder remains the same; d/D is 0.5.

4. Result and Discussion

This section presents the results of numerical simulations for two cylinders in tandem with different spacing ratios (L/D) and diameter ratios (d/D). To confirm the results, the visualization of computational results was conducted and it showed the details of flow field. Micro AVS by CYBERNET SYSTEMS was used to plot Figures 5–14.

4.1. Validation. In order to validate the current modeling, numerical results of two cylinders of the same diameter with different spacing ratios (L/D) were compared to well-recognized results. Figures 6 and 7 show the comparison of current results with the numerical results reported by Jester and Kallinderis [4].

As is shown in Figure 6, there is a recirculation between the two cylinders at $L/D = 2$. The vortex shedding occurs only behind the downstream cylinder.

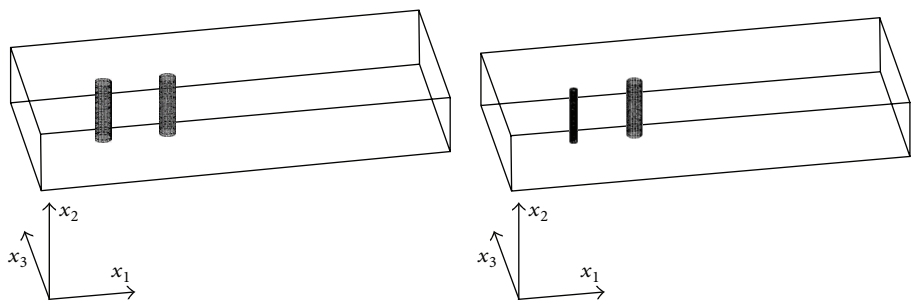
At $L/D = 2.15$, both reattachment and two vortex streets occur between two cylinder, which presents a bistable state [4]. When reattachment occurs, the vortices shed from the upstream cylinder fluctuate and reattach to the downstream cylinder; see Figure 7. Recirculation is observed between the two cylinders and vortex shedding is observed behind the downstream cylinder all the time at steady state.

At $L/D = 2.5$, the vortex shedding occurs behind both upstream and downstream cylinders, as is shown in Figure 8. The vortex sheds from the upstream cylinder sharply and impinges to the downstream cylinder.

The computational results are also validated by comparing with the experiment conducted by Wu et al. [2], and the comparison of the vortex contour is shown in Figure 9.

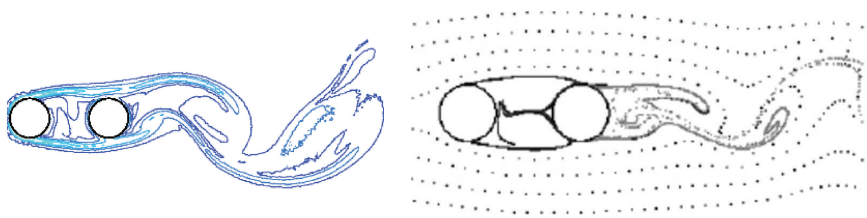
As can be seen from Figure 9, the numerical and experimental results agree with each other very well: the vortices separate from the upstream cylinder reattach to the downstream cylinder. The vortex shedding occurs only after the downstream cylinder. There is no independent vortex street between two cylinders.

The computational results present the same characteristic and tendency with recognized results of other researchers. The characteristics of the flow and the vortex shedding are both performed as expected. The numerical experiments also convinced us that the parallel computation with Domain Decomposition Method has good accuracy and efficiency in large-scale simulation, which encouraged us to move forward to investigate the effect of the diameter ratio.



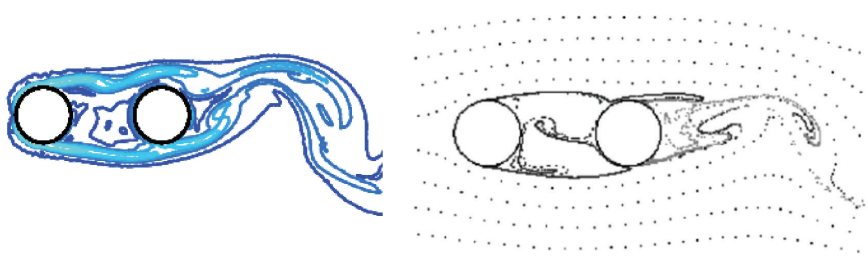
(a) (b)

Figure 5: Two cylinders in tandem with different diameter ratios. ((a) $d/D = 1$; (b) $d/D = 0.5$).



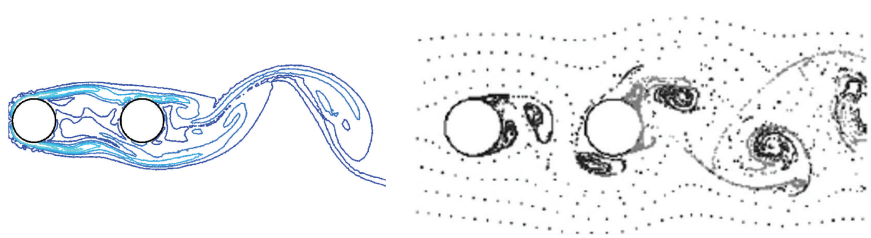
(a) (b)

Figure 6: Streamlines of Model 1–1 at $t = 8.00$ s. ((a) Current results; (b) Jester and Kallinderis).



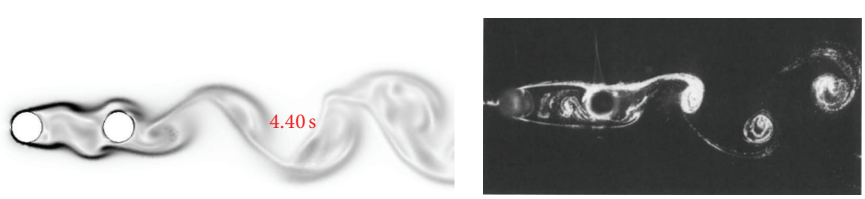
(a) (b)

Figure 7: Streamlines of Model 1–2 at $t = 6.40$ s. ((a) Current results; (b) Jester and Kallinderis).



(a) (b)

Figure 8: Streamlines of Model 1–3 at $t = 8.00$ s. ((a) Current results; (b) Jester and Kallinderis).



(a) (b)

Figure 9: Contour of Model 1–1 at $t = 4.40$ s. ((a) Current results; (b) Wu et al.).

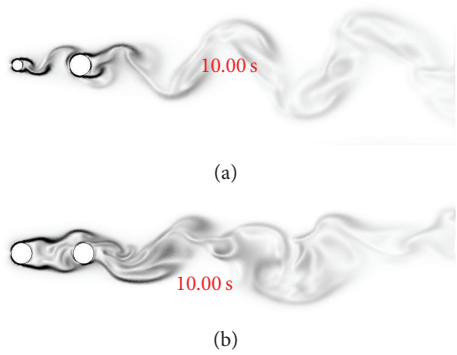


Figure 10: Contour at $L/D = 3$, $t = 10.00$ s. ((a) Model 2–1; (b) Model 2–2).

4.2. The Effect of Diameter Ratio. According to the research conducted by Ding-Yong et al. [20], the diameter ratio has strong effects on the characteristic of flow field around two cylinders in tandem. In the research, the downstream cylinder had smaller diameter than the upstream one. With the increasing of the diameter ratio (d/D), the critical spacing ratio increases. Mahbub Alam and Zhou also investigated flow past two cylinders in tandem with different diameters experimentally in wind tunnel [19]. The diameter of upstream cylinder varied from 0.24 to 1 of the downstream cylinder diameter. However, the spacing ratio kept 5.5 times of upstream cylinder diameter.

To determine the influence of the diameter ratio (d/D) to the flow field more comprehensively, this paper conducted the study by decreasing the diameter of the upstream cylinder to half of the downstream cylinder. The space between two cylinders remains the same as the corresponding simulation at $d/D = 1$. The visualization of the flow field shows the characteristic between the gap and vortices street. In all arrangements, the vortex shedding occurs behind both upstream and downstream cylinders. However, the volume of the vortices and the complexity of wake flow are different.

As can be seen from Figure 10, at $L/D = 3$, the vortices are smaller and but the wake behind downstream cylinder is wider at $d/D = 0.5$ compared with $d/D = 1$. With $d/D = 1$, the vortices shed from the upstream cylinder sharply and impinged to the downstream cylinder. At $d/D = 0.5$, the interference on the upstream cylinder is less than that at $d/D = 1$.

In Figure 11, at $L/D = 4$, the vortices are smaller at $d/D = 0.5$ compared with $d/D = 1$. The wake behind the downstream cylinder shares about the same area for $d/D = 0.5$ and $d/D = 1$. The interference on the upstream cylinder is smaller at $d/D = 0.5$ than that at $d/D = 1$.

At $L/D = 5$, the vortices and wake behind the downstream cylinder appears smaller and narrower at $d/D = 0.5$ compared with $d/D = 1$. The interference on the upstream cylinder is much smaller at $d/D = 0.5$ than that at $d/D = 1$; see Figure 12.

At $L/D = 6$, the vortices and wake behind the downstream cylinder are much smaller and narrower at $d/D = 0.5$ compared with $d/D = 1$. The interference on the upstream cylinder is negligible at $d/D = 0.5$ than that at $d/D = 1$; see Figure 13.

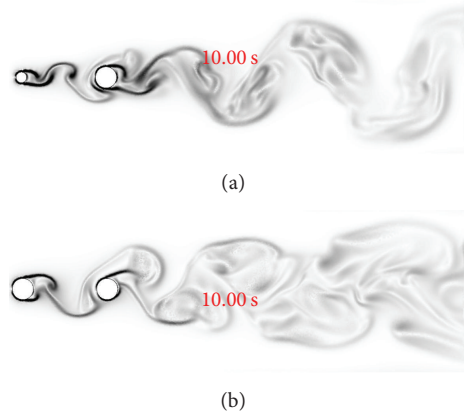


Figure 11: Contour at $L/D = 4$, $t = 10.00$ s. ((a) Model 2–3; (b) Model 2–4).

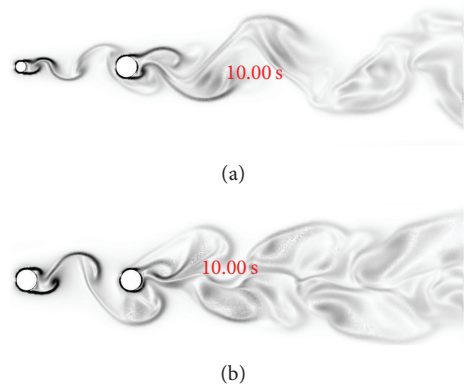


Figure 12: Contour at $L/D = 5$, $t = 10.00$ s. ((a) Model 2–5; (b) Model 2–6).

Overall, the wake after the downstream cylinder at $d/D = 0.5$ is clearer than that at $d/D = 1$. The vortices that shed from the smaller upstream cylinder impinge to the downstream cylinder and lead to a simpler behavior of wake. It is supposed that there exists a critical spacing ratio dominating the flow regime. Additionally, the different diameter ratio also gave rise to the change of critical spacing ratio. From Figures 10–13, it is known that

(1) When the spacing ratio is less than the critical value, there are only few vortices shedding after the downstream cylinder but no vortex shedding occurs after the upstream cylinder; see Figure 10(b).

(2) When the spacing ratio reaches the critical value, the vortex shedding occurs after both cylinders; see Figure 11(b).

(3) When the spacing ratio keeps increasing, the vortices shed from the upstream cylinder will not attach to the downstream cylinder, and there will be two independent vortex streets; see Figure 13(a).

It can be seen that the critical spacing ratio increases with the increasing of diameter ratio. Moreover, the critical

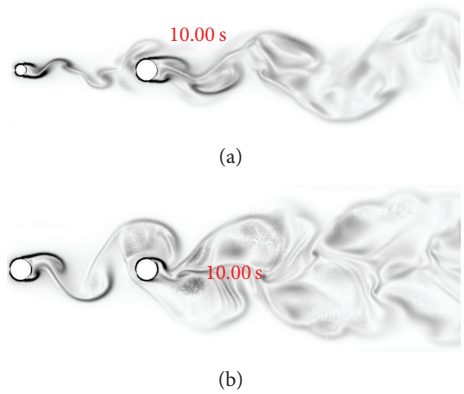


Figure 13: Contour at $L/D = 6$, $t = 10.00$ s. ((a) Model 2–7; (b) Model 2–8).

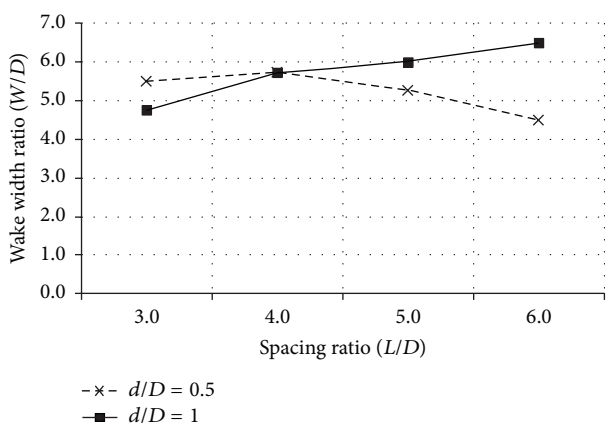


Figure 14: Wake width ratio versus spacing ratio.

spacing ratio at $d/D = 1$ is between 3 and 4, and the critical spacing ratio at $d/D = 0.5$ is smaller than 3. This phenomenon coincides with what we know: decreasing d/D leads to the narrower wake and smaller vortices after the upstream cylinder [19]; before L/D reaches the critical value, the vortex street is suppressed between the two cylinders [30]. To explain the mechanism behind this, wake width (W) is measured for Figures 10–13 and wake width ratio is defined by W/D; see Figure 14.

At $d/D = 0.5$, the wake and vortices behind upstream cylinder are narrower and smaller, which reduces the interference between the two cylinders, leading to the narrower wake behind downstream cylinder as well. At $d/D = 1$ and $L/D = 3$, the vortex street is suppressed behind the upstream cylinder. Consequently, the wake width is narrower compared to the case at $d/D = 0.5$, $L/D = 3$. However, the wake behind the upstream cylinder is no longer suppressed when keep increasing the spacing ratio. Thus, the interference increases. At $d/D = 1$, the wake behind the upstream cylinder is wider than at $d/D = 0.5$. The increasing of spacing ratio will even cause wider wake behind the downstream cylinder due to stronger interference, until two independent vortex streets occur.

5. Conclusions

This paper tested the parallel computation with Domain Decomposition Method in large-scale computation by simulating flow past two cylinders in tandem. The arrangements of spacing ratio L/D between 2 and 6 with diameter ratio d/D equal to 1 and 0.5 were studied.

(a) By applying this method to simulate flow past two cylinders in tandem and comparing the results with others' work, it showed that the results have better accuracy and the computation costs less time. The credibility and the viability were verified. The method will have broad application in large-scale computation in the future.

(b) The diameter ratio plays an important role in flow past two cylinders in tandem. The upstream cylinder with smaller diameter will decrease the vortex volume and the interference between two cylinders. With a decreasing of d/D, the vortices behind the upstream cylinder appear more but smaller. Meanwhile, with an increasing spacing ratio L/D, the wake becomes narrower at $d/D = 0.5$ but becomes wider at $d/D = 1$.

(c) The critical spacing ratio increases with the diameter ratio increasing. Moreover, the critical spacing ratio at $d/D = 1$ is between 3 and 4, and it is smaller than 3 at $d/D = 0.5$.

In real word engineering, the structure of cylinders is more complicated, which should be considered in future study.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

The authors would like to thank Mr. Cansen Zhen for some of numerical experiments at early stage. The National Science Foundation of China, Grants 11202248 and 11572356, supports this work.

References

[1] D. Sumner, "Two circular cylinders in cross-flow: a review," *Journal of Fluids and Structures*, vol. 26, no. 6, pp. 849–899, 2010.

[2] J. Wu, L. W. Welch, M. C. Welsh, J. Sheridan, and G. J. Walker, "Spanwise wake structures of a circular cylinder and two circular cylinders in tandem," *Experimental Thermal and Fluid Science*, vol. 9, no. 3, pp. 299–308, 1994.

[3] S. Mittal, V. Kumar, and A. Raghuvanshi, "Unsteady incompressible flows past two cylinders in tandem and staggered arrangements," *International Journal for Numerical Methods in Fluids*, vol. 25, no. 11, pp. 1315–1344, 1997.

[4] W. Jester and Y. Kallinderis, "Numerical study of incompressible flow about fixed cylinder pairs," *Journal of Fluids and Structures*, vol. 17, no. 4, pp. 561–577, 2003.

[5] Y. Bao, D. Zhou, and C. Huang, "Numerical simulation of flow over three circular cylinders in equilateral arrangements at low

Reynolds number by a second-order characteristic-based split finite element method," *Computers and Fluids*, vol. 39, no. 5, pp. 882–899, 2010.

[6] F. Xu and J.-P. Ou, "Numerical simulation of vortex-induced vibration of three cylinders subjected to a cross flow in equilateral arrangement," *Acta Aerodynamica Sinica*, vol. 28, no. 5, pp. 582–590, 2010.

[7] A. Zhang and L. Zhang, "Numerical simulation of three equispaced circular cylinders," *Chinese Journal of Applied Mechanics*, vol. 20, no. 1, pp. 31–36, 2003.

[8] K. K. Chen, J. H. Tu, and C. W. Rowley, "Variants of dynamic mode decomposition: boundary condition, Koopman, and Fourier analyses," *Journal of Nonlinear Science*, vol. 22, no. 6, pp. 887–915, 2012.

[9] H. Ding, C. Shu, K. S. Yeo, and D. Xu, "Numerical simulation of flows around two circular cylinders by mesh-free least-square-based finite difference methods," *International Journal for Numerical Methods in Fluids*, vol. 53, no. 2, pp. 305–332, 2007.

[10] Z. Han, D. Zhou, Y. Chen, X. Gui, and J. Li, "Numerical simulation of cross-flow around multiple circular cylinders by spectral element method," *Acta Aerodynamica Sinica*, vol. 32, no. 1, pp. 21–37, 2014.

[11] P. Hao, G. Li, L. Yang, and G. Chen, "Large eddy simulation of the circular cylinder flow in different regimes," *Chinese Journal of Applied Mechanics*, vol. 29, no. 4, pp. 437–443, 2012.

[12] A. G. Kravchenko and P. Moin, "Numerical studies of flow over a circular cylinder at ReD=3900," *Physics of Fluids*, vol. 12, no. 2, pp. 403–417, 2000.

[13] G. V. Papaioannou, D. K. P. Yue, M. S. Triantafyllou, and G. E. Karniadakis, "Three-dimensionality effects in flow around two tandem cylinders," *Journal of Fluid Mechanics*, vol. 558, pp. 387–413, 2006.

[14] A. Kumar and R. K. Ray, "Numerical study of shear flow past a square cylinder at reynolds numbers 100, 200," *Procedia Engineering*, vol. 127, pp. 102–109, 2015.

[15] J. O. Pralits, F. Giannetti, and L. Brandt, "Three-dimensional instability of the flow around a rotating circular cylinder," *Journal of Fluid Mechanics*, vol. 730, pp. 5–18, 2013.

[16] G. Schewe, "Reynolds-number-effects in flow around a rectangular cylinder with aspect ratio 1:5," *Journal of Fluids and Structures*, vol. 39, pp. 15–26, 2013.

[17] F. X. Trias, A. Gorobets, and A. Oliva, "Turbulent flow around a square cylinder at Reynolds number 22,000: a DNS study," *Computers and Fluids*, vol. 123, pp. 87–98, 2015.

[18] M. Zhao, L. Cheng, B. Teng, and G. Dong, "Hydrodynamic forces on dual cylinders of different diameters in steady currents," *Journal of Fluids and Structures*, vol. 23, no. 1, pp. 59–83, 2007.

[19] M. Mahbub Alam and Y. Zhou, "Strouhal numbers, forces and flow structures around two tandem cylinders of different diameters," *Journal of Fluids and Structures*, vol. 24, no. 4, pp. 505–526, 2008.

[20] Y. Ding-Yong, L. Hong-Chao, and W. Chang-Hai, "Numerical simulation of viscous flow past two tandem circular cylinders of different diameters," *Periodical of Ocean University of China*, no. 2, pp. 160–165, 2012.

[21] P. Zhang, Z.-D. Su, L.-D. Guang, and Y.-L. Li, "Effect of the diameter change on the flow past three cylinders," *Journal of Hydrodynamics*, vol. 27, no. 5, pp. 554–560, 2012.

[22] Q. Yao and H. Kanayama, "A coupling analysis of thermal convection problems based on a characteristic curve method," *Theoretical and Applied Mechanics Japan*, vol. 59, pp. 257–264, 2011.

[23] Q. Yao and Q. Zhu, "A pressure-stabilized Lagrange-Galerkin method in a parallel domain decomposition system," *Abstract and Applied Analysis*, vol. 2013, Article ID 161873, 13 pages, 2013.

[24] Q. Yao and H. Zhu, "Numerical simulation of hydrogen dispersion behaviour in a partially open space by a stabilized balancing domain decomposition method," *Computers and Mathematics with Applications*, vol. 69, no. 10, pp. 1068–1079, 2015.

[25] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer, New York, NY, USA, 1991.

[26] Q.-H. Yao and Q.-Y. Zhu, "Investigation of the contamination control in a cleaning room with a moving AGV by 3D large-scale simulation," *Journal of Applied Mathematics*, vol. 2013, Article ID 570237, 10 pages, 2013.

[27] F. Brezzi and J. Douglas, "Stabilized mixed methods for the Stokes problem," *Numerische Mathematik*, vol. 53, no. 1-2, pp. 225–235, 1988.

[28] J. Mandel, "Balancing domain decomposition," *Communications in Numerical Methods in Engineering*, vol. 9, no. 3, pp. 233–241, 1993.

[29] A. Toselli and O. B. Widlund, *Domain Decomposition Methods—Algorithms and Theory*, Springer, Berlin, Germany, 2005.

[30] M. M. Zdravkovich, "REVIEW—review of flow interference between two circular cylinders in various arrangements," *Journal of Fluids Engineering*, vol. 99, no. 4, pp. 618–633, 1977.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 1673959, 8 pages
http://dx.doi.org/10.1155/2016/1673959

Research Article

An MDADT-Based Approach for L_2-Gain Analysis of Discrete-Time Switched Delay Systems

Honglei Xu,[1] Xiang Xie,[1] and Lilian Shi[2]

[1] *School of Energy and Power Engineering, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China*
[2] *School of Engineering, Shaoxing University, Shaoxing, Zhejiang 312000, China*

Correspondence should be addressed to Xiang Xie; xiang_xie@outlook.com and Lilian Shi; sllian@sina.com

Received 8 October 2015; Accepted 14 February 2016

Academic Editor: Herve G. E. Kadji

We study the L_2-gain analysis problem for a class of discrete-time switched systems with time-varying delays. A mode-dependent average dwell time (MDADT) approach is applied to analyze the L_2-gain performance for these discrete-time switched delay systems. Combining a multiple Lyapunov functional method with the MDADT approach, sufficient conditions expressed in form of a set of feasible linear matrix inequalities (LMIs) are established to guarantee the L_2-gain performance. Finally, a numerical example will be provided to demonstrate the validity and usefulness of the obtained results.

1. Introduction

Switched systems consist of a finite number of subsystems and a logical law which orchestrates the switching behaviors between these subsystems. These dynamical systems can mathematically model many practical engineering applications with switching characteristics in a variety of disciplines; see, for example, [1–7].

A constrained switching signal can be regarded as a powerful tool to stabilize and control these switched systems [8–10]. Among them, the average dwell time (ADT) switching is the most common and typical one. It guarantees that the number of types of switching in a finite interval be bounded and the average time between any two types of consecutive switching not be less than a positive constant [11, 12]. In recent years, it has been recognised that ADT is flexible and efficient for dynamics analysis of many switched systems [8, 13–16]. However, the ADT switching's property that the average time interval between any two types of consecutive switching should be greater than a positive number τ_a makes the dwell time independent of the system modes. Hence whether the dwelling at some classes of subsystems will deteriorate the disturbance attenuation cannot be predicted.

As shown in [17], the minimum of admissible ADT is computed by two mode-independent parameters: the increase coefficient of the Lyapunov-like function and the decay rate of the Lyapunov function, which will cause certain conservativeness. To solve the problem, more recently, a new mode-dependent ADT concept has been introduced in [18]. Two mode-independent parameters can be set in a mode-dependent manner, which will reduce the conservativeness.

Even though stability analysis for the switched systems with MDADT has been investigated extensively (see, e.g., [17, 18]), how to solve the L_2-gain problem of the switched systems with MDADT is interesting and worthwhile to study. This has motivated our study in this paper.

The rest of the paper is as follows. In Section 2, we introduce the class of discrete-time switched system, some necessary definitions, and lemmas. In Section 3, sufficient conditions for ensuring L_2-gain for the discrete-time switched delay system are constructed. In Section 4, a numerical example is presented to illustrate the obtained results. Conclusion remarks are given in Section 5.

2. Preliminaries and Problem Statement

Consider a discrete-time switched system with a time-varying delay:

$$L_i: \begin{cases} x(t+1) = A_i x(t) + B_i x(t-d(t)) + C_i w(t), \\ x_{t_0}(l) = x(t_0 + l) = \phi(l), \\ z(t) = D_i x(t) + E_i w(t), \end{cases} \quad l = -d_M, -d_M+1, -d_M+2, \ldots, 0, \tag{1}$$

where $x(t) \in R^n$ is the system state, $z(t) \in R^m$ is the controlled output, $\phi(l)$ is a vector-valued initial function, t_0 is the initial time, and $w(t)$ is the disturbance input which belongs to $L_2[0, +\infty)$. $d(t)$ is the time-varying delay and satisfies $0 < d_m < d(t) \le d_M$, where d_m and d_M denote the upper and the lower bounds of the delays. i is the switching signal, which takes its values in the finite set $S = \{1, \ldots, M\}$, where M is the number of subsystems. When $t \in [t_i, t_{i+1})$, $i \in \mathbb{N}$, we call the ith subsystem active. A_p, B_p, C_p, D_p, and E_p are constant matrices with appropriate dimension. When $i = p = 1, \ldots, m$, it represents the pth subsystem or pth mode of (1).

To proceed, we need the following definitions and lemmas.

Definition 1 (see [11]). For any $T_2 > T_1 \ge 0$ and any switching signal i, $T_1 \le t < T_2$, let $N_i(T_1, T_2)$ denote the number of types of switching of i over (T_1, T_2). If $N_i(T_1, T_2) \le N_0 + T_2 - T_1/T_a$ holds for $N_0 \ge 0$ and $T_a > 0$, then T_a is the average dwell time and N_0 is the chatter bound. Without loss of generality, we choose $N_0 = 0$.

Definition 2 (see [18]). For a switching signal i and any $T \ge t \ge 0$, let $N_{ip}(T, t)$ be the switching numbers in which the pth subsystem is activated over the interval $[t, T]$ and let $T_p(T, t)$ denote the total running time of the pth subsystem over the interval $[t, T]$, $p \in S$. We say that i has a mode-dependent average dwell time (MDADT) τ_{ap} if there exist positive numbers N_{op} and τ_{ap} such that

$$N_{ip}(T, t) \le N_{op} + \frac{T_p(T, t)}{\tau_{ap}}, \quad \forall T \ge t \ge 0 \tag{2}$$

and we call N_{op} the mode-dependent chatter bounds. Here, we choose $N_{op} = 0$ as well.

Definition 3. For $\gamma > 0$, the switched delay system (1) is said to have L_2-gain property, if, under zero initial condition $\phi(l) = 0$, $l \in [t_0 - d_M, t_0]$, it holds that

$$\int_0^\infty z^T(s) z(s) ds \le \gamma^2 \int_0^\infty w^T(s) w(s) ds. \tag{3}$$

Lemma 4. *For any given matrices $X, Y \in R^{n\times n}$, it holds that*

$$X^T Y + Y^T X \le \delta X^T X + \delta^{-1} Y^T Y, \tag{4}$$

where δ is any given positive constant.

Lemma 5 (see [6]). *Let A, D, E, F, and P be real matrices of appropriate dimensions with $P > 0$ and F satisfying $F^T F \le I$. Then for any scalar $\varepsilon > 0$ satisfying $P^{-1} - \varepsilon^{-1} DD^T > 0$, one has*

$$\begin{aligned} &(A + DFE)^T P (A + DFE) \\ &\quad \le A^T \left(P^{-1} - \varepsilon^{-1} DD^T\right)^{-1} A + \varepsilon E^T E. \end{aligned} \tag{5}$$

Lemma 6 (Schur complement). *Let M, P, and Q be given matrices such that $Q > 0$. Then*

$$\begin{aligned} &\begin{bmatrix} P & M \\ * & -Q \end{bmatrix} < 0 \Longleftrightarrow \\ &P + MQ^{-1}M^T < 0. \end{aligned} \tag{6}$$

Lemma 7 (see [13]). *Let $\phi(k) \in R^n$ be a vector-valued function. If there exist any matrices $R > 0$, G_1, G_2, and a scalar $d \ge 0$, then the following inequality*

$$\begin{aligned} &-\sum_{s=k-d}^{k-1} N^T(s) R N(s) \\ &\quad \le \eta^T(k) \begin{bmatrix} G_1 + G_1^T & -G_1^T + G_2 \\ * & -G_2 - G_2^T \end{bmatrix} \eta(k) \\ &\quad + \eta^T(k) \begin{bmatrix} G_1^T \\ G_2^T \end{bmatrix} dR^{-1} \begin{bmatrix} G_1 & G_2 \end{bmatrix} \eta(k) \end{aligned} \tag{7}$$

holds, where $N(s) = \phi(s+1) - \phi(s)$ and $\eta(t) = \left[\begin{smallmatrix} \phi(t) \\ \phi(t-d) \end{smallmatrix}\right]$.

3. L_2-Gain Analysis

Firstly, we will introduce two important lemmas for the L_2-gain analysis of the switched delay system (1). The first lemma will provide the decay estimation of the Lyapunov functional $V_i(t)$ along the trajectory of the switched delay system without disturbances.

Lemma 8. *Consider the switched delay system (1) with $w(t) = 0$. For given positive integers d_M, d_m, and λ_i, suppose that there exist matrices G_1, G_2, Ω_1, Ω_2, and Ω_3 such that*

(i)

$$\Omega_3 \le 0. \tag{8}$$

(ii)

$$\Omega_1 - \Omega_2 \Omega_3^{-1} \Omega_2^T \leq 0, \tag{9}$$

where

$$\begin{aligned}
\Omega_1 &= A_i^T P_i A_i - P_i + \lambda_i^{-2} Q_i + (d_M - d_m) \lambda_i^{-2} Q_i \\
&\quad + \lambda_i^{-2} d_M \left[\lambda_i^2 A_i^T R_i A_i - \lambda_i R_i A_i - \lambda_i A_i^T R_i + R_i\right] \\
&\quad + \lambda_i^{-2} \left(G_1 + G_1^T + d_M G_1^T R_i^{-1} G_1\right), \\
\Omega_2 &= A_i^T P_i B_i + d_M \left(A_i^T R_i B_i - \lambda_i^{-1} R_i B_i\right) \\
&\quad + \lambda_i^{-2} \left(-G_1^T + G_2 + d_M G_1^T R_i^{-1} G_2\right), \\
\Omega_3 &= B_i^T P_i B_i - \lambda_i^{-2(1+d_M)} Q_i + d_M B_i^T R_i B_i \\
&\quad + \lambda_i^{-2} \left(-G_2 - G_2^T + d_M G_2^T R_i^{-1} G_2\right)
\end{aligned} \tag{10}$$

with P_i, Q_i, and R_i being symmetric positive definite matrices; then the Lyapunov functional $V_i(t)$ along the trajectory of the switched delay system (1) will satisfy

$$V_i(t) \leq \lambda_i^{-2(t-t_0)} V_i(t_0). \tag{11}$$

Proof. Choose the following Lyapunov functional candidate:

$$V_i(t) = V_{i_1}(t) + V_{i_2}(t) + V_{i_3}(t) + V_{i_4}(t). \tag{12}$$

Here,

$$\begin{aligned}
V_{i_1}(t) &= x^T(t) P_i x(t), \\
V_{i_2}(t) &= \sum_{s=t-d(t)}^{t-1} \lambda_i^{2(s-t)} x^T(s) Q_i x(s), \\
V_{i_3}(t) &= \sum_{\theta=-d_M+2}^{-d_m+1} \sum_{s=t-1+\theta}^{t-1} \lambda_i^{2(s-t)} x^T(s) Q_i x(s), \\
V_{i_4}(t) &= \sum_{\theta=-d_M+1}^{0} \sum_{s=t-1+\theta}^{t-1} \lambda_i^{2(s-t)} y^T(s) R_i y(s),
\end{aligned} \tag{13}$$

where P_i, Q_i, and R_i are symmetric positive definite matrices, $\lambda_i > 1$ is a given constant, and $y(s) = \lambda_i x(s+1) - x(s)$. Next, we will estimate the difference of $V_i(t)$ along the trajectory of the switched delay system (1):

$$\begin{aligned}
\Delta V_{i_1}(t) &= V_{i_1}(t+1) - V_{i_1}(t) \\
&= x^T(t+1) P_i x(t+1) - x^T(t) P_i x(t) \\
&= x^T(t) A_i^T P_i A_i x(t) \\
&\quad + x^T(t-d(t)) B_i^T P_i A_i x(t) \\
&\quad + x^T(t) A_i^T P_i B_i x(t-d(t)) \\
&\quad + x^T(t-d(t)) B_i^T P_i B_i x(t-d(t)) \\
&\quad - x^T(t) P_i x(t).
\end{aligned} \tag{14}$$

Then, we have

$$\begin{aligned}
\Delta V_{i_1}(t) &= \begin{bmatrix} x(t) \\ x(t-d(t)) \end{bmatrix}^T \\
&\quad \cdot \begin{bmatrix} A_i^T P_i A_i - P_i & A_i^T P_i B_i \\ B_i^T P_i A_i & B_i^T P_i B_i \end{bmatrix} \begin{bmatrix} x(t) \\ x(t-d(t)) \end{bmatrix},
\end{aligned} \tag{15}$$

$$\begin{aligned}
\Delta V_{i_2}(t) &= V_{i_2}(t+1) - V_{i_2}(t) \leq V_{i_2}(t+1) - \lambda_i^{-2} V_{i_2}(t) \\
&= \sum_{s=t+1-d(t+1)}^{t} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s) \\
&\quad - \sum_{s=t-d(t)}^{t-1} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s) = \lambda_i^{-2} x^T(t) Q_i x(t) \\
&\quad + \sum_{s=t+1-d(t+1)}^{t-1} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s) \\
&\quad - \sum_{s=t-d(t)}^{t-1} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s).
\end{aligned} \tag{16}$$

Since the delay $d(t)$ satisfies $0 < d_m < d(t) \leq d_M$, we can consider the following two cases.

When $d_m > 1$, it holds that

$$\begin{aligned}
&\sum_{s=t+1-d(t+1)}^{t-1} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s) \\
&\quad \leq \sum_{s=t+1-d_m}^{t-1} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s) \\
&\qquad + \sum_{s=t+1-d(t+1)}^{t-d_m} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s) \\
&\quad \leq \sum_{s=t+1-d(t)}^{t-1} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s) \\
&\qquad + \sum_{s=t+1-d_M}^{t-d_m} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s).
\end{aligned} \tag{17}$$

When $d_m = 1$,

$$\begin{aligned}&\sum_{s=t+1-d(t+1)}^{t-1} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s)\\ &\quad\le \sum_{s=t+1-d(t)}^{t-1} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s)\\ &\qquad+ \sum_{s=t+1-d_M}^{t-d_m} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s)\end{aligned} \tag{18}$$

is satisfied as well.

So from (16) and (17) we can obtain

$$\begin{aligned}\Delta V_{i_2}(t) &\le \lambda_i^{-2} x^T(t) Q_i x(t)\\ &\quad+ \sum_{s=t+1-d(t)}^{t-1} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s)\\ &\quad+ \sum_{s=t+1-d_M}^{t-d_m} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s)\\ &\quad- \sum_{s=t-d(t)}^{t-1} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s)\\ &= \lambda_i^{-2} x^T(t) Q_i x(t)\\ &\quad- \lambda_i^{2(-1-d(t))} x^T(t-d(t)) Q_i x(t-d(t))\\ &\quad+ \sum_{s=t+1-d_M}^{t-d_m} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s).\end{aligned} \tag{19}$$

Since $-\lambda_i^{2(-1-d(t))} \le -\lambda_i^{2(-1-d_M)}$, we get

$$\begin{aligned}\Delta V_{i_2}(t) &\le \begin{bmatrix} x(t) \\ x(t-d(t)) \end{bmatrix}^T\\ &\quad\cdot \begin{bmatrix} \lambda_i^{-2} Q_i & 0 \\ 0 & -\lambda_i^{2(-1-d_M)} Q_i \end{bmatrix} \begin{bmatrix} x(t) \\ x(t-d(t)) \end{bmatrix}\\ &\quad+ \sum_{s=t+1-d_M}^{t-d_m} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s).\end{aligned} \tag{20}$$

The derivation process of $\Delta V_{i_3}(t)$ is similar to $\Delta V_{i_2}(t)$, and then we have

$$\begin{aligned}\Delta V_{i_3}(t) &= V_{i_3}(t+1) - V_{i_3}(t) \le V_{i_3}(t+1) - \lambda_i^{-2} V_{i_3}(t)\\ &= \sum_{\theta=-d_M+2}^{-d_m+1} \sum_{s=t+\theta}^{t-1} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s)\\ &\quad+ \sum_{\theta=-d_M+2}^{-d_m+1} \lambda_i^{-2} x^T(t) Q_i x(t)\\ &\quad- \sum_{\theta=-d_M+2}^{-d_m+1} \sum_{s=t-1+\theta}^{t-1} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s)\\ &\le -\sum_{s=t+1-d_M}^{t-d_m} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s)\\ &\quad+ (d_M - d_m) \lambda_i^{-2} x^T(t) Q_i x(t)\end{aligned} \tag{21}$$

which is equal to

$$\begin{aligned}\Delta V_{i_3} &\le \begin{bmatrix} x(t) \\ x(t-d(t)) \end{bmatrix}^T\\ &\quad\cdot \begin{bmatrix} (d_M - d_m)\lambda_i^{-2} Q_i & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ x(t-d(t)) \end{bmatrix}\\ &\quad- \sum_{s=t+1-d_M}^{t-d_m} \lambda_i^{2(s-t-1)} x^T(s) Q_i x(s),\end{aligned} \tag{22}$$

$$\begin{aligned}\Delta V_{i_4}(t) &= V_{i_4}(t+1) - V_{i_4}(t) \le V_{i_4}(t+1) - \lambda_i^{-2} V_{i_4}(t)\\ &\le \sum_{\theta=-d_M+1}^{0} \sum_{s=t+\theta}^{t-1} \lambda_i^{2(s-t-1)} y^T(s) R_i y(s) + d_M\\ &\quad\cdot \lambda_i^{-2} y^T(t) R_i y(t)\\ &\quad- \sum_{\theta=-d_M+1}^{0} \sum_{s=t-1+\theta}^{t-1} \lambda_i^{2(s-t-1)} y^T(s) R_i y(s) \le d_M\\ &\quad\cdot \lambda_i^{-2} y^T(t) R_i y(t) - \sum_{s=t-d_M}^{t-1} \lambda_i^{2(s-t-1)} y^T(s) R_i y(s).\end{aligned} \tag{23}$$

Since $y(s) = \lambda_i x(s+1) - x(s)$, we substitute it into (23) and obtain

$$\begin{aligned}\Delta V_{i_4}(t) &\le \lambda_i^{-2} d_M \begin{bmatrix} x(t) \\ x(t-d(t)) \end{bmatrix}^T \begin{bmatrix} \left(\lambda_i A_i^T - I\right) R_i \left(\lambda_i A_i - I\right) & \lambda_i \left(\lambda_i A_i^T - I\right) R_i B_i \\ B_i^T R_i \left(\lambda_i A_i - I\right) \lambda_i & \lambda_i^2 B_i^T R_i B_i \end{bmatrix} \begin{bmatrix} x(t) \\ x(t-d(t)) \end{bmatrix}\\ &\quad- \sum_{s=t-d_M}^{t-1} \lambda_i^{2(s-t-1)} y^T(s) R_i y(s),\end{aligned} \tag{24}$$

where we apply the transformation $\phi(s) = \lambda_i^{(s-t-1)} x(s)$. Then we have $\lambda_i^{s-t-1} y(s) = \phi(s+1) - \phi(s)$; by Lemma 7 we continue to have

$$-\sum_{s=t-d_M}^{t-1} \lambda_i^{2(s-t-1)} y^T(s) R_i y(s) \le \begin{bmatrix} \phi(t) \\ \phi(t-d(t)) \end{bmatrix}^T \cdot \begin{bmatrix} G_1 + G_1^T & -G_1^T + G_2 \\ * & -G_2 - G_2^T \end{bmatrix} \begin{bmatrix} \phi(t) \\ \phi(t-d(t)) \end{bmatrix} + \begin{bmatrix} \phi(t) \\ \phi(t-d(t)) \end{bmatrix}^T \cdot \begin{bmatrix} G_1^T \\ G_2^T \end{bmatrix} d_M R_i^{-1} \begin{bmatrix} G_1 & G_2 \end{bmatrix} \begin{bmatrix} \phi(t) \\ \phi(t-d(t)) \end{bmatrix}. \tag{25}$$

Due to the fact that $\phi(t) = \lambda_i^{-1} x(t)$, $\phi(t-d(t)) = \lambda_i^{-(d(t)+1)} x(t-d(t)) \le \lambda_i^{-1} x(t-d(t))$, it holds that

$$\begin{aligned} \Delta V_{i_4}(t) \le{}& \lambda_i^{-2} d_M \begin{bmatrix} x(t) \\ x(t-d(t)) \end{bmatrix}^T \begin{bmatrix} \left(\lambda_i A_i^T - I\right) R_i \left(\lambda_i A_i - I\right) & \lambda_i \left(\lambda_i A_i^T - I\right) R_i B_i \\ B_i^T R_i \left(\lambda_i A_i - I\right) \lambda_i & \lambda_i^2 B_i^T R_i B_i \end{bmatrix} \begin{bmatrix} x(t) \\ x(t-d(t)) \end{bmatrix} \\ &+ \lambda_i^{-2} \begin{bmatrix} x(t) \\ x(t-d(t)) \end{bmatrix}^T \begin{bmatrix} G_1 + G_1^T & -G_1^T + G_2 \\ * & -G_2 - G_2^T \end{bmatrix} \begin{bmatrix} x(t) \\ x(t-d(t)) \end{bmatrix} \\ &+ \lambda_i^{-2} \begin{bmatrix} x(t) \\ x(t-d(t)) \end{bmatrix}^T \begin{bmatrix} G_1^T \\ G_2^T \end{bmatrix} d_M R_i^{-1} \begin{bmatrix} G_1 & G_2 \end{bmatrix} \begin{bmatrix} x(t) \\ x(t-d(t)) \end{bmatrix}. \end{aligned} \tag{26}$$

Let $\xi(t) = \left[\begin{smallmatrix} x(t) \\ x(t-d(t)) \end{smallmatrix}\right]$; then we add (15), (20), (22), and (26) together to yield

$$\Delta V_i \le \xi^T(t) \Omega \xi(t), \tag{27}$$

where $\Omega = \left[\begin{smallmatrix} \Omega_1 & \Omega_2 \\ * & \Omega_3 \end{smallmatrix}\right]$,

$$\begin{aligned} \Omega_1 ={}& A_i^T P_i A_i - P_i + \lambda_i^{-2} Q_i + \left(d_M - d_m\right) \lambda_i^{-2} Q_i \\ &+ \lambda_i^{-2} d_M \left[\lambda_i^2 A_i^T R_i A_i - \lambda_i R_i A_i - \lambda_i A_i^T R_i + R_i\right] \\ &+ \lambda_i^{-2} \left(G_1 + G_1^T + d_M G_1^T R_i^{-1} G_1\right), \\ \Omega_2 ={}& A_i^T P_i B_i + d_M \left(A_i^T R_i B_i - \lambda_i^{-1} R_i B_i\right) \\ &+ \lambda_i^{-2} \left(-G_1^T + G_2 + d_M G_1^T R_i^{-1} G_2\right), \\ \Omega_3 ={}& B_i^T P_i B_i - \lambda_i^{-2(1+d_M)} Q_i + d_M B_i^T R_i B_i \\ &+ \lambda_i^{-2} \left(-G_2 - G_2^T + d_M G_2^T R_i^{-1} G_2\right). \end{aligned} \tag{28}$$

By (8) and (9) and Lemma 6, we can obtain

$$\Omega = \begin{bmatrix} \Omega_1 & \Omega_2 \\ * & \Omega_3 \end{bmatrix} \le 0. \tag{29}$$

It follows from (27) and (29) that

$$V_i(t+1) \le \lambda_i^{-2} V_i(t). \tag{30}$$

Therefore,

$$V_i(t) \le \lambda_i^{-2} V_i(t-1) \le \cdots \le \lambda_i^{-2(t-t_0)} V_i(t_0). \tag{31}$$

This completes the proof. □

Lemma 9. *For given constants λ_i and γ_0, suppose that there exist matrices Ξ_1, Ξ_2, and Ξ_3 such that*

(i)

$$\Xi_3 \le 0 \tag{32}$$

(ii)

$$\Xi_1 - \Xi_2 \Xi_3^{-1} \Xi_2^T \le 0 \tag{33}$$

and $\gamma_0 > 0$, $\varepsilon_1 > 0$, and $\varepsilon_2 > 0$ satisfying

$$\gamma_0^2 I \ge \varepsilon_1^{-1} I + \varepsilon_2^{-1} I + C_i^T P_i C_i + d_M C_i^T R_i C_i + E_i^T E_i; \tag{34}$$

then along the trajectory of system (1), one has

$$V_i(t+1) \le \lambda_i^{-2} V_i(t) + \gamma_0^2 w^T(t) w(t) - Z^T(t) Z(t), \tag{35}$$

where

$$\begin{aligned} \Xi_1 &= \Omega_1 + \varepsilon_1 \varphi_{1_i}^T \varphi_{1_i} + D_i^T D_i, \\ \Xi_2 &= \Omega_2, \\ \Xi_3 &= \Omega_3 + \varepsilon_2 \varphi_{2_i}^T \varphi_{2_i}, \\ \varphi_{1_i} &= C_i^T P_i A_i + d_M \left(C_i^T R_i A_i - \lambda_i^{-1} C_i^T R_i\right) + E_i^T D_i \\ \varphi_{2_i} &= C_i^T P_i B_i + d_M C_i^T R_i B_i. \end{aligned} \tag{36}$$

Proof. Using Lemma 8 and (1), we have

$$\begin{aligned}
&V_i(t+1) - \lambda_i^{-2} V_i(t) + Z^T(t) Z(t) - \gamma_0^2 w^T(t) w(t) \\
&\quad \le \xi^T(t) \Omega \xi(t) + x^T(t) \\
&\quad \cdot \left[A_i^T P_i C_i + d_M \left(A_i^T R_i C_i - \lambda_i^{-1} R_i C_i\right) + D_i^T E_i\right] \\
&\quad \cdot w(t) + w^T(t) \\
&\quad \cdot \left[C_i^T P_i A_i + d_M \left(C_i^T R_i A_i - \lambda_i^{-1} C_i^T R_i\right) + E_i^T D_i\right] \\
&\quad \cdot x(t) + x^T(t-d(t)) \left[B_i^T P_i C_i + d_M B_i^T R_i C_i\right] w(t) \\
&\quad + w^T(t) \left[C_i^T P_i B_i + d_M C_i^T R_i B_i\right] x(t-d(t)) \\
&\quad + x^T(t) D_i^T D_i x(t) + w^T(t) \\
&\quad \cdot \left(C_i^T P_i C_i + d_M C_i^T R_i C_i + E_i^T E_i - \gamma_0^2 I\right) w(t).
\end{aligned} \tag{37}$$

Based on Lemmas 4 and 5, it holds that

$$\begin{aligned}
&x^T(t) \left[A_i^T P_i C_i + d_M \left(A_i^T R_i C_i - \lambda_i^{-1} R_i C_i\right) + D_i^T E_i\right] \\
&\quad \cdot w(t) + w^T(t) \\
&\quad \cdot \left[C_i^T P_i A_i + d_M \left(C_i^T R_i A_i - \lambda_i^{-1} C_i^T R_i\right) + E_i^T D_i\right] \\
&\quad \cdot x(t) \le \varepsilon_1 x^T(t) \varphi_{1_i}^T \varphi_{1_i} x(t) + \varepsilon_1^{-1} w^T(t) w(t). \\
&x^T(t-d(t)) \left[B_i^T P_i C_i + d_M B_i^T R_i C_i\right] w(t) + w^T(t) \\
&\quad \cdot \left[C_i^T P_i B_i + d_M C_i^T R_i B_i\right] x(t-d(t)) \\
&\quad \le \varepsilon_2 x^T(t-d(t)) \varphi_{2_i}^T \varphi_{2_i} x(t-d(t)) + \varepsilon_2^{-1} w^T(t) \\
&\quad \cdot w(t).
\end{aligned} \tag{38}$$

Then, it follows from (35) and (38) that

$$\begin{aligned}
&V_i(t+1) - \lambda_i^{-2} V_i(t) + Z^T(t) Z(t) - \gamma_0^2 w^T(t) w(t) \\
&\quad \le \xi^T(t) \\
&\quad \cdot \begin{bmatrix} \Omega_1 + \varepsilon_1 \varphi_{1_i}^T \varphi_{1_i} + D_i^T D_i & \Omega_2 \\ * & \Omega_3 + \varepsilon_2 \varphi_{2_i}^T \varphi_{2_i} \end{bmatrix} \xi(t) \\
&\quad + w^T(t) \left[\varepsilon_1^{-1} I + \varepsilon_2^{-1} I + C_i^T P_i C_i + d_M C_i^T R_i C_i \right. \\
&\quad \left. + E_i^T E_i - \gamma_0^2 I\right] w(t).
\end{aligned} \tag{39}$$

Combining (32), (33) with (34) will lead to

$$V_i(t+1) \le \lambda_i^{-2} V_i(t) + \gamma_0^2 w^T(t) w(t) - Z^T(t) Z(t). \tag{40}$$

This completes the proof. □

Now, our L_2-gain analysis results can be presented as follows.

Theorem 10. *For given constants λ_i and γ_0, suppose that there exist matrices Ξ_1, Ξ_2, and Ξ_3 such that*

(i)

$$\Xi_3 \le 0 \tag{41}$$

(ii)

$$\Xi_1 - \Xi_2 \Xi_3^{-1} \Xi_2^T \le 0 \tag{42}$$

and $\gamma_0 > 0$, $\varepsilon_1 > 0$, and $\varepsilon_2 > 0$ satisfying

$$\gamma_0^2 I \ge \varepsilon_1^{-1} I + \varepsilon_2^{-1} I + C_i^T P_i C_i + d_M C_i^T R_i C_i + E_i^T E_i. \tag{43}$$

Then the switched delay system (1) has a L_2-gain with MDADT $\tau_{ap} > \tau_{ap}^* = \ln \mu_p / 2 \ln \lambda_p$, where $\mu_p \ge 1$ satisfying (35) and φ_{1_i}, φ_{2_i}, Ξ_1, Ξ_2, and Ξ_3 are defined in Lemma 9.

Proof. Choose the Lyapunov functional candidate (12). From (41) and (42) and Lemma 9, we have

$$V_i(t+1) \le \lambda_i^{-2} V_i(t) + \gamma_0^2 w^T(t) w(t) - Z^T(t) Z(t). \tag{44}$$

Let $\Gamma(t) = \gamma_0^2 w^T(t) w(t) - Z^T(t) Z(t)$. From (35), since $t_{i-1} = t_i - 1$, we have

$$\begin{aligned}
V_{\sigma(t)}(t) &\le \lambda_{\sigma(t_i)}^{-2(t-t_i)} V_{\sigma(t_i)}(t_i) + \sum_{j=t_i}^{t-1} \lambda_{\sigma(t_i)}^{-2(t-j-1)} \Gamma(j) \\
&\le \mu_{\sigma(t_i)} \lambda_{\sigma(t_i)}^{-2(t-t_i)} V_{\sigma(t_{i-1})}(t_i) + \sum_{j=t_i}^{t-1} \lambda_{\sigma(t_i)}^{-2(t-j-1)} \Gamma(j) \\
&\le \mu_{\sigma(t_i)} \lambda_{\sigma(t_i)}^{-2(t-t_i)} \left\{ \lambda_{\sigma(t_{i-1})}^{-2(t_i-t_{i-1})} V_{\sigma(t_{i-1})}(t_{i-1}) \right. \\
&\quad \left. + \sum_{j=t_{i-1}}^{t_i-1} \lambda_{\sigma(t_{i-1})}^{-2(t_i-j-1)} \Gamma(j) \right\} + \sum_{j=t_i}^{t-1} \lambda_{\sigma(t_i)}^{-2(t-j-1)} \Gamma(j) \le \cdots \\
&\le \left(\prod_{s=1}^{i} \mu_{\sigma(t_s)}\right) \cdot \exp\left(-2 \sum_{s=1}^{i} \ln \lambda_{\sigma(t_{s-1})} (t_s - t_{s-1})\right) \\
&\quad \cdot V_{\sigma(t_0)}(t_0) + \sum_{k=1}^{i} \left[\left(\prod_{s=k}^{i} \mu_{\sigma(t_s)}\right)\right. \\
&\quad \left. \cdot \exp\left(-2 \sum_{s=k}^{i} \ln \lambda_{\sigma(t_{s-1})} (t_s - t_{s-1})\right)\right] \Gamma(t_{k-1}) \\
&= \exp\left[\sum_{s=1}^{i} \left(\ln \mu_{\sigma(t_s)} - 2 \ln \lambda_{\sigma(t_{s-1})} (t_s - t_{s-1})\right)\right] \\
&\quad \cdot V_{\sigma(t_0)}(t_0) \\
&\quad + \sum_{k=1}^{i} \left\{\exp\left[\sum_{s=k}^{i} \left(\ln \mu_{\sigma(t_s)} - 2 \ln \lambda_{\sigma(t_{s-1})} (t_s - t_{s-1})\right)\right]\right\} \\
&\quad \cdot \Gamma(t_{k-1})
\end{aligned} \tag{45}$$

which combined with Definition 2 and the MDADT scheme $N_{\sigma p}(T,t) \le T_p(T,t)/\tau_{ap}$ yields

$$\begin{aligned} V_{\sigma(t)}(t) \le{} & \exp\left[\sum_{p=1}^{m}\left(\ln \mu_p^{N_{\sigma p}(t,t_0)} - 2\ln\lambda_p T_p(t,t_0)\right)\right] \\ & \cdot V_{\sigma(t_0)}(t_0) \\ & + \sum_{k=1}^{i}\left\{\exp\left[\sum_{p=1}^{m}\left(\ln \mu_p^{N_{\sigma p}(t,t_{k-1})} - 2\ln\lambda_p T_p(t,t_{k-1})\right)\right]\right\} \\ & \cdot \Gamma(t_{k-1}) \le e^{-\beta T_p(t,t_0)} V_{\sigma(t_0)}(t_0) + \sum_{k=1}^{i} e^{-\beta T_p(t,t_{k-1})}\Gamma(t_{k-1}), \end{aligned} \tag{46}$$

where $\beta = \sum_{p=1}^{m}(2\ln\lambda_p - \ln\mu_p/\tau_{\sigma p}) > 0$.

Under zero initial condition, from (46), one obtains

$$0 \le V_{\sigma(t)}(t) \le \sum_{k=1}^{i} e^{-\beta T_p(t,t_0)}\Gamma(t_{k-1}) \tag{47}$$

which implies that

$$\begin{aligned} & \sum_{k=1}^{i} e^{-\beta T_p(t,t_{k-1})} Z^T(t_{k-1}) Z(t_{k-1}) \\ & \quad \le \sum_{k=1}^{i} e^{-\beta T_p(t,t_{k-1})}\gamma_0^2 w^T(t_{k-1}) w(t_{k-1}). \end{aligned} \tag{48}$$

Then, we multiply both sides by $e^{-\beta T_p(t_{k-1},t_0)}$ to get

$$\begin{aligned} & \sum_{k=1}^{i} e^{-\beta T_p(t,t_0)} Z^T(t_{k-1}) Z(t_{k-1}) \\ & \quad \le \sum_{k=1}^{i} e^{-\beta T_p(t,t_0)}\gamma_0^2 w^T(t_{k-1}) w(t_{k-1}). \end{aligned} \tag{49}$$

Thus,

$$\sum_{k=0}^{i} Z^T(t_k) Z(t_k) \le \sum_{k=0}^{i}\gamma_0^2 w^T(t_k) w(t_k). \tag{50}$$

This completes the proof. □

4. A Numerical Example

Consider the switched delay system (1) with the following specifications:

$$\begin{aligned} A_1 &= \begin{bmatrix} -0.2 & 0.3 \\ 0.1 & -0.5 \end{bmatrix}, \\ B_1 &= \begin{bmatrix} 0.4 & 0 \\ 0.1 & -0.5 \end{bmatrix}, \\ A_2 &= \begin{bmatrix} -0.1 & 1 \\ 0 & -0.6 \end{bmatrix}, \\ B_2 &= \begin{bmatrix} -0.7 & 0.1 \\ 1 & 0.2 \end{bmatrix}, \\ C_1 &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \\ C_2 &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \\ D_1 &= [1, 1], \\ D_2 &= [0, 1], \\ E_1 &= E_2 = [0.2, 0.8], \end{aligned} \tag{51}$$

and $d(t) = \sin(t\pi/2) + 1$, so that $d_M = 2$, $d_m = 0$. The disturbance input is defined as

$$w(t) = \begin{cases} 1, & 0 < t \le 20, \\ 0, & t > 20. \end{cases} \tag{52}$$

Let $\mu_1 = \mu_2 = 12$; by the LMI Control Toolbox and Theorem 10, we obtain

$$\begin{aligned} P_1 &= \begin{bmatrix} 12.5739 & 5.0613 \\ 5.0613 & 4.8703 \end{bmatrix}, \\ Q_1 &= \begin{bmatrix} 2.6676 & 1.2445 \\ 1.2445 & 0.9452 \end{bmatrix}, \\ R_1 &= \begin{bmatrix} 4.4703 & -0.6996 \\ -0.6996 & 9.3862 \end{bmatrix}, \\ P_2 &= \begin{bmatrix} 12.3445 & 18.3565 \\ 18.3565 & 30.2222 \end{bmatrix}, \\ Q_2 &= \begin{bmatrix} 2.3941 & 3.7907 \\ 3.7907 & 6.1097 \end{bmatrix}, \\ R_2 &= \begin{bmatrix} 16.1813 & 18.8134 \\ 18.8134 & 22.4268 \end{bmatrix}, \end{aligned} \tag{53}$$

and $\lambda_1 = 27.2485$, $\lambda_2 = 38.7807$, where $\tau_{a1}^* = \ln\mu_1/2\ln\lambda_1 = 0.3759$ and $\tau_{a2}^* = \ln\mu_2/2\ln\lambda_2 = 0.3397$. Now, we choose

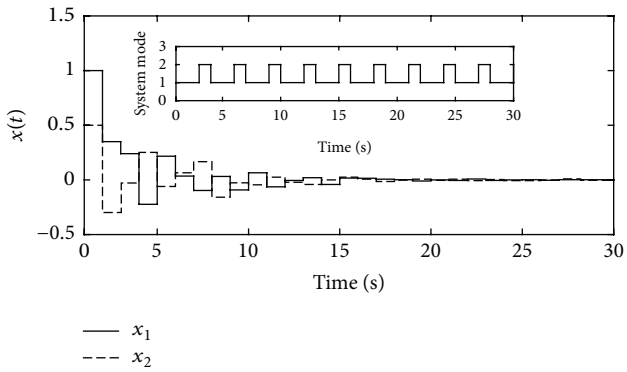


Figure 1: State trajectories of the switched delay system (1) under MDADT switching.

the switching periods $\tau_{a1} = 2$, $\tau_{a2} = 1$ and take the initial state condition $\psi(l) = [1; 2]$ for all $l = -2, -1, 0$. Then the numerical simulations can be shown in Figure 1.

It can be seen from Figure 1 that under the designed MDADT switching signals the switched delay system can achieve better dynamics performance and disturbance tolerance capability, which shows the potentiality of our results in practice.

5. Conclusions

In this paper, the problem of L_2-gain analysis for discrete-time switched systems with MDADT switching has been investigated. By combining with the multiple Lyapunov function method, sufficient conditions are established to ensure L_2-gain performance for discrete-time switched delay system, and the admissible MDADT switching signals are also designed accordingly. Finally, a numerical example is given to demonstrate the usefulness of the obtained results.

Competing Interests

The authors declare that they have no competing interests.

References

[1] S. Arik, "Stability analysis of delayed neural networks," *IEEE Transactions on Circuits and Systems. I. Fundamental Theory and Applications*, vol. 47, no. 7, pp. 1089–1092, 2000.

[2] M. S. Branicky, "Multiple Lyapunov functions and other analysis tools for switched and hybrid systems," *IEEE Transactions on Automatic Control*, vol. 43, no. 4, pp. 475–482, 1998.

[3] W.-H. Chen, Z.-H. Guan, and X. Lu, "Delay-dependent output feedback guaranteed cost control for uncertain time-delay systems," *Automatica*, vol. 40, no. 7, pp. 1263–1268, 2004.

[4] J. P. Hespanha and A. Stephen Morse, "Stability of switched systems with average Dwell-time," in *Proceedings of the 38th Conference on Decision & Control*, pp. 2655–2660, Phoenix, Ariz, USA, December 1999.

[5] D. Liberzon and A. S. Morse, "Basic problems in stability and design of switched systems," *IEEE Control Systems Magazine*, vol. 19, no. 5, pp. 59–70, 1999.

[6] X. Xie, H. Xu, and R. Zhang, "Exponential stabilization of impulsive switched systems with time delays using guaranteed cost control," *Abstract and Applied Analysis*, vol. 2014, Article ID 126836, 8 pages, 2014.

[7] S. Pettersson, "Synthesis of switched linear systems," in *Proceedings of the 42nd IEEE Conference on Decision and Control*, pp. 5283–5288, IEEE, December 2003.

[8] X.-L. Zhu, H. Yang, Y. Wang, and Y.-L. Wang, "New stability criterion for linear switched systems with time-varying delay," *International Journal of Robust and Nonlinear Control*, vol. 24, no. 2, pp. 214–227, 2014.

[9] D. Liberzon, J. P. Hespanha, and A. S. Morse, "Stability of switched systems: a Lie-algebraic condition," *Systems & Control Letters*, vol. 37, no. 3, pp. 117–122, 1999.

[10] H. Lin and P. J. Antsaklis, "Stability and stabilizability of switched linear systems: a survey of recent results," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 308–322, 2009.

[11] D. Liberzon, *Switching in Systems and Control*, Springer, 2003.

[12] X.-M. Sun, J. Zhao, and D. J. Hill, "Stability and L_2-gain analysis for switched delay systems: a delay-dependent method," *Automatica*, vol. 42, pp. 1769–1774, 2006.

[13] Q.-X. Chen, L. Yu, and W.-A. Zhang, "Delay-dependent output feedback guaranteed cost control for uncertain discrete-time systems with multiple time-varying delays," *IET Control Theory and Applications*, vol. 1, no. 1, pp. 97–103, 2007.

[14] J. Zhang, Z. Han, F. Zhu, and J. Huang, "Stability and stabilization of positive switched systems with mode-dependent average dwell time," *Nonlinear Analysis: Hybrid Systems*, vol. 9, pp. 42–55, 2013.

[15] Z. Li, H. Gao, and H. R. Karimi, "Stability analysis and H_∞ controller synthesis of discrete-time switched systems with time delay," *Systems & Control Letters*, vol. 66, no. 1, pp. 85–93, 2014.

[16] H. Yan and X. Zhang, "Robust exponential stability and L2-gain for switched discrete-time nonlinear cascade systems," in *Proceedings of the 33rd Chinese Control Conference (CCC '14)*, pp. 4198–4203, Nanjing, China, July 2014.

[17] X. Zhao, H. Liu, and Z. Wang, "Weighted H_∞ performance analysis of switched linear systems with mode-dependent average dwell time," *International Journal of Systems Science*, vol. 44, no. 11, pp. 2130–2139, 2013.

[18] X. Zhao, L. Zhang, P. Shi, and M. Liu, "Stability and stabilization of switched linear systems with mode-dependent average dwell time," *IEEE Transactions on Automatic Control*, vol. 57, no. 7, pp. 1809–1815, 2012.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 5641478, 11 pages
http://dx.doi.org/10.1155/2016/5641478

Research Article

Robust Control of Underactuated Systems: Higher Order Integral Sliding Mode Approach

Sami ud Din,[1,2] Qudrat Khan,[3,4] Fazal ur Rehman,[1] and Rini Akmeliawati[3]

[1] *Department of Electrical Engineering, Capital University of Science and Technology (CUST), Kahuta Road, Express Highway, Islamabad 44000, Pakistan*
[2] *Department of Electrical Engineering, The University of Lahore (UOL), Japan Road, Express Highway, Islamabad 44000, Pakistan*
[3] *Department of Mechatronics Engineering, International Islamic University, 50728 Kuala Lumpur, Malaysia*
[4] *Center for Advanced Studies in Telecommunications, COMSATS Institute of Information Technology, Islamabad 44000, Pakistan*

Correspondence should be addressed to Sami ud Din; engrsamiuddin@gmail.com

Received 25 September 2015; Revised 8 January 2016; Accepted 12 January 2016

Academic Editor: Wenguang Yu

This paper presents a robust control design for the class of underactuated uncertain nonlinear systems. Either the nonlinear model of the underactuated systems is transformed into an input output form and then an integral manifold is devised for the control design purpose or an integral manifold is defined directly for the concerned class. Having defined the integral manifolds discontinuous control laws are designed which are capable of maintaining sliding mode from the very beginning. The closed loop stability of these systems is presented in an impressive way. The effectiveness and demand of the designed control laws are verified via the simulation and experimental results of ball and beam system.

1. Introduction

The control design of underactuated systems was the main focus of the researchers in the current and last decade. These systems, by definition, contain less number of control inputs/actuators as compared to the degree of freedom [1]. This feature makes them quite different from the other nonlinear plants where the systems operate with the same number of inputs and outputs, the so-called fully actuated systems. The control design of these systems is quite demanding because of their vital theoretical and practical applications in the areas of aerospace systems, marine systems, humanoids, locomotive systems, manipulators of different kinds, and so forth [2]. This family also includes ball and beam system [3], TORA (translational oscillator with rotational actuator) [4], and inverted pendulum system [5]. These systems are used in order to have minimum weight, cost, and energy usage while still retaining the key features of the processes. In addition, another significant feature of underactuated systems is less damage in case of collision with other objects which in turn provides more safety to actuators [6]. Underactuation can be raised due to the hardware failure; this hardware solution to actuator failures can be achieved by equipping the vehicle with redundant actuators [2]. Note that, in case of fully actuated systems, there exists a broad range of design techniques in order to improve performance and robustness. These include adaptive control, optimal control, feedback linearization, and passivity. However, it may be difficult to apply such techniques in large class of underactuated systems because sometimes these systems are not linearizable using smooth feedback [7] also due to the existence of unstable hidden modes in some systems. Brockett [8] also provided a necessary condition for the hold of stable smooth feedback law, but this condition is not satisfied in the majority of underactuated systems. Nevertheless, control design experts have employed approximate feedback linearization [9–11] and backstepping control [12]. Passivity-based methodology is also used to control such systems but the main drawback in this technique is its narrow range of applications [13]. Sliding mode control is also proposed for the class of underactuated systems [6] but the problem with sliding mode control is presence of chattering.

The aforementioned design strategies were quite suitable and resulted in satisfactory results but it is worthy to note that the system often becomes too sensitive to disturbance in the reaching phase of sliding mode strategy that the system may even become unstable. Therefore, in order to get rid of this issue the integral sliding mode strategy was proposed [14–16]. In this paper a robust integral sliding mode control (RISMC) approach for underactuated systems is proposed. The benefit of this strategy is enhancement of robustness from initial time instant. It also suppresses the well-known chattering phenomenon across the manifold. Before the design presentation, the system is suitably transformed into special formats. An integral sliding mode strategy is proposed for both the cases along with their comprehensive stability analysis. The proposed technique is practically implemented on the ball and beam system to authenticate the affectivity and efficiency of the designed algorithm. Note that in this paper our contributions are twofold. The first one is the development of the RISMC and the second one is the practical results of the system on the said system. The rest of the paper is organized as follows. In Section 2, the problem is formulated into two special formats which further simplify the design methodology. In Section 3, the integral sliding mode strategy for both the cases is discussed in detail accompanied by their respective stability analysis in terms of Lyapunov theory. Section 4 presents the development of the control laws, simulation, and practical results of the ball and beam system. Section 5 concludes the overall efforts being made in this study. In the end more relevant recent articles are enlisted.

2. Problem Formulation

The dynamic equations which govern the motion of the class of underactuated system can be presented as

$$\begin{aligned} &J(q)\ddot{q} + C(q,\dot{q})\dot{q} + G(q) + F(\dot{q}) \\ &\quad = B(\tau + \delta(q,\dot{q},t)), \end{aligned} \tag{1}$$

where q, $\dot{q}$, and $\ddot{q}$ are n-dimensional position, velocity, and acceleration vectors and $J(q)$, $C(q,\dot{q})$, $G(q)$, and $F(\dot{q})$ represent the inertia, Coriolis, gravitational, and fractional torques matrices, respectively. τ is the measured control input, and $\delta(q,\dot{q},t)$ represents the uncertainties in the control input channel whereas B is the control input channel.

It is assumed that $\operatorname{rank}(J^{-1}(q)B) = m$ and the origin is considered to be the equilibrium point for the aforementioned system. Now, the system in (1) can be rewritten in alternate form as follows:

$$\begin{aligned} m_{11}(q)\ddot{q}_1 + m_{12}(q)\ddot{q}_2 + h_1(q,\dot{q}) &= 0, \\ m_{21}(q)\ddot{q}_1 + m_{12}(q)\ddot{q}_2 + h_2(q,\dot{q}) &= \tau, \end{aligned} \tag{2}$$

where $q = [q_1, q_2]^T$ represents the states of the system and q and $\dot{q}$ point to the states. In order to design a control law, the system in (2) can be transformed into two formats which are described in the subsequent study.

2.1. System in Cascaded Form. Following some algebraic manipulations, the system in (2) may be written in cascaded form as follows [17]:

$$\begin{aligned} \dot{x}_1 &= x_2 + d_1, \\ \dot{x}_2 &= f_1(x_1,x_2,x_3,x_4) + d_2, \end{aligned} \tag{3}$$

$$\begin{aligned} \dot{x}_3 &= x_4, \\ \dot{x}_4 &= f_2(x_1,x_2,x_3,x_4) + b(x_1,x_2,x_3,x_4)\tau + d_3, \end{aligned} \tag{4}$$

where x_1, x_2, x_3, x_4 are measurable states of the systems such that x_1 and x_2 are pointing to the position and velocity of the indirect actuated system (3) while x_3 and x_4 represent the position and velocity of the directly actuated system (4). τ represents the controlled signal, as already discussed, to the system (4) input. Owing to the assumption stated immediately after (1), the inverse of b exists. The nonlinear functions $f_1, f_2 : R^{4n} \to R^n$, $b : R^{4n} \to R^{n\times n}$ are smooth in nature. Now, following the procedure of [6], the disturbances $d_1 d_2 d_3$ are deliberately introduced to get an approximate controllable canonical form. Note that practical systems like inverted pendulum [18], TORA [4], VTOL (vertical take-off and landing) aircraft [17], and quad rotor [19] can be put in the form presented in (3) and (4). Before proceeding to the control design of the above cascaded form, the following assumptions are made.

Assumption 1. Assume that

$$f_1(0,0,0,0) = 0. \tag{5}$$

This condition is necessary for the system origin to be in equilibrium point when the system is operated in closed loop.

Assumption 2. $\partial f_1/\partial x_3$ is invertible or $\partial f_1/\partial x_4$ is invertible.

Assumption 3. $f_1(0,0,x_3,x_4) = 0$ is an asymptotically stable manifold, that is, x_3, and x_4 approaches zero.

Note that Assumptions 2 and 3 lie in the category of nonnecessary conditions. These are only used when one needs to furnish the closed loop system with a sliding mode controller (see for details [6]).

2.2. Input Output Form. The system in (3) and (4) can be transformed into the following input output form while following the procedure reported in [16]. Let us assume that the system has a nonlinear output $y = h(x)$. To this end we denote

$$\begin{aligned} L_f h(x) &= \frac{\partial h(x)}{\partial x} f(x) = \nabla h(x) f(x), \\ L_{f_\tau} h(x) &= \frac{\partial h(x)}{\partial x} f_\tau = \nabla h(x) f_\tau. \end{aligned} \tag{6}$$

Recursively, it can be written as

$$\begin{aligned} L_f^0 h(x) &= h(x), \\ L_f^j h(x) &= L_f\left(L_f^{j-1} h(x)\right) = \nabla\left(L_f^{j-1} h(x)\right) f(x). \end{aligned} \tag{7}$$

Assume that the system reported in (3)-(4) has a relative degree "r" with respect to the defined nonlinear output. Therefore, owing to [20], one has

$$y^{(r)} = L_f^r h(x) + L_g\left(L_f^{r-1} h(x)\right)\tau + \zeta(x,t) \tag{8}$$

subject to the following conditions:

(1) $L_g(L_f^i h(x)) = 0\ \forall x \in B$, where B indicates the neighborhood of x_0 for $i < r-1$;

(2) $L_g(L_f^{r-1} h(x)) \neq 0$, where $\zeta(x,t)$ represents the matched unmodeled uncertainties. System (8), by defining the transformation $y^{(i-1)} = \xi_i$ [21], can be put in the following form:

$$\begin{aligned} \dot{\xi}_1 &= \xi_2, \\ \dot{\xi}_2 &= \xi_3, \\ &\vdots \\ \dot{\xi}_n &= \varphi\left(\widehat{\xi}, \widehat{\tau}\right) + \gamma\left(\widehat{\xi}\right)\left\{\tau + \Delta G_m\left(\widehat{\xi}, \widehat{\tau}, t\right)\right\}, \end{aligned} \tag{9}$$

where the transformed states $\widehat{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)$ are phase variables, τ is the control input, and $\Delta G_m(\widehat{\xi}, \widehat{\tau}, t)$ represents matched uncertainties. It is worthy to notice that the inverted pendulum and the ball and beam systems can be replaced in the aforementioned form.

Note that both the formats are ready to design the control law for these systems. In the next section, we outline the design procedure for both the forms.

3. Control Law Design

The control design for the forms presented in (3)-(4) and (9) is carried out in this section which we claim as our main contribution in this paper. The main objective in this work is to enhance the robustness of the system from the very beginning of the process which is the beauty of integral sliding mode control. In general, the integral sliding mode control law appears as follows [14]. In the subsequent subsections, the authors aim to present the design procedure.

3.1. Integral Sliding Mode. This variant of sliding mode possesses the main features of the sliding mode like robustness and the existence chattering across the switching manifold. On the other hand, the sliding mode occurs from the very start which, consequently, provides insensitivity of disturbance from the beginning. The control law can be expressed as follows:

$$\tau = \tau_0 + \tau_1, \tag{10}$$

where the first component on the right hand side of (10) governs the systems dynamics in sliding modes whereas the second component compensates the matched disturbances. Now, the aim is to present the design of the aforesaid control components.

3.1.1. Control Design for Case-1. This control design for case-1 is the main obstacle in this subsection. To define both the components, the following terms are defined:

$$\begin{aligned} e_1 &= x_1, \\ e_2 &= x_2, \\ e_3 &= f_1(x_1, x_2, x_3, x_4), \\ e_4 &= \frac{\partial f_1}{\partial x_1} x_2 + \frac{\partial f_1}{\partial x_2} f_1 + \frac{\partial f_1}{\partial x_3} x_4. \end{aligned} \tag{11}$$

Using these new variables, the components of the controller are designed in the following subsection. For the sake of completeness the design of this component is worked out via simple pole placement. Following the design procedure of pole placement method, one gets

$$\tau_0 = -k_1 e_1 - k_2 e_2 - k_3 e_3 - k_4 e_4, \tag{12}$$

where k_i $i = 1, 2, 3, 4$ are the gains of this control component. This control component steers the states of the nominal system to their defined equilibrium. Now, in the subsequent study the design of the uncertainties compensating term is presented. An integral manifold is defined as follows:

$$\sigma = c_1 e_1 + c_2 e_2 + c_3 e_3 + e_4 + z = \sigma_0 + z, \tag{13}$$

where $\sigma_0 = c_1 e_1 + c_2 e_2 + c_3 e_3 + e_4$ represents the conventional sliding manifold which is Hurwitz by definition.

Now, computing $\dot{\sigma}$ along (3)-(4), one has

$$\begin{aligned} \dot{\sigma} = {} & c_1(x_2 + d_1) + c_2(f_1(x_1, x_2, x_3, x_4) + d_2) \\ & + c_3\left(\frac{df_1(x_1, x_2, x_3, x_4)}{dt}\right) + \frac{d}{dt}\left(\frac{\partial f_1}{\partial x_1} x_2\right) \\ & + \frac{\partial f_1}{\partial x_1}\dot{x}_2 + \frac{d}{dt}\left(\frac{\partial f_1}{\partial x_2} f_1\right) + \frac{\partial f_1}{\partial x_2}\dot{f}_1 \\ & + \frac{d}{dt}\left(\frac{\partial f_1}{\partial x_3} x_4\right) + \frac{\partial f_1}{\partial x_3} f_2(x_1, x_2, x_3, x_4) \\ & + \frac{\partial f_1}{\partial x_3} b(x_1, x_2, x_3, x_4)\tau_0 \\ & + \frac{\partial f_1}{\partial x_3} b(x_1, x_2, x_3, x_4)\tau_1 + \frac{\partial f_1}{\partial x_3} d_3. \end{aligned} \tag{14}$$

Now, choose the dynamics of the integral term as follows:

$$\begin{aligned} \dot{z} = {} & -c_1 x_2 - c_2 f_1(x_1, x_2, x_3, x_4) \\ & - c_3\left(\frac{df_1(x_1, x_2, x_3, x_4)}{dt}\right) - \frac{d}{dt}\left(\frac{\partial f_1}{\partial x_1} x_2\right) \\ & - \frac{\partial f_1}{\partial x_1}\dot{x}_2 - \frac{d}{dt}\left(\frac{\partial f_1}{\partial x_2} f_1\right) - \frac{\partial f_1}{\partial x_2}\dot{f}_1 \\ & - \frac{d}{dt}\left(\frac{\partial f_1}{\partial x_3} x_4\right) - \frac{\partial f_1}{\partial x_3}(x_1, x_2, x_3, x_4)\tau_0. \end{aligned} \tag{15}$$

The expression of the term which compensates the uncertainties may be written as follows:

$$\tau_1 = -\left(\frac{\partial f_1}{\partial x_3} b\left(x_1, x_2, x_3, x_4\right)\right)^{-1} \cdot \left(\frac{\partial f_1}{\partial x_3} f_2\left(x_1, x_2, x_3, x_4\right) + K\operatorname{sign}(\sigma)\right). \tag{16}$$

The overall controller will look like

$$\tau = -k_1 e_1 - k_2 e_2 - k_3 e_3 - k_4 e_4 - \left(\frac{\partial f_1}{\partial x_3} b\left(x_1, x_2, x_3, x_4\right)\right)^{-1} \cdot \left(\frac{\partial f_1}{\partial x_3} f_2\left(x_1, x_2, x_3, x_4\right) + K\operatorname{sign}(\sigma)\right). \tag{17}$$

The constants c_i's are control gains which are selected intelligently according to bounds. In the forthcoming paragraph, the stability of the presented integral sliding mode is carried out in the presence of the disturbances and uncertainties. Consider the following Lyapunov candidate function:

$$V = \frac{1}{2}\sigma^2. \tag{18}$$

The time derivative of this function along dynamics (11) becomes

$$\begin{aligned}\dot{V} = \sigma\dot{\sigma} = \sigma\Bigg(& c_1\left(x_2 + d_1\right) \\ &+ c_2\left(f_1\left(x_1, x_2, x_3, x_4\right) + d_2\right) \\ &+ c_3\left(\frac{df_1\left(x_1, x_2, x_3, x_4\right)}{dt}\right) + \frac{d}{dt}\left(\frac{\partial f_1}{\partial x_1} x_2\right) \\ &+ \frac{\partial f_1}{\partial x_1}\dot{x}_2 + \frac{d}{dt}\left(\frac{\partial f_1}{\partial x_2} f_1\right) + \frac{\partial f_1}{\partial x_2}\dot{f}_1 + \frac{d}{dt}\left(\frac{\partial f_1}{\partial x_3} x_4\right) \\ &+ \frac{\partial f_1}{\partial x_3} f_2\left(x_1, x_2, x_3, x_4\right) + \frac{\partial f_1}{\partial x_3}\left(x_1, x_2, x_3, x_4\right)\tau_0 \\ &+ \frac{\partial f_1}{\partial x_3} b\left(x_1, x_2, x_3, x_4\right)\tau_1 + \frac{\partial f_1}{\partial x_3} d_3\Bigg).\end{aligned} \tag{19}$$

The substitution of (15)-(16) results in the following form:

$$\begin{aligned}\dot{V} &\le -|\sigma|\,\eta_1 < 0 \\ \text{or } \dot{V} &+ \sqrt{2}\eta_1\sqrt{V} < 0\end{aligned} \tag{20}$$

subject to $K \ge [\|(\partial f_1/\partial x_3)d_3 + c_1 d_1 + c_2 d_2\| + \eta]$.

This expression confirms the enforcement of the sliding mode from the very beginning of the process, that is, $\sigma \to 0$ in finite time. Now, we proceed to the actual system's stability. If one considers e_1 as the output of the system, then e_2, e_3, and e_4 become the successive derivatives of e_1. Whenever $\sigma = 0$ is achieved, the dynamics of the transformed system (11) will converge asymptotically to zero under the action of the control component (12) [22]. That is, in closed loop, the transformed system dynamics will be operated under (12) as follows:

$$\begin{bmatrix}\dot{e}_1\\ \dot{e}_2\\ \dot{e}_3\\ \dot{e}_4\end{bmatrix} = \begin{bmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ -k_1 & -k_2 & -k_3 & -k_4\end{bmatrix}\begin{bmatrix}e_1\\ e_2\\ e_3\\ e_4\end{bmatrix} \tag{21}$$

and the disturbances will be compensated via (16).

The asymptotic convergence of e_1, e_2, e_3, and e_4 to zero means the convergence of the indirectly actuated system (3) to zero. On the other hand, the states of the directly actuated system (4) will remain bounded; that is, state of (4) will have some nonzero value in order to keep e_1 at zero. Thus, the overall system is stabilized and the desired control objective is achieved.

3.2. Control Design for Case-2. The nominal system related to (9) can be replaced in the subsequent alternative form

$$\begin{aligned}\dot{\xi}_1 &= \xi_2\\ \dot{\xi}_2 &= \xi_3\\ &\vdots\\ \dot{\xi}_r &= \chi\left(\widehat{\xi}, \tau\right) + \tau,\end{aligned} \tag{22}$$

where $\chi(\widehat{\xi}, \tau) = \varphi(\xi, \tau) + (\gamma(\widehat{\xi}) - 1)\tau$. It is assumed that $\chi(\widetilde{\xi}, \widehat{\tau}, \tau^{(k)}) = 0$ at $t = 0$ in addition to the next supposition that (22) is governed by τ_0:

$$\begin{aligned}\dot{\xi}_1 &= \xi_2\\ \dot{\xi}_2 &= \xi_3\\ &\vdots\\ \dot{\xi}_r &= \tau_0\end{aligned} \tag{23}$$

or

$$\dot{\xi} = A\xi + B\tau_0, \tag{24}$$

where

$$A = \begin{bmatrix}0_{(r-1)\times 1} & I_{(r-1)\times(r-1)}\\ 0_{1\times 1} & 0_{1\times(r-1)}\end{bmatrix}, \quad B = \begin{bmatrix}0_{(r-1)\times 1}\\ 1\end{bmatrix}. \tag{25}$$

Once again, following the pole placement procedure, one may have, for the sake of simplicity, the input τ_0 which is designed via pole placement, that is,

$$\tau_0 = -K_0^T \xi. \tag{26}$$

Now to get the desired robust performance, the following sliding manifold of integral type [14] is defined:

$$\sigma(\xi) = \sigma_0(\xi) + z, \tag{27}$$

where $\sigma_0(\xi)$ is the usual sliding surface and z is the integral term. The time derivative of (27) along (9) yields

$$\begin{aligned} \dot{z} &= -\left(\sum_{i=1}^{r-1} c_i \xi_{i+1} + \tau_0\right), \\ z(0) &= -\sigma_0(\xi(0)), \end{aligned} \tag{28}$$

$$\tau_1 = \frac{1}{\gamma(\widehat{\xi})}\left(-\varphi(\widehat{\xi}, \tau) - \left(\gamma(\widehat{\xi}) - 1\right)\tau_0 - K \operatorname{sign}\sigma\right). \tag{29}$$

This control law enforces sliding mode along the sliding manifold defined in (27). The constant K can be selected according to the subsequent stability analysis.

Thus, the final control law becomes

$$\begin{aligned} \tau_1 &= -K_0^T \xi \\ &\quad + \frac{1}{\gamma(\widehat{\xi})}\left(-\varphi(\widehat{\xi}, u) - \left(\gamma(\widehat{\xi}) - 1\right)\tau_0 - K \operatorname{sign}\sigma\right). \end{aligned} \tag{30}$$

Theorem 4. *Consider that $|\Delta G_m(\widehat{y}, \widehat{u}, t)| \leq \beta_1$ are satisfied; then the sliding mode against the switching manifold $\sigma = 0$ can be ensured and one has*

$$K \geq \left[K_M \beta_1 + \eta_1\right], \tag{31}$$

where η_1 is a positive constant.

Proof. To prove that the sliding mode can be enforced in finite time, differentiating (22) along the dynamics of (3)-(4), and then substituting (30), one has

$$\begin{aligned} \dot{\sigma}(\xi) &= \sum_{i=1}^{r-1} c_i \xi_{i+1} + \tau_0 - K \operatorname{sign}\sigma + \gamma(\widehat{\xi}) \Delta G_m(\widehat{\xi}, \widehat{\tau}, t) \\ &\quad + \dot{z}. \end{aligned} \tag{32}$$

Substituting (28) in (32), and then rearranging, one obtains

$$\dot{\sigma}(\xi) = -K \operatorname{sign}\sigma + \gamma(\widehat{\xi}) \Delta G_m(\widehat{\xi}, \widehat{\tau}, t). \tag{33}$$

Now, the time derivative of the Lyapunov candidate function $V = (1/2)\sigma^2$, with the use of the bounds of the uncertainties, becomes

$$\dot{V} \leq -|\sigma| \left[-K + \left|\gamma(\widehat{\xi}) \Delta G_m(\widehat{\xi}, \widehat{\tau}, t)\right|\right]. \tag{34}$$

This expression may also be written as

$$\begin{aligned} &\dot{V} \leq -|\sigma| \eta_1 < 0 \\ &\text{or } \dot{V} + \sqrt{2}\eta_1 \sqrt{V} < 0, \end{aligned} \tag{35}$$

provided that

$$K \geq \left[K_M \beta_1 + \eta_1\right]. \tag{36}$$

The inequality in (35) presents that $\sigma(\xi)$ approaches zero in a finite time t_s [23], such that

$$t_s \leq \sqrt{2}\eta_1^{-1} \sqrt{V}(\sigma(0)) \tag{37}$$

which completes the proof. □

4. Illustrative Example

The control algorithms presented in Section 3 are applied to the control design of a ball and beam system. The assessment of the proposed controller, for the ball and beam system, is carried out on the basis of output tracking, robustness enhancement via the elimination of reaching phase, and chattering-free control input in the presence of uncertainties.

4.1. Description of the Ball and Beam System. The ball and beam system is a very sound candidate of the class of underactuated nonlinear system. It is famous because of its nonlinear nature and due to its wide range of applications in the existing era like passenger cabin balancing in luxury cars, balancing of liquid fuel in vertical take-off objects. In terms of control scenarios, it is an ill-defined relative degree system which, to some extent, does not support input output linearization. A schematic diagram with their typical parameters of the ball and beam system is displayed in the adjacent Figure 1 and Table 1, respectively. In this study the authors use the equipment manufacture by GoogolTech. In general this system is equipped with a metallic ball, which is let free to roll on a rod having a specified length, having one end fixed and the other end moved up and down via an electric servomotor. The position of the ball can be measured via different techniques. The measured position is used as feedback to the system and accordingly the motor moves the beam to balance the ball at user defined location.

The motion governing equations of this system are given below which are adopted from [24]:

$$\begin{aligned} &\left(mr^2 + C_1\right)\ddot{\beta} + \left(2mr\dot{r} + C_2\right)\dot{\beta} \\ &\quad + \left(mgr + \frac{L}{2}Mg\right)\cos\beta = \tau, \\ &C_4 \ddot{r} - r\dot{\beta}^2 + g\sin\beta = 0, \end{aligned} \tag{38}$$

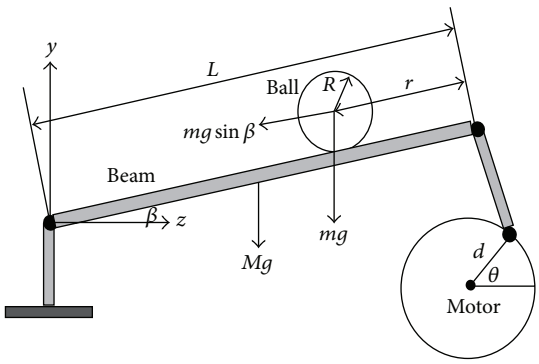


Figure 1: Schematic diagram of the ball and beam system.

where $\theta(t)$ angle is subtended to make the ball stable, the lever angle is represented by $\beta(t)$, $r(t)$ is the position of the ball on the beam, and $v_{\text{in}}(t)$ is the input voltage of the motor whereas the controlled input appears mathematically via the expression $\tau(t) = C_3 v_{\text{in}}(t)$ in the dynamic model.

The derived parameters used in the dynamic model of this system are represented by C_1, C_2, C_3, and C_4 with the following mathematical relations [25]:

$$C_1 = \frac{R_m \times J_m \times L}{C_m \times C_b \times d} + J_1, \tag{39}$$

$$C_2 = \frac{L}{d}\left(\frac{C_m \times C_b}{R_m} + C_b + \frac{R_m \times J_m}{C_m \times C_g}\right), \tag{40}$$

$$C_3 = 1 + \frac{C_m}{R_m}, \tag{41}$$

$$C_4 = \frac{7}{5}. \tag{42}$$

The equivalent state space model of this is described as follows by assuming $x_1 = r$ (position of ball), $x_2 = \dot{r}$ (rate of change of position), $x_3 = \beta$ (beam angle), and $x_4 = \dot{\beta}$ (the rate of change of angle of the motor):

$$\begin{aligned}
\dot{x_1} &= x_2, \\
\dot{x_2} &= \frac{1}{C_4}\left(-g\sin\left(x_3\right)\right), \\
\dot{x_3} &= x_4, \\
\dot{x_4} &= \frac{1}{mx_1^2 + C_1}\Big(\tau - \left(2mx_1x_2 + C_2\right)x_4 \\
&\quad - \left(mgx_1 + \frac{L}{2}Mg\right)\cos x_3\Big).
\end{aligned} \tag{43}$$

Now, the output of interest is $y = x_1$, which represents the position of the ball. This representation is similar to that reported in (3)-(4). In the next discussion the controller design is outlined.

4.2. Controller Design. Following the procedure outlined in Section 3, the authors proceed as follows:

$$\begin{aligned}
y &= x_1, \\
\dot{y} &= x_2, \\
\ddot{y} &= -\frac{g}{C_4}\sin\left(x_3\right), \\
y^{(3)} &= -\frac{g}{C_4}x_4\cos\left(x_3\right), \\
y^{(4)} &= \frac{1}{C_4\left(mx_1^2 + C_1\right)}\left[-\tau\cos x_3 + \left(2mx_1x_2 + C_2\right)x_4\cos x_3 + \left(mgx_1 + \frac{L}{2}Mg\right)\cos^2 x_3 + x_4^2\left(mx_1^2 + C_1\right)\sin x_3\right], \\
y^{(4)} &= f_s + h_s\tau, \\
f_s &= \frac{g}{C_4}\left[\frac{\left(2mx_1x_2 + C_2\right)x_4 + \left(mgx_1 + (L/2)Mg\right)\cos^2 x_3 + x_4^2\sin x_3}{mx_1^2 + C_1}\right], \\
h_s &= \frac{-g\cos x_3}{C_4\left(mx_1^2 + C_1\right)}.
\end{aligned} \tag{44}$$

Table 1: Parameters and values used in equations.

Parameter	Description	Nominal values	Units
g	Gravitational acceleration	9.81	m/s^2
m	Mass of ball	0.07	kg
M	Mass of beam	0.15	kg
L	Length of beam	0.4	m
R_m	Resistance of armature of the motor	9	Ω
J_m	Moment of inertia of motor	7.35×10^{-4}	Nm/rad/s^2
C_m	Torque constant of motor	0.0075	Nm/A
C_g	Gear ratio	4.28	—
d	Radius of arm connected to servomotor	0.04	m
J_1	Moment of inertia of beam	0.001	kg m^2
C_b	Back emf constant value	0.5625	V/rad/s

Now, writing this in the controllable canonical form (phase variable form), one may have

$$\begin{aligned} \dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= \xi_3 \\ &\vdots \\ \dot{\xi}_4 &= \varphi\left(\widehat{\xi}\right) + \gamma\left(\widehat{\xi}\right)\tau + \gamma\left(\widehat{\xi}\right)\Delta G_m\left(\widehat{\xi}, \widehat{\tau}, t\right), \end{aligned} \tag{45}$$

where $y^{(i-1)} = \xi_i$,

$$\begin{aligned} \varphi\left(\widehat{\xi}\right) = \frac{1}{C_4\left(mx_1^2 + C_1\right)} &\left[\left(2mx_1x_2 + C_2\right)x_4\cos x_3\right. \\ &+ \left(mgx_1 + \frac{L}{2}Mg\right)\cos^2 x_3 \\ &\left. + x_4^2\left(mx_1^2 + C_1\right)\cos x_3\right], \end{aligned} \tag{46}$$

$\gamma(\widehat{\xi})\tau = -\tau\cos x_3$, and $\gamma(\widehat{\xi})\Delta G_m(\widehat{\xi}, \widehat{\tau}, t)$ represents the model uncertainties. Here we discuss ISMC on ball and beam system with fixed step tracking as well as variable step tracking. The integral manifold is defined as follows:

$$\sigma = c_1\xi_1 + c_2\xi_2 + c_3\xi_3 + \xi_4 + z. \tag{47}$$

The expression of the overall controller which becomes $\dot{\sigma}$ will be as follows:

$$\begin{aligned} \tau_1 &= -k_1\xi_1 - k_2\xi_2 - k_3\xi_3 - k_4\xi_4 \\ &\quad + \frac{1}{\gamma\left(\widehat{\xi}\right)}\left(-\varphi\left(\widehat{\xi}\right) - \left(\gamma\left(\widehat{\xi}\right) - 1\right)\tau_0 - K\operatorname{sign}\sigma\right), \\ \dot{\sigma} &= c_1\dot{\xi}_1 + c_2\dot{\xi}_2 + c_3\dot{\xi}_3 + f_s + h_s\tau_0 + h_s\tau_1 + \dot{z}, \\ \dot{z} &= -c_1x_2 + \frac{c_2 g}{C_4}\sin x_3 + \frac{c_3 g}{C_4}x_4\cos x_3 - \gamma\left(\widehat{\xi}\right)\tau_0 \\ &\quad - \varphi\left(\widehat{\xi}\right). \end{aligned} \tag{48}$$

As the authors are performing the reference tracking here, therefore, the integral manifold and the controller will appear as follows:

$$\sigma = c_1\left(\xi_1 - r_d\right) + c_2\xi_2 + c_3\xi_3 + \xi_4 + z, \tag{49}$$

$$\begin{aligned} \tau_1 &= -k_1\left(\xi_1 - r_d\right) - k_2\xi_2 - k_3\xi_3 - k_4\xi_4 \\ &\quad + \frac{1}{\gamma\left(\widehat{\xi}\right)}\left(-\varphi\left(\widehat{\xi}, \tau\right) - \left(\gamma\left(\widehat{\xi}\right) - 1\right)\tau_0 - K\operatorname{sign}(\sigma)\right), \end{aligned} \tag{50}$$

where r_d is the desired reference with $\dot{r}_d, \ddot{r}_d, \dddot{r}_d$ being bounded.

4.3. Simulation Results. The simulation study of the system is carried out by considering the reference tracking of a square wave signal and sinusoidal wave signal. In the subsequent paragraph their respective results will be demonstrated in detail.

In case the efforts are directed to track a fixed square wave signal in the presence of disturbances, the initial conditions of the system were set to $x_1(0) = 0.4$, $x_2(0) = x_3(0) = x_4(0) = 0$. Furthermore, the square wave was defined in the simulation code as follows:

$$r_d(t) = \begin{cases} 20\text{ cm} & 0 \le t \le 19 \\ 14\text{ cm} & 20 \le t \le 39 \\ 20\text{ cm} & 40 \le t \le 60. \end{cases} \tag{51}$$

The gains of the proposed controller presented from (39) to (41) are chosen according to Table 2.

The output tracking performance of the proposed control input, when a square wave is used as desired reference output, is shown in Figure 2. It can be clearly examined that the performance is very appealing in this case. The corresponding sliding manifold profile is displayed in Figure 3 which clearly indicates that the sliding mode is established from the very beginning of the processes which in turn results in enhanced robustness. The controlled input signal's profile is depicted in Figure 4 with its zoomed profile as shown in Figure 5. It is obvious from both the figures that the control input derives the system with suppressed chattering phenomenon which is

Table 2: Parametric values used in the square wave tracking.

Constants	C_1	C_2	C_3	K_1	K_2	K_3	K_4	K
Values	1.2	1.2	0.11	402.98	250.18	60	4.1	5

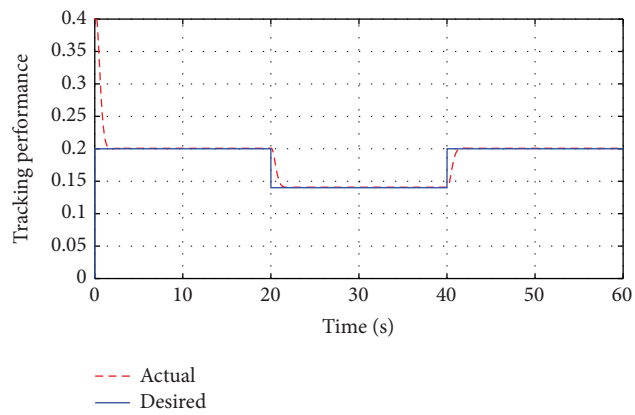


Figure 2: Output tracking performance when a square wave is used as reference/desired output.

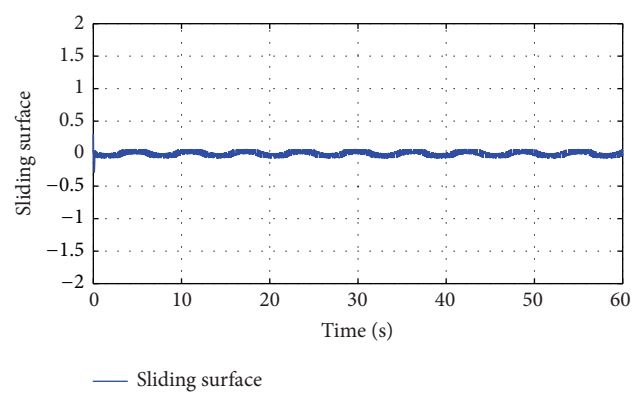


Figure 3: Sliding manifold convergence profile in case of square wave tracking.

tolerable for the system actuators health. Now, from this case study, it is concluded that integral sliding mode approach is an interesting candidate for this class.

In this case study, once again, efforts are focused on the tracking of a sinusoidal signal, which is defined as $r_d(t) = \sin(t)$, in the presence of disturbances. Like the previous case study, the initial conditions of the system were set to $x_1(0) = 0.4$, $x_2(0) = x_3(0) = x_4(0) = 0$. In addition, the gains of the proposed controller presented in (50) are chosen according to Table 3.

The output tracking performance of the proposed control input, when a sinusoidal signal is considered as desired reference output, is shown in Figure 6. It can be clearly seen that the performance is excellent in this scenario. The corresponding sliding manifold profile is displayed in Figure 7 which confirms the establishment of sliding modes from the starting instant and, consequently, enhancement of robustness. The controlled input signal's profile is depicted in Figure 8. It is obvious from the figure that the control input

Table 3: Parametric values used in the sinusoid wave tracking.

Constants	C_1	C_2	C_3	K_1	K_2	K_3	K_4	K
Values	1.2	1.2	0.11	402.98	250.18	230	4.9	5

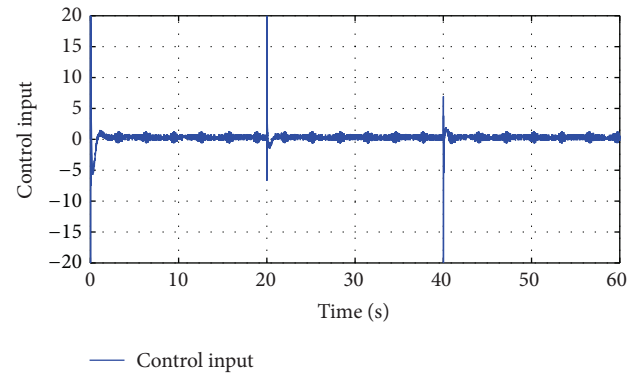


Figure 4: Control input in square wave reference tracking.

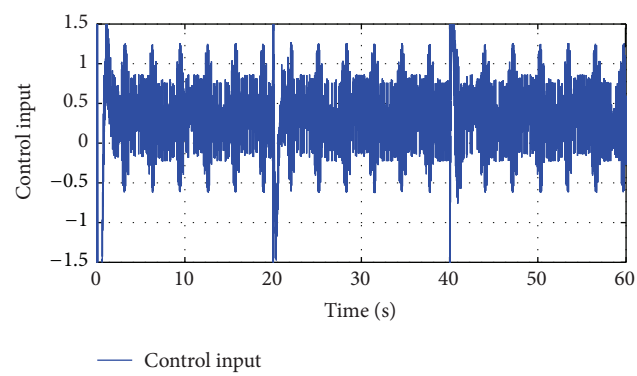


Figure 5: Zoom profile of the control input depicted in Figure 4.

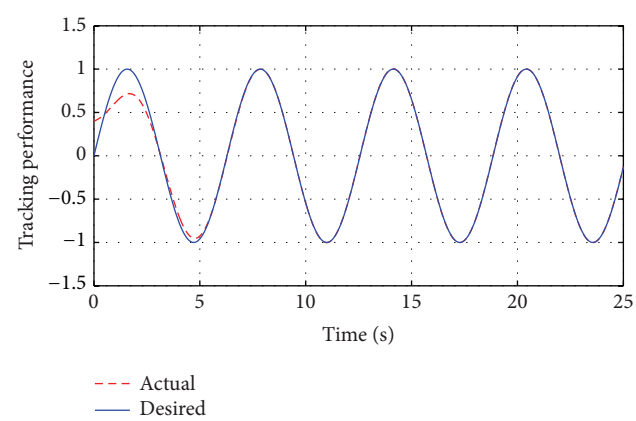


Figure 6: Output tracking performance when a sinusoidal wave is used as reference/desired output.

evolves with suppressed chattering phenomenon which, once again, makes this design strategy a good candidate for the class of these underactuated systems.

4.4. Implementation Results. The control technique proposed in this paper is implemented on the actual apparatus using

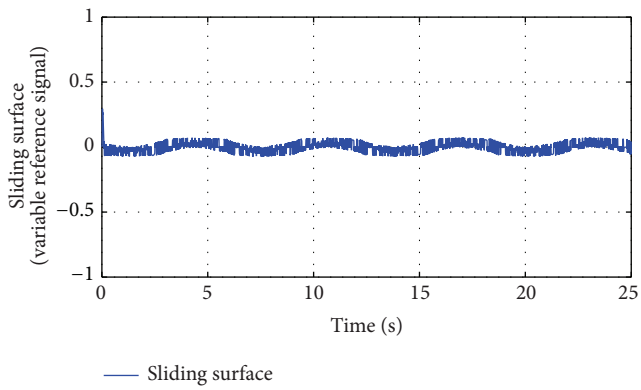


Figure 7: Sliding manifold convergence profile in case of sinusoidal wave tracking.

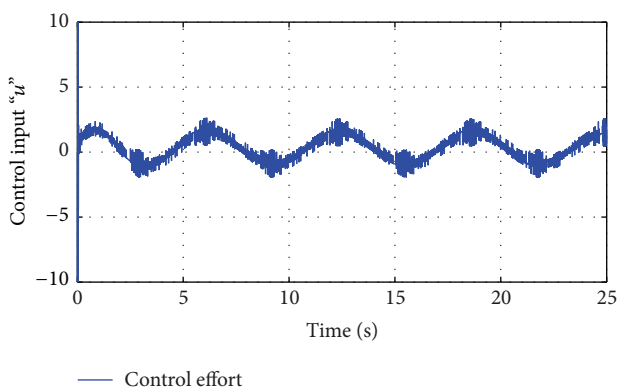


Figure 8: Control input in sinusoidal wave reference tracking.

the MATLAB environment. The detailed discussions are presented below.

4.4.1. Experimental Setup Description. The experiment setup is equipped by GoogolTech GBB1004 with an electronic control box. The beam length is 40 cm along with mass of ball, that is, 28 g, and an intelligent IPM100 servo driver which is used for moving the ball on the beam. The experimental setup is shown in Figure 9.

The input given to apparatus is the voltage $v_{\text{in}}(t)$ and the output is the position of the motor $\theta(t)$, which, in other words, is an input for the positioning of the ball on the beam. This apparatus uses potentiometer mounted within a slot inside the beam to sense the position of the ball on the beam. The measured position along the beam is fed to the A/D converter of IPM100 motion drive.

The power module used in GoogolTech requires 220 V and 10 A input. Note that the control accuracy of this manufactured apparatus lies within the range of ±1 mm. The typical parameters values are listed in Table 1. The environment used here includes Windows XP as an operating system and MATLAB 7.12/Simulink 7.7. Furthermore, the sampling time used in forthcoming practical results was 2 ms. In the experimental processes, the proposed controllers need velocity measurements which are, in general, not available. One may

Figure 9: Experimental setup of the ball and beam equipped via GoogolTech GBB1004.

use different kind of velocity observers/differentiator for the velocity estimation [16]. In order to make the implementation easy and simple, a derivative block of the Simulink environment is used to provide the corresponding velocities measurements. Now, we are ready to discuss the results of the system.

In this experiment, the initial conditions were set to $x_1(0) = 0.28$, $x_2(0) = x_3(0) = x_4(0) = 0$. The reference signal which is needed to be tracked is being defined in (51). In Figures 10 and 11, the tracking performance is shown. The results reveal that the actual signal $x_1(t)$ is pretty close to the desired signal $r_d(t)$ with a steady state error which is approximately ±0.001 m. The existence of this error is because of the apparatus.

The observations of these tracking results make it clear that the practically implemented results have very close resemblance with the simulation result presented in Figure 2. The error convergence depends on the initial conditions of the ball on the beam. If the ball is placed very close to the desired reference value then it will take little time to reach the desired position. On the other hand, the convergence to the desired position will take considerable time if the initial condition is chosen far away from the desired values. This phenomenon of convergence is according to the equipment design and structure.

The sliding manifold convergence and the control input are shown in Figures 12 and 13, respectively. The control input and the sliding manifolds show some deviations in the first second. This deviation occurs because the ball on the beam, being placed anywhere on the beam, is first moved to one side of the beam and then ball moved to the desired position. The zoomed profile of the control input, being displayed in Figure 14, shows high frequency vibration (chattering) of magnitude ±0.07. This makes the proposed control design algorithm an appealing candidate for this class of nonlinear systems. The gains of the controller being used in this experiment are displayed in Table 4.

5. Conclusion

The control of underactuated systems, because of their less number of actuators than the degree of freedom, is an

Table 4: Parametric values used in implementation.

Constants	C_1	C_2	C_3	K_1	K_2	K_3	K_4	K
Values	8	5	1	3	15	3	1	4

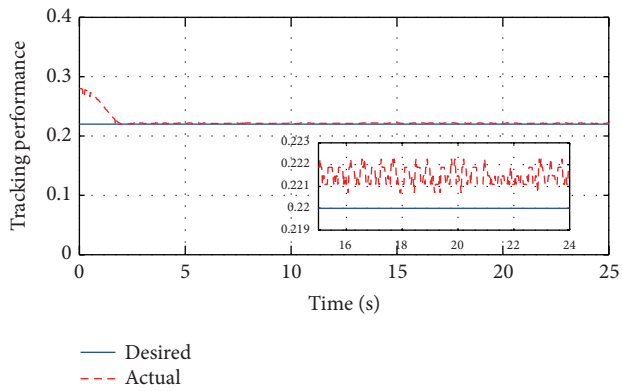


Figure 10: Output tracking performance when $r_d = 22$ cm is set as reference/desired output.

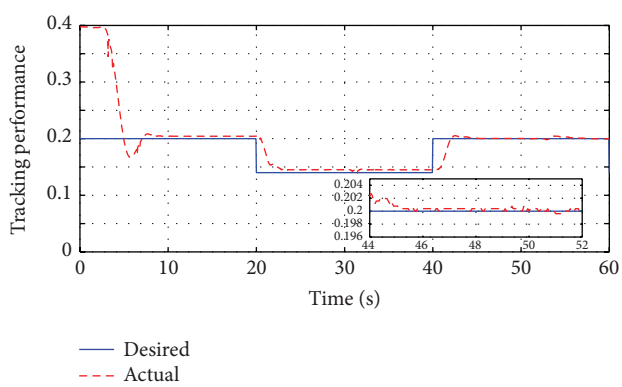


Figure 11: Output tracking performance when a square wave is used as reference/desired output.

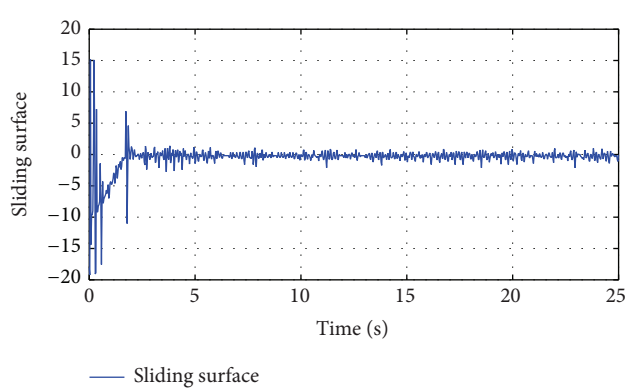


Figure 12: Sliding surface of practical system.

interesting objective among the researchers. In this work, an integral sliding mode control approach, due to its robustness from the very beginning of the process, is employed for the control design of this class. The design of the integral manifold relied upon a transformed form. The benefit of the transformed form is that it makes the design strategy easy and

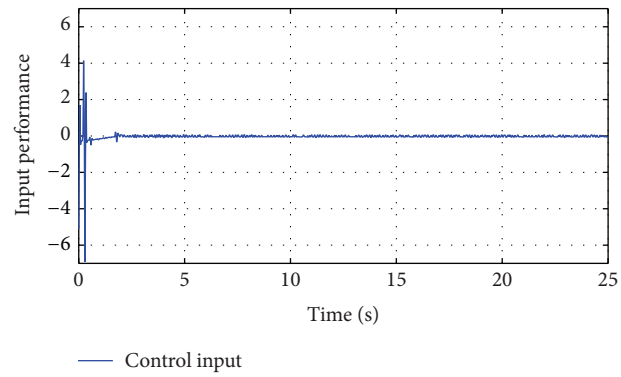


Figure 13: Control input for reference tracking.

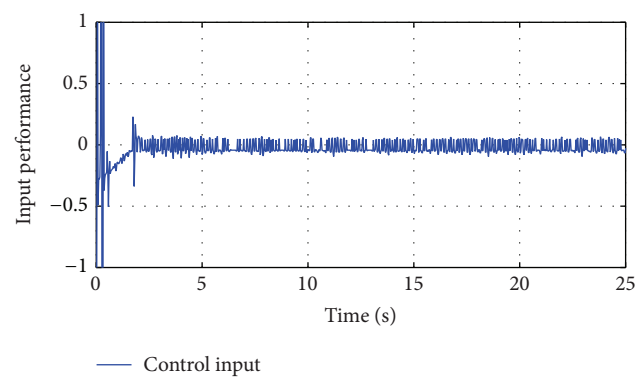


Figure 14: Zoom profile of the control input depicted in Figure 13.

simple. The stability analysis and experimental results of the proposed control laws are presented, which convey the good features and demand the proposed approach when the system operates under uncertainties.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] S. Mahjoub, F. Mnif, and N. Derbel, "Set point stabilization of a 2DOF underactuated manipulator," *Journal of Computers*, vol. 6, no. 2, pp. 368–376, 2011.

[2] F. Mnif, "VSS control for a class of underactuated Mechanical systems," *International Journal of Computational Cognition*, vol. 3, no. 2, pp. 14–18, 2005.

[3] J. Hauser, S. Sastry, and P. Kokotovic, "Nonlinear control via approximate input–output linearization: the ball and beam example," *IEEE Transactions on Automatic Control*, vol. 37, no. 3, pp. 392–398, 1992.

[4] M. Jankovic, D. Fontanine, and P. V. Kokotovic, "TORA example: cascade and passitivity-based control designs," *IEEE Transactions on Control Systems Technology*, vol. 6, pp. 4347–4351, 1996.

[5] T. Sugie and K. Fujimoto, "Control of inverted pendulum systems based on approximate linearization: design and experiment," in *Proceedings of the 33rd IEEE Conference on Decision and Control*, vol. 2, pp. 1647–1648, December 1994.

[6] R. Xu and Ü. Özgüner, "Sliding mode control of a class of underactuated systems," *Automatica*, vol. 44, no. 1, pp. 233–241, 2008.

[7] M. W. Spong, "The swing-up control problem for the acrobat," *IEEE Transactions on Systems Technology*, vol. 15, no. 1, pp. 49–55, 1995.

[8] R. W. Brockett, "Asymptotic stability feedback and stabilization," in *Differential Geometric Control Theory*, pp. 181–191, Birkhäuser, Boston, Mass, USA, 1983.

[9] C. J. Tomlin and S. S. Sastry, "Switching through singularities," *Systems and Control Letters*, vol. 35, no. 3, pp. 145–154, 1998.

[10] W.-H. Chen and D. J. Ballance, "On a switching control scheme for nonlinear systems with ill-defined relative degree," *Systems & Control Letters*, vol. 47, no. 2, pp. 159–166, 2002.

[11] F. Zhang and B. Fernndez-Rodriguez, "Feedback linearization control of systems with singularities," in *Proceedings of the 6th International Conference on Complex Systems (ICCS '06)*, Boston, Mass, USA, June 2006.

[12] D. Seto and J. Baillieul, "Control problems in super-articulated mechanical systems," *IEEE Transactions on Automatic Control*, vol. 39, no. 12, pp. 2442–2453, 1994.

[13] M. W. Spong, *Energy Based Control of a Class of Underactuated Mechanical Systems*, IFAC World Congress, 1996.

[14] V. I. Utkin, *Sliding Mode Control in Electromechanical Systems*, Taylor & Francis, 1999.

[15] M. Rubagotti, A. Estrada, F. Castanos, A. Ferrara, and L. Fridman, "Integral sliding mode control for nonlinear systems with matched and unmatched perturbations," *IEEE Transaction on Automatic Control*, vol. 56, no. 11, pp. 2699–2704, 2011.

[16] Q. Khan, A. I. Bhatti, S. Iqbal, and M. Iqbal, "Dynamic integral sliding mode for MIMO uncertain nonlinear systems," *International Journal of Control, Automation and Systems*, vol. 9, no. 1, pp. 151–160, 2011.

[17] R. Olfati-Saber, "Normal forms for underactuated mechanical systems with symmetry," *IEEE Transactions on Automatic Control*, vol. 47, no. 2, pp. 305–308, 2002.

[18] R. Lozano, I. Fantoni, and D. J. Block, "Stabilization of the inverted pendulum around its homoclinic orbit," *Systems and Control Letters*, vol. 40, no. 3, pp. 197–204, 2000.

[19] E. Altug, J. P. Ostrowski, and R. Mahony, "Control of a quadrotor helicopter using visual feedback," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '02)*, vol. 1, pp. 72–77, Washington, DC, USA, 2002.

[20] A. Isidori, *Nonlinear Control Systems*, Communications and Control Engineering Series, Springer, Berlin, Germany, 3rd edition, 1995.

[21] H. Sira-Ramirez, "On the dynamical sliding mode control of nonlinear systems," *International Journal of Control*, vol. 57, no. 5, pp. 1039–1061, 1993.

[22] Q. Khan, A. I. Bhatti, and A. Ferrara, "Dynamic sliding mode control design based on an integral manifold for nonlinear uncertain systems," *Journal of Nonlinear Dynamics*, vol. 2014, Article ID 489364, 10 pages, 2014.

[23] C. Edwards and S. K. Spurgeon, *Sliding Modes Control: Theory and Applications*, Taylor & Francis, London, UK, 1998.

[24] A. M. Khan, A. I. Bhatti, S. U. Din, and Q. Khan, "Static & dynamic sliding mode control of ball and beam system," in *Proceedings of the 9th International Bhurban Conference on Applied Sciences and Technology (IBCAST '12)*, pp. 32–36, Islamabad, Pakistan, January 2012.

[25] N. B. Almutairi and M. Zribi, "On the sliding mode control of a Ball on a Beam system," *Nonlinear Dynamics*, vol. 59, no. 1-2, pp. 222–239, 2010.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 365049, 8 pages
http://dx.doi.org/10.1155/2015/365049

Research Article

Global Stability of a Variation Epidemic Spreading Model on Complex Networks

De-gang Xu, Xi-yang Xu, Chun-hua Yang, and Wei-hua Gui

College of Information Science and Engineering, Central South University, Hunan, Changsha 410083, China

Correspondence should be addressed to De-gang Xu; dgxu@csu.edu.cn

Received 20 September 2015; Accepted 18 November 2015

Academic Editor: Xinguang Zhang

Epidemic spreading on networks becomes a hot issue of nonlinear systems, which has attracted many researchers' attention in recent years. A novel epidemic spreading model with variant factors in complex networks is proposed and investigated in this paper. One main feature of this model is that virus variation is investigated in the process of epidemic dynamical spreading. The global dynamics of this model involving an endemic equilibrium and a disease-free equilibrium are, respectively, discussed. Some sufficient conditions are given for the existence of the endemic equilibrium. In addition, the global asymptotic stability problems of the disease-free equilibrium and the endemic equilibrium are also investigated by the Routh-Hurwitz stability criterion and Lyapunov stability criterion. And the uniform persistence condition of the new system is studied. Finally, numerical simulations are provided to illustrate obtained theoretical results.

1. Introduction

The research of infectious disease has always been a hot issue of nonlinear systems with applications. The popular dynamics on complex network is the epidemic spreading, which describes how infections spread throughout a network [1]. In recent years, much research work has been done about the viral dynamics of epidemic spreading [2, 3]. These results are helpful for preventing and controlling most emerging infectious diseases like SARS, HIV/AIDS, H5N1, and H1N1. They are also meaningful to provide important information for the research in the field of rumor spreading [4–7], traffic dynamics [8–10], computer viruses [11, 12], biology mechanism [13], and medicine developing [14–16].

In real world, the population size is large enough such that the mixing of individuals can be considered to be homogeneous. Social and biological systems can be properly described as complex networks with nodes representing individuals and links mimicking the interactions [17, 18]. Suitable mathematical models of the infectious disease spreading in complex homogeneous networks are of great practical value to analyze the detailed spreading process. Because epidemic spreading usually brings great harm to society, it is very urgent to establish accurate propagation models considering the infection contagion spreading problems. In past decades, complicated *SIR* models were formulated from different perspectives of epidemiology [19, 20]. In these models, S, I, and R denote, respectively, the number of individuals susceptible to the disease, the number of infectious individuals, and the number of individuals who are recovered from being infectious. The process of epidemic spreading can be further modeled with differential equations, such as *SIS*, *SIR*, and *SIRS* model [21–26]. In these research works the network topological structure is simplified presumptively to regular network or sufficient mixing homogeneous network, where the relationship between the network structure and the epidemic spreading is discussed. Kephart et al. [27] established a virus spreading model based on a homogeneous network by characterizing the average degree as the network metrics and obtained a virus spreading threshold $\tau_c = 1/\langle k\rangle$, while Pastor-Satorras [28–31] studied the epidemic outbreaks in complex heterogeneous network, which chose the degree and average degree as the network metrics and obtained the epidemic spreading threshold $\tau_c = \langle k\rangle/\langle k^2\rangle$. Moore and Newman [32] applied the percolation theory to analyze the epidemic spreading behaviors in small-world network and

showed the differences of spreading action between small-world network and regular network. These research results illustrate that the different network topological structure can affect the epidemic spreading.

Apart from the network topological structure, one of the most important characteristics of epidemic spreading models is the dynamical stability which can reflect the development of the spreading behaviors of infectious disease. Hence, the stability problems of these epidemic models need to be investigated. Kuniya [21] applied a discretization method to prove the global asymptotic stability of the *SIR* model with the age structure. Zhang and Feng [22] deal with the global analysis of a dynamical model describing the spread of tuberculosis with isolation and incomplete treatment. Lahrouz et al. [23] studied a nonlinear *SIRS* model with saturated birth and death rates, and the global asymptotic stability of the model is also discussed. Xu et al. [24] analyzed a time-delayed *SIRS* model with temporary immunity, and some conditions for the globally asymptotically stability of the disease-free equilibrium and the endemic equilibrium are given. Besides, for the epidemic model with time-delay Kang and Fu [25] presented a new *SIS* model with an infective vector on scale-free networks and the global stability of equilibrium is proved. The influences of treatment and vaccination efforts on a dynamic disease model in presence of incubation delays and relapse are studied and sufficient conditions for the local stability of the equilibrium are derived [26].

However, few papers are available in the literature to consider variant factors in the epidemic spreading from a systematic framework. In real world, certain variants exist in the infectious disease transmitting, resulting from some factors including gene mutation and cell division environment. Viruses evolve rapidly because they have strong ability of propensity for genetic variation and short generation time, which leads to evading human immunization response and obtaining drug resistance. For example, influenza viruses can be classified into three major types (A, B, and C). There are many different virus forms because of mutation; type A infects many animal species including humans, while type B and type C viruses are mainly human pathogens. If individuals are affected by viruses, not all infected individuals can be recovered. Some of them may suffer from other diseases because virus variation or the infectious individuals contacted with variants. In fact, some infectious persons, who may be infected by some diseases, would have certain probability to become variant members of another group. It is necessary to propose a new model considering this condition. How to build models with variant factors in the epidemics spreading becomes a challenge. Therefore, the paper presents a novel *SIVRS* epidemic spreading model considering variant factors, where S stands for the susceptible and I, V, and R stand for the infectious, the variant, and the recovered, respectively.

Given the mechanism of the *SIVRS* model in a homogeneous network, which is only composed by blank nodes initially, the entire population can be divided into four groups described by the symbols of S, I, V, and R, respectively. They denote four epidemiological statuses: susceptible, infectious, variant, and recovered. All new individuals are supposed to be blank nodes in complex networks. When a susceptible individual contacts the other infected individual, this individual may be infectious with certain probability. Then an infectious individual would have only three states including infectious, variant, and recovered. The infectious individuals would become the variants with certain probability affected by some factors such as gene mutation and the indeterminacy of cell division. Similarly, an infectious person may become a variant with certain probability after contacting with a variant. Usually, human body can be protected by one's immune system. Some infectious individuals with recovery probability may become the recovered, while others will keep the infectious status. We assume that the four groups have the same mortality rate.

In this paper, a novel *SIVRS* epidemic spreading model with virus variation in complex homogeneous network is proposed and investigated. The rest of this paper is organized as follows. In Section 2, the propagation mechanism of the *SIVRS* model in complex networks is presented, and mean-field equations are used to describe the dynamics of epidemic spreading model with virus variation. The existence of endemic equilibrium is considered. Section 3 is devoted to discuss the global stability of the disease-free equilibrium, which is followed by the discussion of the system uniform persistence in Section 4. Then the proofs of global stability of an endemic equilibrium are presented in Section 5. In Section 6, numerical simulations are performed to illustrate obtained theoretical results. Finally, conclusions are given in Section 7.

2. Epidemic Spreading Model and Its Property

2.1. The SIVRS Model. As described above, in the paper a new *SIVRS* model is established. The model involves a new variant group which is caused by the infectious variation. Assume the number of nodes is N in a closed complex network, which includes four statuses susceptible S, infectious I, variant V, and recovered V as well as some initial blank nodes. All new nodes produced from blank nodes are susceptible.

The flow chart of epidemic spreading is shown in Figure 1.

Assume in a homogeneous network only composed by blank nodes initially the susceptible individuals are produced by the blank nodes; the probability is characterized as δ. Others come from the recovered group with the probability ϕ. A susceptible individual will become infectious with probability α if he/she contacts the infected individual. Then an infectious individual may perhaps become one variant when he/she has tight relation with variants or is affected by other factors. We assume that the variant probability is γ when contacting with a variant. In the process of epidemic spreading, an infectious individual may become the variant with internal probability η. Some infectious individuals with recovery probability β may recover and others will keep the infectious status. In this paper, the four groups are supposed to have the same mortality rate μ. Then the *SIVRS* epidemic spreading rules can be summarized as follows.

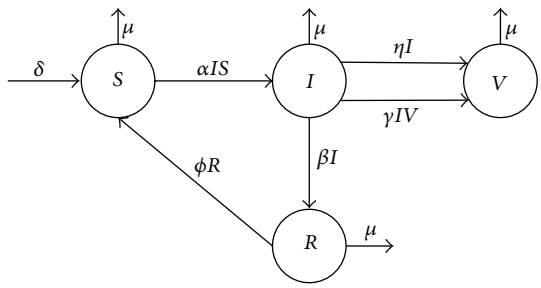


Figure 1: A schematic representation of *SIVRS* epidemic spreading model.

(1) Apart from the four groups in a closed network, there are blank nodes which exist in initial network. The blank nodes may become susceptible ones with probability δ, namely, crude birth rate.

(2) The susceptible individual becomes infectious with probability α when contacting with an infectious one, namely, infection rate.

(3) An infectious individual can be recovered with probability β, namely, recovery rate.

(4) The variants coming from some of the infectious nodes at a variation rate η (internal variation rate) can reflect the variation factors. When an infectious individual contacts with a variant, the individual may become a variant with probability γ (contact variant rate).

(5) The recovered node turns into susceptible with probability ϕ after a period of time due to the loss of immunity. For the four groups in the network, all individuals will become blank with probability μ, namely, natural mortality rate.

A closed and homogeneous network consisting of N individuals is investigated in this paper. Individuals in the network can be represented with nodes and the contact between different individuals can be denoted by edges. Then the network can be described by an undirected graph $G = (V, E)$, where V and E denote the set of nodes and edges, respectively. Therefore, a differential equation model is derived based on the aforementioned rules and the basic assumptions:

$$\begin{aligned} \frac{dS}{dt} &= \delta(1-N) - \alpha\langle k\rangle SI + \phi R - \mu S, \\ \frac{dI}{dt} &= \alpha\langle k\rangle SI - \gamma\langle k\rangle IV - (\eta+\beta+\mu) I, \\ \frac{dV}{dt} &= \gamma\langle k\rangle IV + \eta I - \mu V, \\ \frac{dR}{dt} &= \beta I - \phi R - \mu R, \end{aligned} \tag{1}$$

where $\langle k\rangle$ denotes the average degree of the network.

The total population satisfies $N = S + I + V + R$, and the following equation is obtained:

$$\frac{dN}{dt} = \delta - (\delta+\mu) N \tag{2}$$

which is derived by adding the four equations in (1).

In (2) N will eventually tend to $N_0 = \delta/(\delta+\mu)$ with the exponential decay. Therefore, assume that $N(0) = N_0$. The closed and positively invariant set for (1) is $\Sigma = \{(S, I, V, R) \in \mathbb{R}_+^4 : 0 \le S + I + V + R = N_0 \le 1\}$, where $\mathbb{R}_+^4$ denotes the nonnegative cone of $\mathbb{R}^4$ with its lower dimensional faces. Use $\partial\Sigma$ and Σ° to denote the boundary and interior of Σ in $\mathbb{R}_+^4$, respectively.

2.2. Existence of Equilibrium. The system (1) has a disease-free equilibrium (DFE) E_0, where

$$E_0 = (S_0, I_0, V_0, R_0) = \left(\frac{\delta}{\delta+\mu}, 0, 0, 0\right). \tag{3}$$

Denote the basic reproduction number parameter as

$$R_0 = \frac{\alpha\langle k\rangle S_0 - \gamma\langle k\rangle V_0}{\eta+\beta+\mu} = \frac{\alpha\langle k\rangle\delta}{(\delta+\mu)(\eta+\beta+\mu)}. \tag{4}$$

The following theorem summarizes the parameter restrictions on the existence of equilibrium.

Theorem 1. *If $R_0 > 1$ and the inequality*

$$1 > \eta > \frac{\alpha(1-\gamma\langle k\rangle)}{\alpha+\gamma} \tag{5}$$

is satisfied, there are two endemic equilibria for system (1).

Proof. Assume that $E^* = (S^*, I^*, V^*, R^*)$ is an endemic equilibrium (EE) of system (1). According to system (1), we have

$$\begin{aligned} \delta(1-N^*) - \alpha\langle k\rangle S^* I^* + \phi R^* - \mu S^* &= 0, \\ \alpha\langle k\rangle S^* I^* - \gamma\langle k\rangle I^* V^* - (\eta+\beta+\mu) I^* &= 0, \\ \gamma\langle k\rangle I^* V^* + \eta I^* - \mu V^* &= 0, \\ \beta I^* - \phi R^* - \mu R^* &= 0. \end{aligned} \tag{6}$$

For (6) a straightforward calculation leads to

$$\begin{aligned} S^* &= \frac{\eta+\beta+\mu}{\alpha\langle k\rangle} + \frac{\gamma\eta I^*}{\alpha(\mu-\gamma\langle k\rangle I^*)}, \\ V^* &= \frac{\eta I^*}{\mu-\gamma\langle k\rangle I^*}, \\ R^* &= \frac{\beta}{\mu+\phi} I^*. \end{aligned} \tag{7}$$

From (7) $\mu - \gamma\langle k\rangle I^* > 0$, which implies that

$$0 < I^* < \min\left\{1, \frac{\mu}{(\gamma\langle k\rangle)}\right\}, \tag{8}$$

the component I^* is a positive solution of

$$p(I^*) = AI^{*2} + BI^* + C = 0, \tag{9}$$

where

$$\begin{aligned} A &= \alpha\gamma \langle k\rangle^2 (\mu+\phi+\beta), \\ B &= -\eta \langle k\rangle (\mu+\phi)(\alpha+\gamma) - \alpha \langle k\rangle \mu(\mu+\phi+\beta) \\ &\quad + (\mu+\phi)(\eta+\beta+\mu)(1-\gamma\langle k\rangle R_0), \\ C &= \mu(\eta+\beta+\mu)(\mu+\phi)(R_0-1). \end{aligned} \tag{10}$$

Therefore, consider the following:

(1) If $R_0 < 1$, we have $C < 0$; (9) has only one positive solution.

(2) If $R_0 = 1$, we have $C = 0$; (9) has only one positive solution $-B/A$.

(3) If $R_0 > 1$, we have $C > 0$:

(i) if $B > 0$, the positive solution of (9) does not exist;

(ii) if $B < 0$ and $B > -2\sqrt{AC}$, the solution of (9) does not exist;

(iii) if $B < 0$ and $B = -2\sqrt{AC}$, (9) has only one positive solution;

(iv) if $B < 0$ and $B < -2\sqrt{AC}$, (9) has two positive solutions.

Therefore, if $R_0 > 1$, the solution of (9) exists only when $B \le -2\sqrt{AC}$.

According to the inequality $a + b \ge 2\sqrt{ab}$, $\forall a, b \in Z^+$, assume $R_0 > 1$, choose $a = \alpha\langle k\rangle\mu(\mu+\phi+\beta)$, $b = \gamma\langle k\rangle(\eta+\beta+\mu)(\mu+\phi)(R_0-1)$, and then

$$\begin{aligned} &2\sqrt{AC} \\ &= 2\sqrt{\alpha\gamma\langle k\rangle^2 \mu(\mu+\phi+\beta)(\eta+\beta+\mu)(\mu+\phi)(R_0-1)} \\ &< \alpha\langle k\rangle \mu(\mu+\phi+\beta) \\ &\quad + \gamma\langle k\rangle(\eta+\beta+\mu)(\mu+\phi)(R_0-1) \\ &= \alpha\langle k\rangle \mu(\mu+\phi+\beta) + \gamma\langle k\rangle R_0(\eta+\beta+\mu)(\mu+\phi) \\ &\quad - \gamma\langle k\rangle(\eta+\beta+\mu)(\mu+\phi). \end{aligned} \tag{11}$$

If the inequality $\gamma\langle k\rangle(\eta+\beta+\mu)(\mu+\phi) > (\mu+\phi)[\eta+\beta+\mu-\eta\langle k\rangle(\alpha+\gamma)]$ is satisfied, we have

$$\eta+\beta+\mu < \frac{\eta\langle k\rangle(\alpha+\gamma)}{1-\gamma\langle k\rangle}. \tag{12}$$

Then the following equation is obtained:

$$\begin{aligned} 2\sqrt{AC} &< \alpha\langle k\rangle \mu(\mu+\phi+\beta) \\ &\quad + \gamma\langle k\rangle R_0(\eta+\beta+\mu)(\mu+\phi) \\ &\quad - \gamma\langle k\rangle(\eta+\beta+\mu)(\mu+\phi) \\ &< \alpha\langle k\rangle \mu(\mu+\phi+\beta) \\ &\quad + \gamma\langle k\rangle R_0(\eta+\beta+\mu)(\mu+\phi) \\ &\quad - [\eta+\beta+\mu-\eta\langle k\rangle(\alpha+\gamma)](\mu+\phi) \\ &< \alpha\langle k\rangle \mu(\mu+\phi+\beta) \\ &\quad + (\eta+\beta+\mu)(\mu+\phi)(\gamma\langle k\rangle R_0 - 1) \\ &\quad + \eta\langle k\rangle(\mu+\phi)(\alpha+\gamma) < -B \end{aligned} \tag{13}$$

which implies system (1) has two endemic equilibria.

When $R_0 > 1$,

$$\frac{\alpha\langle k\rangle}{\eta+\beta+\mu} > R_0 = \frac{\alpha\langle k\rangle\delta}{(\delta+\mu)(\eta+\beta+\mu)} > 1 \tag{14}$$

which can transfer to inequality $\alpha\langle k\rangle > \eta+\beta+\mu$.

If $\alpha < \eta(\alpha+\gamma)/(1-\gamma\langle k\rangle)$, we have $1 > \eta > \alpha(1-\gamma\langle k\rangle)/(\alpha+\gamma)$. Then inequality (12) is satisfied. □

Remark 2. According to this theorem, if the reproduction number parameter is above the threshold, then the endemic equilibrium is globally asymptotically stable, which will be discussed further in Section 5.

3. Global Stability of the Disease-Free Equilibrium

Definition 3. If the equilibrium is stable under the meaning of Lyapunov, for $\delta(\varepsilon, t_0)$ and $\forall\mu > 0$, there is real number $T(\mu, \delta, t_0) > 0$ which makes any initial value x_0 of inequality $\|x_0 - x_e\| \le \delta(\varepsilon, t_0)$, $t \ge t_0$, satisfy the following inequality:

$$\|\phi(t; x_0, t_0 - x_e)\| \le \mu, \quad \forall t \ge t_0 + T(\mu, \delta, t_0); \tag{15}$$

then the equilibrium is asymptotically stable.

The Jacobian matrix at the disease-free equilibrium E_0 of system (1) is

$$J(E_0) = \begin{pmatrix} -\delta-\mu & -\delta-\alpha\langle k\rangle S_0 & -\delta & -\delta+\phi \\ 0 & \alpha\langle k\rangle S_0 - (\eta+\beta+\mu) & 0 & 0 \\ 0 & \eta & -\mu & 0 \\ 0 & \beta & 0 & -\mu-\phi \end{pmatrix}. \tag{16}$$

Obviously, if $R_0 < 1$, all eigenvalues of matrix (16) are negative. Then the disease-free equilibrium E_0 is locally asymptotically stable in Σ. Moreover, if $R_0 > 1$, there is one positive eigenvalue and E_0 is unstable.

Theorem 4. *If $R_0 < 1$, the disease-free equilibrium (DFE) E_0 is globally asymptotically stable in Σ and if $R_0 > 1$, the disease-free equilibrium (DFE) E_0 is unstable in Σ.*

Proof. Let $L(S, I, V, R) = I > 0$ as a Lyapunov function; then $L(E_0) = 0$. When $R_0 < 1$

$$\begin{aligned}\frac{dL}{dt}(S, I, V, R) &= \alpha \langle k \rangle SI - \gamma \langle k \rangle IV - (\eta + \beta + \mu) I \\ &< \alpha \langle k \rangle SI - (\eta + \beta + \mu) I \\ &< I(\alpha \langle k \rangle N_0 - (\eta + \beta + \mu)) \\ &< I(\eta + \beta + \mu)(R_0 - 1) < 0.\end{aligned} \tag{17}$$

L is positive definite and $\dot{L}$ is negative definite. Therefore, the disease-free equilibrium (DFE) E_0 is globally asymptotically stable in Σ; the following result can be given. □

4. Uniform Persistence

In this section, the uniform persistence of system (1) will be discussed when the basic reproduction number $R_0 > 1$.

Definition 5 (see [33]). System (1) is said to be uniformly persistent if there exists a constant $0 < c < 1$, which makes any solution $(S(t), I(t), V(t), R(t))$ with $(S(0), I(0), V(0), R(0)) \in \Sigma^\circ$ satisfy

$$\min\left\{\lim_{t\to\infty}\lim S(t), \lim_{t\to\infty} I(t), \lim_{t\to\infty} V(t), \lim_{t\to\infty} R(t)\right\} \geq c. \tag{18}$$

Let X be a locally compact metric space with metric ∂ and let Γ be a closed nonempty subset of X with boundary $\partial\Gamma$ and interior Γ°. Obviously, $\partial\Gamma$ is a closed subset of Γ. Let Φ_t be a dynamical system defined on Γ. A set B in X is said to be invariant if $\Phi_t(B, t) = B$. Define $M_\partial := \{x \in \partial\Gamma : \Phi_t x \in \partial\Gamma, \ \forall t \geq 0\}$.

Lemma 6 (see [34]). *Assume the following:*

(H1) *Φ_t has a global attractor.*

(H2) *There exists an $M = \{M_1, \ldots, M_k\}$ of pair-wise disjoint, compact, and isolated invariant set on $\partial\Gamma$ such that*

(a) $\bigcup_{x\in M_\partial} \omega(x) \subset \bigcup_{j=1}^{k} M_j$;

(b) *no subsets of M form a cycle on $\partial\Gamma$;*

(c) *each M_j is also isolated in Γ;*

(d) *$W^s(M_j) \cap \Gamma^\circ = \phi$ for each $1 \leq j \leq k$, where $W^s(M_j)$ is the stable manifold of M_j. Then Φ_t is uniformly persistent with respect to Γ°.*

According to Lemma 6, the following result is obtained.

Theorem 7. *When $R_0 > 1$, system (1) is uniformly persistent.*

Proof. Let

$$\begin{aligned}\Gamma &= \Sigma = \left\{(S, I, V, R) \in \mathbb{R}_4^+ \mid 0 \leq S + I + V + R \leq 1\right\}, \\ \Gamma^\circ &= \{(S, I, V, R) \in E : I, V > 0\}, \\ \partial\Gamma &= \frac{\Gamma}{\Gamma^\circ}.\end{aligned} \tag{19}$$

Obviously, $M_\partial = \partial\Gamma$.

Choose $M = \{E_0\}$, $\omega(x) = \{E_0\}$ for all $x \in M_\partial$. On $\partial\Gamma$, system (1) reduces to $S' = \delta - (\delta+\mu)S$, in which $S(t) \to \delta/(\delta+\mu)$ as $t \to \infty$. It is concluded that $M = \{E_0\}$, $\omega(x) = \{E_0\}$ for all $x \in M_\partial$, which indicates that hypotheses (a) and (b) hold. When $R_0 > 1$, the disease-free equilibrium E_0 is unstable according to Theorem 4 $W^s(M) = \partial\Gamma$. Hypotheses (c) and (d) are then satisfied. Due to the ultimate boundedness of all solutions of system (1), there is a global attractor, making (H1) true. □

5. Global Dynamics of Endemic Equilibrium

From the previous analysis, the disease dies out when $R_0 > 1$; then the disease becomes endemic. In this section, Lyapunov asymptotic stability theorem is used to investigate the globally asymptotic stability of the endemic equilibrium E^* when $R_0 > 1$.

Theorem 8. *The endemic equilibrium E^* is globally asymptotically stable in Σ, whenever $R_0 > 1$.*

Proof. Consider the following function:

$$V_1 = \ln\left[(S - S^*) + (I - I^*) + (V - V^*) + (R - R^*) + 1\right]. \tag{20}$$

Then the derivative of V_1 along the solution of (1) is given by

$$\begin{aligned}\dot{V}_1 &= \frac{\partial V_1}{\partial S}\frac{dS}{dt} + \frac{\partial V_1}{\partial I}\frac{dI}{dt} + \frac{\partial V_1}{\partial V}\frac{dV}{dt} + \frac{\partial V_1}{\partial R}\frac{dR}{dt} \\ &= \frac{(dS + dI + dV + dR)(1/dt)}{(S - S^*) + (I - I^*) + (V - V^*) + (R - R^*) + 1}.\end{aligned} \tag{21}$$

From (2), all solutions of (6) satisfy the equality

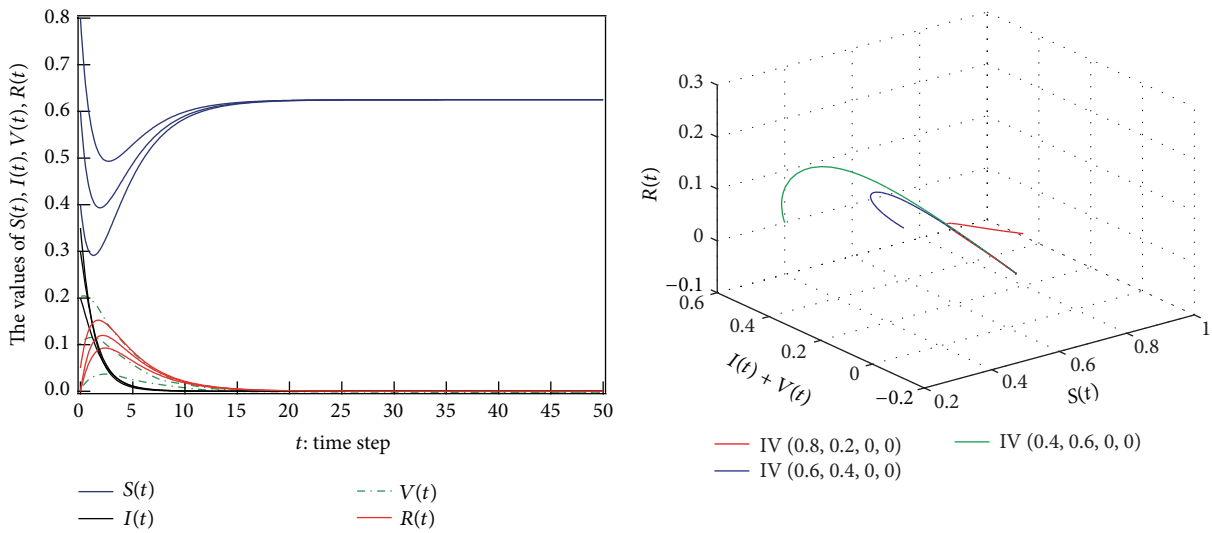
$$N^* = S^* + I^* + V^* + R^* = \frac{\delta}{\delta + \mu} \tag{22}$$

and $N = e^{-(\delta+\mu)t+C} + \delta/(\delta + \mu) \leq \delta/(\delta + \mu)$, where C is the value that makes $N_0 = \delta/(\delta + \mu)$ satisfied.

Hence $V_1 = \ln(N - N^* + 1) \geq 0$; then

$$\begin{aligned}\dot{V}_1 &= \frac{1}{N - \delta/(\delta + \mu) + 1}\frac{dN}{dt} \\ &= \frac{\delta + \mu}{N - \delta/(\delta + \mu) + 1}\left(\frac{\delta}{\delta + \mu} - N\right) \leq 0.\end{aligned} \tag{23}$$

If and only if $N = \delta/(\delta + \mu)$, $V_1 = 0$ and $\dot{V}_1 = 0$ are satisfied.



(a) The states of disease-free equilibrium

(b) The evolution for the disease-free system states

FIGURE 2: (a) and (b) showed that the disease-free equilibrium E_0 of the system (1) is globally asymptotically stable with different initial conditions $(0.8, 0.2, 0, 0)$, $(0.6, 0.3, 0.1, 0)$, and $(0.4, 0.35, 0.2, 0.05)$ and the parameters $\langle k \rangle = 10$, $\alpha = 0.8$, $\beta = 0.5$, $\eta = 0.2$, $\delta = 0.5$, $\gamma = 0.2$, $\phi = 0.01$, and $\mu = 0.3$; $R_0 = 0.5 < 1$. The value of DFE is $E_0 = (0.625, 0, 0, 0)$.

V_1 is positive definite and $\dot{V}_1$ is negative definite. Therefore, the function V_1 is a Lyapunov function for system (1) and the endemic equilibrium E^* is globally asymptotically stable by Lyapunov asymptotic stability theorem [35]. The proof is completed. □

6. Numerical Simulation

To demonstrate the theoretical results obtained in this paper, some numerical simulations will be discussed. In this paper, the hypothetical set of initial values (IV) and parameter values will be given as follows.

Consider the initial values of $(S(0), I(0), V(0), R(0))$ are set as $(0.8, 0.2, 0, 0)$, $(0.6, 0.3, 0.1, 0)$, and $(0.4, 0.35, 0.2, 0.05)$, respectively.

(1) The disease-free equilibrium: Set $\langle k \rangle = 10$, $\alpha = 0.08$, $\beta = 0.5$, $\eta = 0.2$, $\delta = 0.5$, $\gamma = 0.02$, $\phi = 0.01$, and $\mu = 0.3$. $R_0 = 0.5 < 1$ and the disease-free equilibrium $E_0 = (0.625, 0, 0, 0)$ from the parameter values above through the calculation. According to Theorem 4, the disease-free equilibrium E_0 of system (1) is globally asymptotically stable in Σ in this case. The simulation results are shown in Figures 2(a) and 2(b).

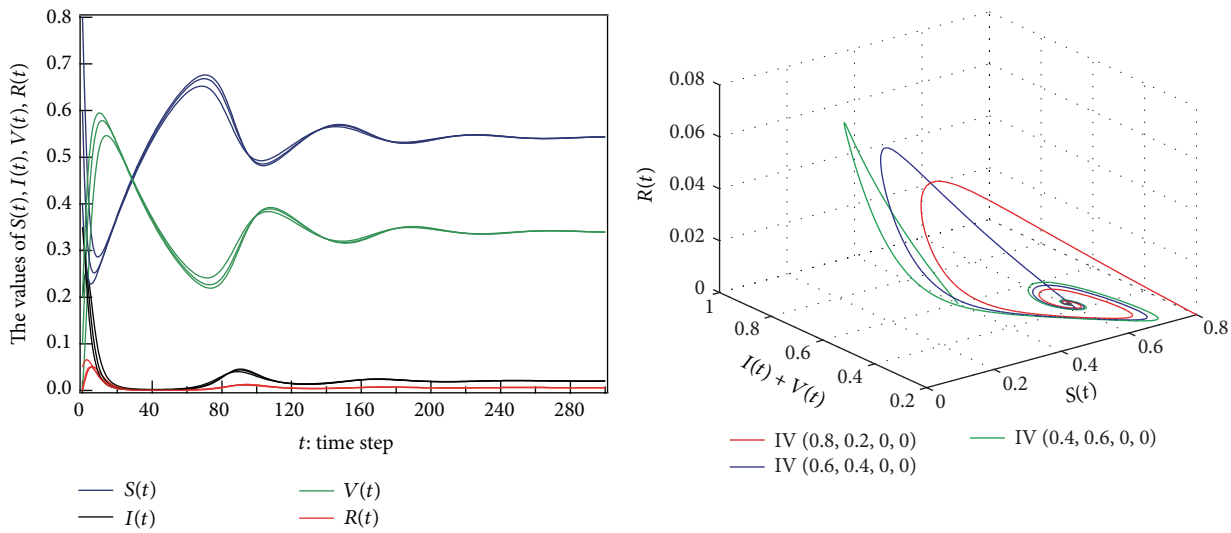
(2) The endemic equilibrium: Set $\langle k \rangle = 10$, $\alpha = 0.08$, $\beta = 0.08$, $\eta = 0.3$, $\delta = 0.2$, $\gamma = 0.01$, $\phi = 0.25$, and $\mu = 0.02$. By direct computation, $R_0 = 1.818 > 1$ and the endemic equilibrium $E^* = (0.5435, 0.0201, 0.3395, 0.00596)$ can be obtained from the parameter values above. According to Theorem 4, the positive endemic equilibrium E^* of system (1) is globally asymptotically stable in Σ°. The simulation results are shown in Figures 3(a) and 3(b).

Figure 2 shows that if $R_0 < 1$, all solutions in Σ would be attracted to the disease-free equilibrium E_0 regardless of the initial values of system (1), which illustrates the validity of Theorem 4. Similarly, it can be seen from Figure 3 that all solutions in Σ° would be attracted to the endemic equilibrium E^* regardless of the initial values of system (1) if $R_0 > 1$ and the conditions of Theorem 8 are satisfied, which is obviously the content of Theorem 4. Moreover, the relationship between the values of equilibrium can be verified as shown in (24), which is coincident with the theoretical results:

$$S_0 + I_0 + V_0 + R_0 = S^* + I^* + V^* + R^* = -\frac{\delta}{\delta + \mu}. \quad (24)$$

7. Conclusion

The stability of the *SIVRS* epidemic spreading model with virus variation in complex networks has been discussed in this paper. The model involves a new variant group which is caused by the infectious variation. By analyzing the model, the disease-free equilibrium E_0 is proved to exist when the basic reproduction number R_0 is less than 1. The analysis result reveals that the infectious disease dies out when R_0 is more than 1 and it becomes endemic. The existing conditions of endemic equilibrium related with the variation rate and the network nodes degree are obtained. Besides, the global asymptotically stability condition of the disease-free equilibrium is obtained by the Routh-Hurwitz stability criterion and the Lyapunov stability criterion. And the condition of the system uniform persistence is also given. The proof of the stability of endemic equilibrium is also illustrated. Finally, a numerical simulation is given to illustrate the correctness of the disease-free equilibrium and the endemic equilibrium results.



(a) The system states of endemic equilibrium

(b) The evolution for the endemic system states

FIGURE 3: (a) and (b) showed that the disease-free equilibrium E^* of system (1) is globally asymptotically stable with different initial conditions $(0.8, 0.2, 0, 0)$, $(0.6, 0.3, 0.1, 0)$, and $(0.4, 0.35, 0.2, 0.05)$ and the parameter values $\alpha = 0.08$, $\beta = 0.08$, $\eta = 0.3$, $\delta = 0.2$, $\gamma = 0.1$, $\phi = 0.25$, $\mu = 0.02$, and $\langle k \rangle = 10$; $R_0 = 1.818 > 1$. The value of EE is $E^* = (0.5435, 0.0201, 0.3395, 0.00596)$.

Conflict of Interests

The authors have declared that no competing interests exist.

Acknowledgments

This research was partially supported by the National Natural Science Foundation of China (Grant nos. 61473319, 61104135), the Science Fund for Creative Research Groups of National Natural Science of China (Grant 61321003), and the Science Fund for Creative Research Project (2015CX007).

References

[1] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of Modern Physics*, vol. 74, no. 1, pp. 47–97, 2002.

[2] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.

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

[4] L. Zhao, W. Xie, H. O. Gao, X. Qiu, X. Wang, and S. Zhang, "A rumor spreading model with variable forgetting rate," *Physica A*, vol. 392, no. 23, pp. 6146–6154, 2013.

[5] F. Chierichetti, S. Lattanzi, and A. Panconesi, "Rumor spreading in social networks," *Theoretical Computer Science*, vol. 412, no. 24, pp. 2602–2610, 2011.

[6] P. Manshour and A. Montakhab, "Contagion spreading on complex networks with local deterministic dynamics," *Communications in Nonlinear Science and Numerical Simulation*, vol. 19, no. 7, pp. 2414–2422, 2014.

[7] C. Liu and Z.-K. Zhang, "Information spreading on dynamic social networks," *Communications in Nonlinear Science and Numerical Simulation*, vol. 19, no. 4, pp. 896–904, 2014.

[8] K. Qiao, P. Zhao, and X.-M. Yao, "Performance analysis of urban rail transit network," *Journal of Transportation Systems Engineering and Information Technology*, vol. 12, no. 4, pp. 115–121, 2012.

[9] O. Lordan, J. M. Sallan, and P. Simo, "Study of the topology and robustness of airline route networks from the complex network approach: a survey and research agenda," *Journal of Transport Geography*, vol. 37, pp. 112–120, 2014.

[10] X. Ling, M.-B. Hu, W.-B. Du, R. Jiang, Y.-H. Wu, and Q.-S. Wu, "Bandwidth allocation strategy for traffic systems of scale-free network," *Physics Letters A*, vol. 374, no. 48, pp. 4825–4830, 2010.

[11] L.-X. Yang and X.-F. Yang, "The spread of computer viruses over a reduced scale-free network," *Physica A: Statistical Mechanics and Its Applications*, vol. 396, pp. 173–184, 2014.

[12] C. Gan, X. Yang, W. Liu, Q. Zhu, and X. Zhang, "An epidemic model of computer viruses with vaccination and generalized nonlinear incidence rate," *Applied Mathematics and Computation*, vol. 222, pp. 265–274, 2013.

[13] R. Steuer, "Computational approaches to the topology, stability and dynamics of metabolic networks," *Phytochemistry*, vol. 68, no. 16–18, pp. 2139–2151, 2007.

[14] A. Li, J. Li, and Y. Pan, "Discovering natural communities in networks," *Physica A*, vol. 436, no. 15, pp. 878–896, 2015.

[15] Y. Maeno, "Discovering network behind infectious disease outbreak," *Physica A*, vol. 389, no. 21, pp. 4755–4768, 2010.

[16] R. Franke and G. Ivanova, "FALCON or how to compute measures time efficiently on dynamically evolving dense complex networks," *Journal of Biomedical Informatics*, vol. 47, no. 2, pp. 62–70, 2014.

[17] J.-H. Li, X.-Y. Shao, Y.-M. Long, H.-P. Zhu, and B. R. Schlessman, "Global optimization by small-world optimization algorithm based on social relationship network," *Journal of Central South University*, vol. 19, no. 8, pp. 2247–2265, 2012.

[18] X.-H. Yang, B. Wang, and B. Sun, "A novel weighted evolving network model based on clique overlapping growth," *Journal of*

Central South University of Technology, vol. 17, no. 4, pp. 830–835, 2010.

[19] O. Diekmann and J. A. P. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*, vol. 251, John Wiley & Sons, Chichester, UK, 2000.

[20] H. R. Thieme, *Mathematics in Population Biology*, vol. 4, Princeton University Press, 2003.

[21] T. Kuniya, "Global stability analysis with a discretization approach for an age-structured multigroup SIR epidemic model," *Nonlinear Analysis: Real World Applications*, vol. 12, no. 5, pp. 2640–2655, 2011.

[22] J. Zhang and G. Feng, "Global stability for a tuberculosis model with isolation and incomplete treatment," *Computational and Applied Mathematics*, vol. 34, no. 3, pp. 1237–1249, 2015.

[23] A. Lahrouz, L. Omari, and D. Kiouach, "Global analysis of a deterministic and stochastic nonlinear SIRS epidemic model," *Nonlinear Analysis: Modelling and Control*, vol. 16, no. 1, pp. 59–76, 2011.

[24] R. Xu, Z. Ma, and Z. Wang, "Global stability of a delayed SIRS epidemic model with saturation incidence and temporary immunity," *Computers & Mathematics with Applications*, vol. 59, no. 9, pp. 3211–3221, 2010.

[25] H. Kang and X. Fu, "Epidemic spreading and global stability of an SIS model with an infective vector on complex networks," *Communications in Nonlinear Science & Numerical Simulation*, vol. 27, no. 1–3, pp. 30–39, 2015.

[26] P. Raja Sekhara Rao and M. Naresh Kumar, "A dynamic model for infectious diseases: the role of vaccination and treatment," *Chaos, Solitons & Fractals*, vol. 75, pp. 34–49, 2015.

[27] J. O. Kephart, S. R. White, and D. M. Chess, "Computers and epidemiology," *IEEE Spectrum*, vol. 30, no. 5, pp. 20–26, 1993.

[28] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, "Epidemic outbreaks in complex heterogeneous networks," *European Physical Journal B*, vol. 26, no. 4, pp. 521–529, 2002.

[29] R. Pastor-Satorras and A. Vespignani, "Epidemic dynamics and endemic states in complex networks," *Physical Review E*, vol. 63, Article ID 066117, 2001.

[30] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," *Physical Review Letters*, vol. 86, no. 14, pp. 3200–3203, 2001.

[31] R. Pastor-Satorras and A. Vespignani, "Epidemic dynamics in finite size scale-free networks," *Physical Review E*, vol. 65, no. 3, Article ID 035108, 2002.

[32] C. Moore and M. E. J. Newman, "Epidemics and percolation in small-world networks," *Physical Review E*, vol. 61, no. 5, pp. 5678–5682, 2000.

[33] G. Butler, H. I. Freedman, and P. Waltman, "Uniformly persistent systems," *Proceedings of the American Mathematical Society*, vol. 96, no. 3, pp. 425–430, 1986.

[34] X.-Q. Zhao, *Dynamical Systems in Population Biology*, Canadian Mathematical Society, Springer, 2003.

[35] A. M. Lyapunov, *The General Problem of the Stability of Motion*, Taylor & Francis, London, UK, 1992.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 589184, 10 pages
http://dx.doi.org/10.1155/2015/589184

Research Article

Optimal Controller and Controller Based on Differential Flatness in a Linear Guide System: A Performance Comparison of Indexes

Fabio Abel Gómez Becerra,[1,2] Víctor Hugo Olivares Peregrino,[1] Andrés Blanco Ortega,[1] and Jesús Linares Flores[3]

[1] *Centro Nacional de Investigación y Desarrollo Tecnológico, 62490 Cuernavaca, MOR, Mexico*
[2] *ITS de PV, 48333 Puerto Vallarta, JAL, Mexico*
[3] *Universidad Tecnológica de la Mixteca, 69000 Huajuapam de León, OAX, Mexico*

Correspondence should be addressed to Fabio Abel Gómez Becerra; fabioabelgo@hotmail.com

Received 27 August 2015; Revised 20 November 2015; Accepted 23 November 2015

Academic Editor: Wenguang Yu

The use of linear slide system has been augmented in recent times due to features granted to supplement electromechanical systems; new technologies have allowed the manufacture of these systems with low coefficients of friction and offer a variety of types of sliding. In this paper, we present a comparison between the performance indexes of two techniques of control applying optimal control LQR (Linear Quadratic Regulator) acronym for STIs in English and the technique of differential flatness controller. The use of linear slide bolt of potency takes into account the dynamics of the DC motor; the Euler-Lagrange formalism was used to establish the mathematical model of the slide. Cosimulation via the MATLAB/Simulink-ADAMS virtual prototype package, including realistic measurement disturbances, is used to compare the performance indexes between the LQR controller versus differential flatness controller for the position tracking of linear guide system.

1. Introduction

The linear slides have been used in different electromechanical systems as actuators; these have great advantages, especially those consisting of screws-and-nuts; some advantages are as follows: the effect of gravity at the beginning of the movement introduces no disturbance of change of position by not overcoming the power screw, the degree of accuracy in positioning is very high, it is possible to track both soft paths laws of classic and modern control, and the force developed by using this type of device is very high usually requiring actuators (DC motor) of lower power than those requiring other types of systems. Blanco Ortega et al. [1] proposed a rehabilitation ankle device based on an XY table which uses two linear slidings, one on each axis, through a Generalized Proportional Integral (GPI) controller. Valdivia et al. [2] proposed a rehabilitation ankle TobiBot, which covers only the movements of dorsiflexion/plantarflexion, and a degree of freedom; it is controlled by an outline of PID control, to perform the movements, and uses a linear slide based on a power screw. Blanco Ortega et al. [3, 4] in turn presented a rehabilitation of the design and construction of an ankle rehabilitation based on a parallel robot of 3 degrees of freedom, which provides the movements of dorsiflexion/plantarflexion and inversion/eversion made by the ankle and uses a PID control technique to perform rehabilitation ankle movements by using three linear power sliding screws. Blanco Ortega et al. [3, 4] have presented an ankle rehabilitation machine, using a linear slide to effect movements dorsiflexion/ankle plantarflexion; the control technique using a PID was smooth trajectory tracking.

As can be seen, linear slides have been widely used as actuators in various robotic devices from a simple XY table. Various control techniques have been implemented for

accurate position with smooth tracking paths. The present work proposes control technique for differential flatness, compared with optimal control, showing both its advantages and disadvantages over the other, considering the control technique for differential flatness as an option with highly acceptable results in tracking soft paths, and highlighting also the lack of optimization for optimum control trajectories mechanical tracking systems of this type; the randomness of the matrix appears when weighing matrixes R and Q. Control technique by feedback employed in this work is based on the concept of differential flatness, which come from differential plan systems (Fliess et al. [5]). This was made known fifteen years ago in France by professor Fliess and his collaborators. Differential flatness has had important uses within the areas of robotics, control processes, aerospace systems, optimization systems, trajectory planning in linear and nonlinear aspects, and systems of infinite dimensions described in partially controlled differential equations with border conditions (Fliess et al. [5, 6]; Linares-Flores and Sira-Ramírez [7]; and Sira-Ramírez and Agrawal [8]). Thounthong et al. [9] suggest the use of a differential based on flatness of a fuel cell system and a hybrid supercapacitors source achieving robustness, stability, and efficiency of the controlled system controller. Jörgl and Gattringer [10] proposed the control of a conveyor belt using a control law based on differential flatness reducing the trajectory tracking error compared with traditional drivers, flatness theory has been used in a variety of nonlinear systems in various engineering disciplines, Thounthong and Pierfederici [11], such as the inverted pendulum control and aircraft vertical rise and fall, Fliess et al. [12]. Danzer et al. [13] proposed driver in such a control system of the pressure of a cathode and oxygen excess of a chemical system. Gensior et al. [14] used a control of tracking of a DC voltage boost converter. Song et al. [15] have shown that based on flatness control is robust and provides improved performance monitoring transience compared to a traditional method of linear control (PI). A nonlinear system is flat if there is a set of independent variables (differentially equal in number to the number of entries) so that all the state variables x and input variables u can be expressed in terms of those output by Syed et al. [16], Rabbani et al. [17], and Agrawal et al. [18].

In order to make a comparison between the performance index of different controllers, this paper also presents an optimal control law applied to the same linear slide (Figure 1); optimal control theory focuses on the design of controllers to perform their objective and concurrently satisfy physical constraints to optimize predetermined performance criteria (Hassani and Lee [19]). With the new trend of seeking a high-performance, sustainable manufacturing, pollution awareness and finding ways for greater energy efficiency, greater emphasis on the optimal design of control systems is made. The optimality design criteria may include minimum fuel, low energy, minimum time (Lewis [20]). Because of this the focus in recent years has been directed to the use of various techniques of optimal control; Yu and Hwang [21] have presented an LQR (Linear Quadratic Regulator) approach for determining a control law PID optimal in order to control the speed of a DC motor; this contribution proposes a systematic approach to design a speed control of a DC motor based on an identification model and LQR design with a nonlinear increase with feedforward compensator. Ruderman et al. [22] propose a methodical approach LQR state feedback control of a DC motor. Likewise LQR optimal control has been used in conjunction with other control techniques such as that of Prasad et al. [23] who proposed a control system high nonlinearity such as the inverted pendulum. It is presented with a linearized dynamics using a PID controller LQR bringing out results in a robust control scheme for optimal system control.

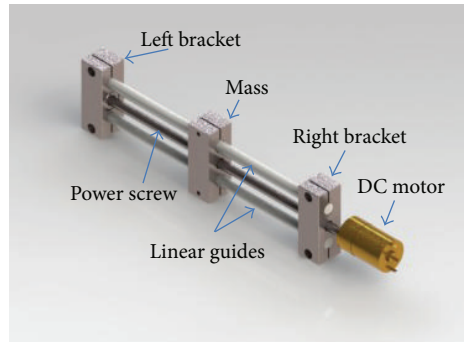


Figure 1: Linear slide system.

The contributions of this paper are as follows: (1) the design of an LQR position tracking controller; (2) the design of a differential flatness position tracking controller; these tracking controllers are compared on the performance index of position tracking error. Thus, we conclude that the two controllers have a high capacity for the position tracking of the linear guide system. This paper is organized as follows. Section 2 presents the mathematical model of linear slide system. In this section, we obtained the mathematical model via Euler-Lagrange formalism and incorporate the dynamic of DC motor to model. The position control based on linear quadratic regulator is presented in Section 3. In Section 4, we present the design of tracking controller based on differential flatness. In Section 5, the simulations results are obtained through the MATLAB/Simulink-ADAMS virtual prototype package, and we compare the performance indexes of the position tracking error of both controllers. In Section 6, the results are shown in the experiment using a physical slider and data acquisition card, using the LabView software with a graphical interface. Finally, in Section 7, we give the conclusions of all of the work.

2. Mathematical Model of Linear Slide

The linear slide control in this work is formed with a motor coupled CD to a power screw through a gearbox speed, screw, and rotating, linearly moving mass m, as shown in Figures 1 and 2.

Motor mathematical model of linear guide system was obtained by applying the Euler-Lagrange formalism. It considers the dynamics of the DC motor. The generalized coordinates are q and θ.

Consider Figure 2 and the notation shown in Notation.

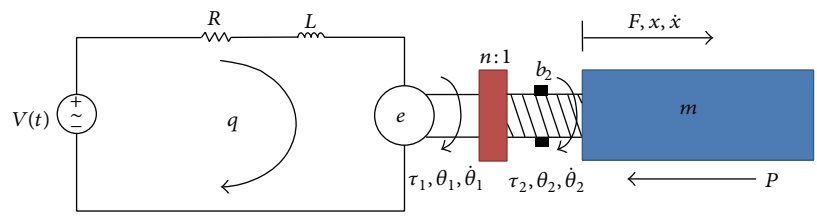


Figure 2: Schematic diagram of a linear slide system-DC motor.

2.1. Electric Motor. Considering Figure 2, the coenergy by storage of effort is given by (1), the coenergy storage of effort by (2), and the energy dissipation by (3).

Coenergy by Storage of Effort. Consider

$$T^* = \frac{1}{2}L\dot{q}^2. \tag{1}$$

The Storage of Energy Flow. Consider

$$U = 0. \tag{2}$$

Energy Dissipation. Consider

$$G = \frac{1}{2}R\dot{q}^2. \tag{3}$$

2.2. Linear Slide. The kinetic energy of the system of linear slide is given by (4) and the dissipated energy is represented by (7).

Kinetic Energy. Consider

$$K = \frac{1}{2}I\dot{\theta}_2^2 + \frac{1}{2}m\dot{x}^2. \tag{4}$$

Knowing that $x = p\theta_2$, $\dot{x} = p\dot{\theta}_2$, and $\ddot{x} = p\ddot{\theta}_2$ and substituting in (4), the following equation is obtained, with kinetic energy remaining in function of the angular velocity of the power screw (see Figure 2):

$$K = \frac{1}{2}I\dot{\theta}_2^2 + \frac{1}{2}m\left(p\dot{\theta}_2\right)^2. \tag{5}$$

Dissipated Energy. Consider

$$D = \frac{1}{2}b_2\dot{\theta}_2^2. \tag{6}$$

From (1) until (5) is represented total system in the kinetic energy and the following equation is obtained:

$$K_T = \frac{1}{2}L\dot{q}^2 + \frac{1}{2}I\dot{\theta}_2^2 + \frac{1}{2}m\left(p\dot{\theta}_2\right)^2, \tag{7}$$

where L is the Lagrangian and is given by

$$L = \frac{1}{2}L\dot{q}^2 + \frac{1}{2}I\dot{\theta}_2^2 + \frac{1}{2}m\left(p\dot{\theta}_2\right)^2. \tag{8}$$

For the generalized coordinate q,

$$\begin{aligned} \frac{\partial L}{\partial \dot{q}} &= L\dot{q}, \\ \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} &= L\ddot{q}, \\ \frac{\partial L}{\partial q} &= 0, \\ \frac{\partial D}{\partial \dot{q}} &= R\dot{q}, \end{aligned} \tag{9}$$

where

$$L\ddot{q} + R\dot{q} = V(t) - e. \tag{10}$$

Considering that $e = k_b\dot{\theta}_1$ and $\dot{\theta}_1 = n\dot{\theta}_2$ as well as $\dot{\theta}_2 = \dot{x}/p$ and replacing them in (10), the following equation is obtained:

$$L\ddot{q} + R\dot{q} = V(t) - k_b\left(\frac{n}{p}\right)\dot{x}. \tag{11}$$

Knowing that

$$q = \int_0^t i\,dt, \tag{12}$$

the following equation is obtained:

$$L\frac{di}{dt} + Ri = V(t) - k_b\left(\frac{n}{p}\right)\dot{x}. \tag{13}$$

For the generalized coordinate θ_2,

$$\begin{aligned} \frac{\partial L}{\partial \dot{\theta}_2} &= I\dot{\theta}_2 + mp\dot{\theta}_2, \\ \frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}_2} &= I\ddot{\theta}_2 + mp\ddot{\theta}_2, \\ \frac{\partial L}{\partial \theta_2} &= 0, \\ \frac{\partial D}{\partial \dot{\theta}_2} &= b_2\dot{\theta}_2. \end{aligned} \tag{14}$$

Equation of Euler-Lagrange formalism is

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} + \frac{\partial D}{\partial \dot{q}_i} = \tau, \quad \text{to } i = 1, 2, \ldots, n. \tag{15}$$

The mathematical model of the linear slide taking into account the dynamics of the DC motor is defined by

$$L\frac{di}{dt} + Ri = V(t) - k_b\left(\frac{n}{p}\right)\dot{x}, \tag{16}$$

$$\left(\frac{I}{p} + mp\right)\ddot{x} + \frac{b_2}{p}\dot{x} = nk_f i - P. \tag{17}$$

3. Optimal Control

3.1. Considering Acceleration State Variable. Based on the mathematical model (16) and (17), i is derived and from (17) it is clear that

$$\begin{aligned} i &= \left(\frac{\alpha}{nk_f}\right)\ddot{x} + \left(\frac{b_2}{pnk_f}\right)\dot{x}, \\ \frac{di}{dt} &= \left(\frac{\alpha}{nk_f}\right)\dddot{x} + \left(\frac{b_2}{pnk_f}\right)\ddot{x}. \end{aligned} \tag{18}$$

Substituting in (16), one has

$$\begin{aligned} &L\left(\frac{\alpha}{nk_f}\right)\dddot{x} + \left(\frac{b_2}{pnk_f}\right)\ddot{x} + R\left(\frac{\alpha}{nk_f}\right)\ddot{x} + \left(\frac{b_2}{pnk_f}\right)\dot{x} \\ &\quad = V(t) - K_b\left(\frac{n}{p}\right)\dot{x}. \end{aligned} \tag{19}$$

Fixing the equation, one gets

$$\begin{aligned} &L\left(\frac{\alpha}{nk_f}\right)\dddot{x} + \left(\frac{b_2}{pnk_f} + \frac{R\alpha}{nk_f}\right)\ddot{x} + \left(\frac{b_2}{pnk_f} + \frac{k_b n}{p}\right)\dot{x} \\ &\quad = u. \end{aligned} \tag{20}$$

The state variables are

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= x_3, \\ \dot{x}_3 &= -\left(\frac{nk_f}{L\alpha}\right)\left(\frac{b_2}{pnk_f} + \frac{k_b n}{p}\right)\dot{x} \\ &\quad - \left(\frac{nk_f}{L\alpha}\right)\left(\frac{b_2}{pnk_f} + \frac{R\alpha}{nk_f}\right)\ddot{x} + \frac{nk_f}{L\alpha}u. \end{aligned} \tag{21}$$

Placing the state variables in the matrix form, one gets

$$\begin{aligned} \dot{x} &= Ax + Bu, \\ y &= Cx, \\ \begin{vmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{vmatrix} &= \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -\left(\frac{nk_f}{L\alpha}\right)\left(\frac{b_2}{pnk_f} + \frac{k_b n}{p}\right) & -\left(\frac{nk_f}{L\alpha}\right)\left(\frac{b_2}{pnk_f} + \frac{R\alpha}{nk_f}\right) \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \\ x_3 \end{vmatrix} + \begin{vmatrix} 0 \\ 0 \\ \frac{nk_f}{L\alpha} \end{vmatrix} u, \\ y &= \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \end{aligned} \tag{22}$$

In control systems, often we want to select the control vector $u(t)$ such that a given performance index is minimized. A quadratic performance index, where the integration limits are 0 to ∞, so that

$$J = \int_0^\infty L(x, u)\, dt, \tag{23}$$

where $L(x, u)$ is a quadratic function or a Hermitian function of x and u, produces linear control laws; that is to say,

$$u = -Kx(t). \tag{24}$$

For weight matrixes for semipositive definite Q and R, the optimal control system is based on minimizing the performance index. This requires numerically solving the Riccati algebraic equation:

$$A^T P + PA - PBR^{-1}B^T P + Q = 0 \tag{25}$$

for a symmetric positive definite matrix P.

Finally, gains are calculated as

$$K = R^{-1}B^T P. \tag{26}$$

In the present case, u is given by

$$u = -k_1x_1 - k_2x_2 - k_3x_3. \tag{27}$$

The state weighting matrix is proposed as

$$Q = C^TC,$$
$$Q = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{28}$$

which meets nonnegative definite. As for the scalar weighting for the control input, it is chosen as

$$R = 0.01,$$
$$R = 1e - 10. \tag{29}$$

3.2. Considering State Variable Motor Current CD. By representing mathematical model equations (16) and (17) in state space, one gets

$$\dot{x}_1 = x_2,$$
$$\alpha\dot{x}_2 = -\frac{b_2}{p}x_2 + nk_fx_3,$$
$$L\dot{x}_3 = -Rx_3 - \left(\frac{nk_b}{p}\right)x_2 + u(t),$$
$$\begin{vmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 0 \\ 0 & -\frac{b_2}{\alpha p} & \frac{nk_f}{\alpha} \\ 0 & -\frac{nk_b}{Lp} & -\frac{R}{L} \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \\ x_3 \end{vmatrix} + \begin{vmatrix} 0 \\ 0 \\ \frac{1}{L} \end{vmatrix} u, \tag{30}$$
$$C = \begin{vmatrix} 1 & 0 & 0 \end{vmatrix},$$
$$D = |0|,$$
$$Q = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

with

$$R = 0.01 \tag{31}$$

for the first case and

$$R = 1e - 10 \tag{32}$$

for the second case.

3.3. LQR Optimal Control in Virtual Prototype. A test was performed using a virtual prototype of the linear slide, Figure 3, in room ADAMS MSC together with Matlab-Simulink in order to test the effectiveness of optimal control LQR; the results of this experiment can be seen in Section 5.

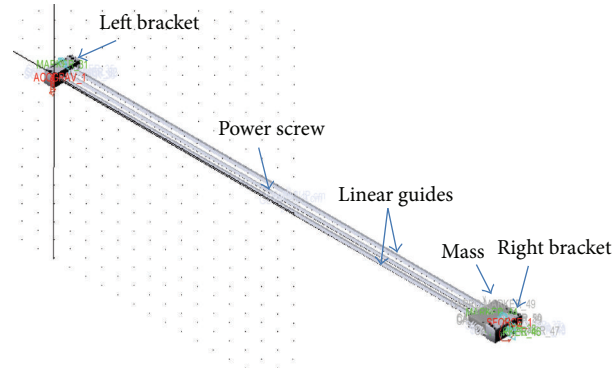


Figure 3: Virtual prototype of the linear slide.

4. Control Based on Differential Flatness

This third-order linear system (30), where its Kalman controllability matrix is calculated by the following expression: $C = [B, AB, A^2B]$, is given as

$$C = \begin{vmatrix} 0 & 0 & \frac{nk_f}{\alpha} \\ 0 & \frac{nk_f}{\alpha} & -\frac{Rnk_f}{L\alpha} - \frac{nb_2k_f}{p\alpha^2} \\ 1 & -\frac{R}{L} & \frac{R^2}{L^2} - \frac{n^2k_bk_f}{Lp\alpha} \end{vmatrix} \det C = -\frac{n^2k_f^2}{\alpha^2} \neq 0. \tag{33}$$

Since the determinant is nonzero, then the system is controllable and, therefore, is differentially flat (Sira-Ramírez and Agrawal [8]). The flat output of a linear system input output (I/O) is obtained by multiplying the inverse matrix of controllability by the state vector x, associated with the system. Column vector obtained by multiplying the last line is chosen to obtain the flat output (Linares-Flores and Sira-Ramírez [7]). In particular for reducing motor-drive CD, flat output system F is calculated as

$$F = \begin{vmatrix} 0 & 0 & 1 \end{vmatrix} C^{-1} \begin{vmatrix} x_1 \\ x_2 \\ x_3 \end{vmatrix},$$
$$F = \begin{vmatrix} 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} \frac{(k_bk_fn^2 + Rb_2)}{Lnpk_f} & \frac{R\alpha}{Lnk_f} & 1 \\ \frac{(Lb_2 + Rp\alpha)}{Lnpk_f} & \frac{\alpha}{nk_f} & 0 \\ \frac{\alpha}{nk_f} & 0 & 0 \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \\ x_3 \end{vmatrix} \tag{34}$$
$$= \frac{\alpha}{nk_f}x_1.$$

Table 1: Parameter values for simulation.

Parameter	Value
n = speed ratio	0.19
P = force opposite to the movement of m	0.01 N
b_2 = coefficient of viscous friction	2
p = pitch of the screw thread power	0.00196 m
m = mass to be displaced	10 kg
J = moment of inertia	0.0000014 kg m^2
V = voltage	12 volts
k_b = constant emf	0.022 (Vs)/rad
R = resistance of DC motor	5.3 ohm
L = motor inductance	0.00058 henries
K_f = constant torque	90 (N-m)/A

Hence, we have chosen x_1 as the flat output. This flat output provides the following differential parametrization of the system variables:

$$\begin{aligned} x_1 &= F, \\ x_2 &= \dot{F}, \\ x_3 &= \left(\frac{\alpha}{nk_f}\right)\ddot{F} + \left(\frac{b_2}{nk_f p}\right)\dot{F} + \frac{P}{nk_f p}, \end{aligned} \tag{35}$$

and the control input:

$$\begin{aligned} u = {} & \left[\frac{L\alpha}{nk_f}\right]\dddot{F} + \left[\frac{Lb_2}{nk_f p} + \frac{R\alpha}{nk_f}\right]\ddot{F} \\ & + \left[\frac{Rb_2}{nk_f p} + \frac{nk_b}{p}\right]\dot{F} + \frac{RP}{nk_f p}. \end{aligned} \tag{36}$$

Considering (36) and in order to verify the stability of the system, applying the Laplace transform, the following holds:

$$\begin{aligned} u = {} & \left[\frac{L\alpha}{nk_f}\right]s^3 + \left[\frac{Lb_2}{nk_f p} + \frac{R\alpha}{nk_f}\right]s^2 \\ & + \left[\frac{Rb_2}{nk_f p} + \frac{nk_b}{p}\right]s + \frac{RP}{nk_f p}. \end{aligned} \tag{37}$$

Taking into account the values of Table 1, the system transfer function is

$$\begin{aligned} & \frac{F(s)}{u(s)} \\ & \quad = \frac{1}{6.8854 * 10^{-7}s^3 + 4.0902 * 10^{-2}s^2 + 318.40s}. \end{aligned} \tag{38}$$

Plotting the locus of roots verifies that the system is stable when these are in the left half of the complex plane (Figure 4).

Hence, we have the fact that all state variables and the control input are in terms of F and its successive derivatives, where it denotes the position of the mass as x. The differential parametrization above allows obtaining the equilibrium for

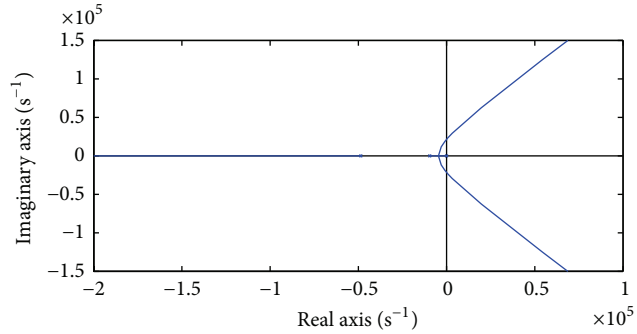


Figure 4: Roots locus.

the system in terms of the equilibrium values of the flat output and the disturbance inputs. Thus,

$$\begin{aligned} \overline{x}_1 &= \overline{F}_d, \\ \overline{x}_2 &= 0, \\ \overline{x}_3 &= \frac{\overline{P}}{nk_f p}, \\ \overline{u} &= \frac{R\overline{P}}{nk_f p}. \end{aligned} \tag{39}$$

From (36), we design the average controller based on differential flatness property. Thus, we replace the higher-order derivative of the flat output by a virtual controller (see Slotine and Li [24]) resulting in the following:

$$\dddot{F} = v_{\text{aux}} = -k_2\ddot{F} - k_1\dot{F} - k_0\left(F - \overline{F}_d\right). \tag{40}$$

For the tracking controller design, we use a nominal desired linear displacement profile F_d that exhibits a rather smooth start for the motor linear slide system. This is specified using an interpolating Bézier polynomial of 10th order where the initial linear displacement is set to be $F_{\text{ini}} = 0$ m valid until $t_{\text{ini}} = 15$ sec and the final desired value of the angular velocity is specified as $F_{\text{fin}} = 0.5$ m to be reached at $t_{\text{fin}} = 45$ sec; that is, we used

$$F_d = \begin{cases} F_{\text{ini}}, & t < t_{\text{ini}}, \\ F_{\text{ini}} + \left(F_{\text{fin}} - F_{\text{ini}}\right)b_x, & t_{\text{ini}} \le t \le t_{\text{fin}}, \\ F_{\text{fin}}, & t > t_{\text{fin}}, \end{cases} \tag{41}$$

where $b_x(t, t_{\text{ini}}, t_{\text{fin}})$ is a polynomial function of time, exhibiting a sufficient number of zero derivatives at times t_{ini} and t_{fin}, while also satisfying $b_x(t_{\text{ini}}, t_{\text{ini}}, t_{\text{fin}}) = 0$ and $b_x(t_{\text{fin}}, t_{\text{ini}}, t_{\text{fin}}) = 1$. For instance, one such polynomial may be given by

$$\begin{aligned} & b_x(t, t_{\text{ini}}, t_{\text{fin}}) \\ & \quad = \beta^5\left[r_1 - r_2\beta + r_3\beta^2 - r_4\beta^3 + r_5\beta^4 - r_6\beta^5\right], \\ & \beta = \left(\frac{t - t_{\text{ini}}}{t_{\text{fin}} - t_{\text{ini}}}\right). \end{aligned} \tag{42}$$

The differential parametrization of the control input, u, in terms of F_d shows that the proposed flat output trajectory tracking task is that of controlling the third derivative of F_d by means of v_{aux}:

$$v_{\text{aux}} = \dddot{F}_d - k_2\left(\ddot{F} - \ddot{F}_d\right) - k_1\left(\dot{F} - \dot{F}_d\right) - k_0\left(F - F_d\right). \quad (43)$$

Therefore, the linear displacement-tracking controller is

$$u = \left[\frac{L\alpha}{nk_f}\right] v_{\text{aux}} + \left[\frac{Lb_2}{nk_f p} + \frac{R\alpha}{nk_f}\right]\ddot{F}_d + \left[\frac{Rb_2}{nk_f p} + \frac{nk_b}{p}\right]\dot{F}_d + \frac{RP}{nk_f p}. \quad (44)$$

Under these circumstances, the closed loop system tracking error, $e = F - F_d$, satisfies the linear differential equation

$$e^{(3)} + k_2\ddot{e} + k_1\dot{e} + k_0 e = 0. \quad (45)$$

The appropriate choice of the constant coefficients $\{k_2, k_1, k_0\}$, as coefficients of the third-order Hurwitz polynomial, guarantees the asymptotic exponential stability to zero of the tracking error, e. One such choice, yielding a characteristic polynomial of the form $(s+\gamma)(s^2 + 2\xi\omega_n s + \omega_n^2)$ with $\gamma > 0$, $0 < \xi < 1$, and $\omega_n > 0$, is given by

$$\begin{aligned} k_0 &= \omega_n^2\gamma, \\ k_1 &= 2\xi\omega_n\gamma + \omega_n^2, \\ k_2 &= 2\xi\omega_n + \gamma. \end{aligned} \quad (46)$$

The virtual controller, v_{aux}, achieves a smooth start for the linear displacement of the system. If we apply an unknown constant load torque on the system, we have to include a term of integral action into the virtual control (43). Thus, we minimize the tracking error near to zero.

5. Simulation Results

These results were obtained using the parameter values of Table 1.

In Figure 5, one can see the movement of the mass 0.5 meters in 60 seconds following a path originated by a tenth-order Bézier polynomial; the tracking error can be seen as zero in the whole path; in this case, the acceleration is taken as a state variable.

Figure 6 shows the graphs of results taking into account the current as state variable; considering the matrix $R = 0.01$ and considering the matrix $R = 1e-10$, Figure 7, in both cases the tracking error can be seen as zero in the whole path.

Figure 8 shows the results of simulation based on differential flatness controller observed; it also shows that it is capable of moving the mass to a position 0.5 meters along a desired path of a Bézier polynomial of tenth order.

The results of the performance index are as follows: the two controllers are shown in Figures 9, 10, and 11; the indexes of performance of the optimal controller for two values of the matrix R are shown.

Figure 12 shows the results using a virtual prototype environment ADAMS MSC and MATLAB-Simulink; we can see that the offset for this test was set at 0.7 meters corresponding to the desired position; a tenth-order Bézier polynomial was used as desired path.

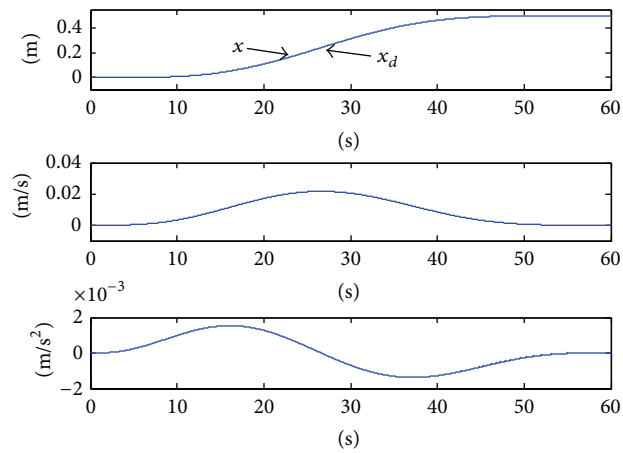


Figure 5: Graphics optimal controller response with $R = 1e-10$.

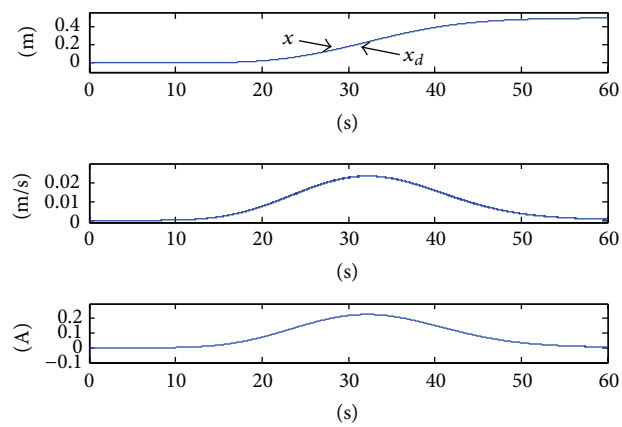


Figure 6: Graphics optimal controller response with $R = 0.01$.

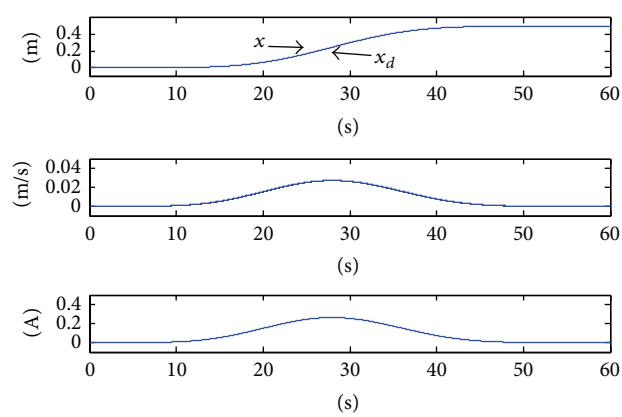


Figure 7: Graphics optimal controller response with $R = 1e-10$.

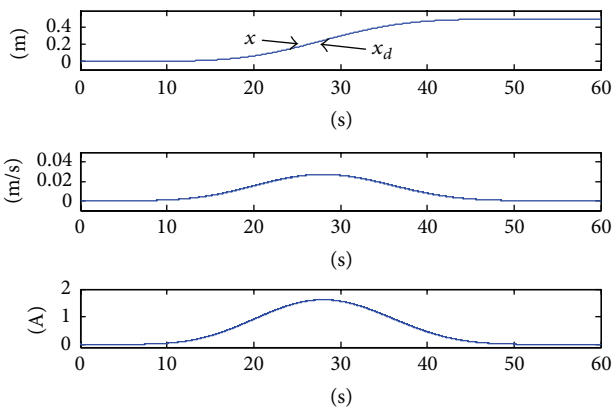


Figure 8: Graphics controller response flatness based on difference.

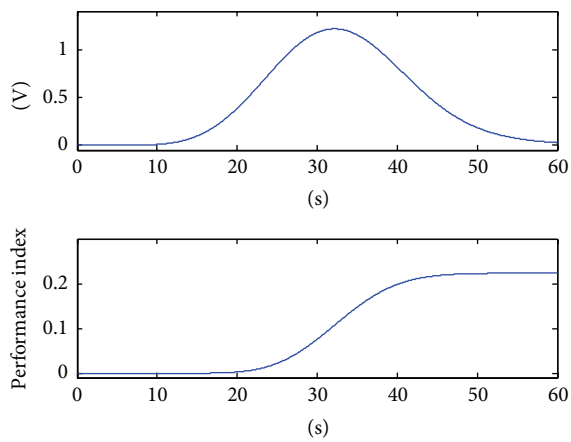


Figure 9: Performance Index LQR optimal controller with $R = 0.01$.

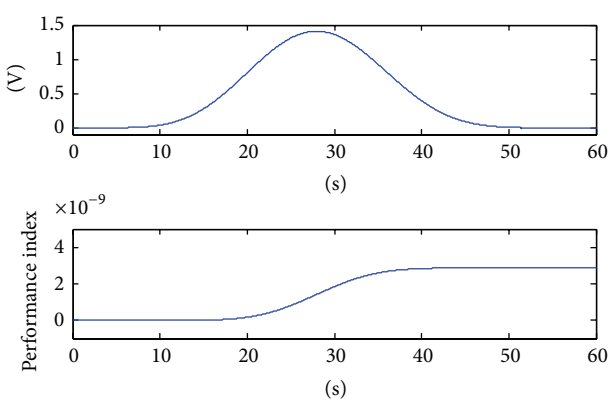


Figure 10: Performance Index LQR optimal controller with $R = 1e-10$.

6. Results with Linear Slide

Experimentation was held using a physical linear slide (Figure 13); the laws of LQR control are implemented based on differential flatness with the same parameter values and gains controllers; the results of the experiment can be seen in Figures 14, 15, and 16, where it can be seen that in all three cases the controller is able to zero the position error; it is clear

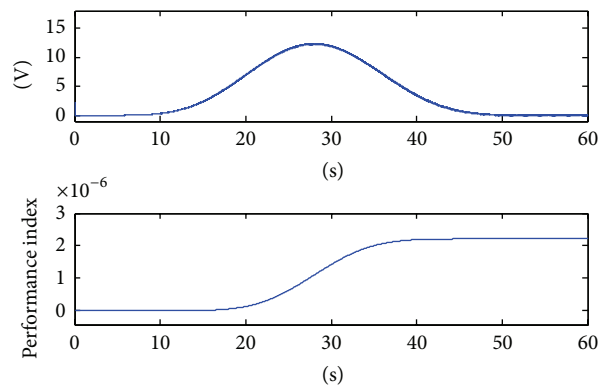


Figure 11: Controller performance index based on differential flatness.

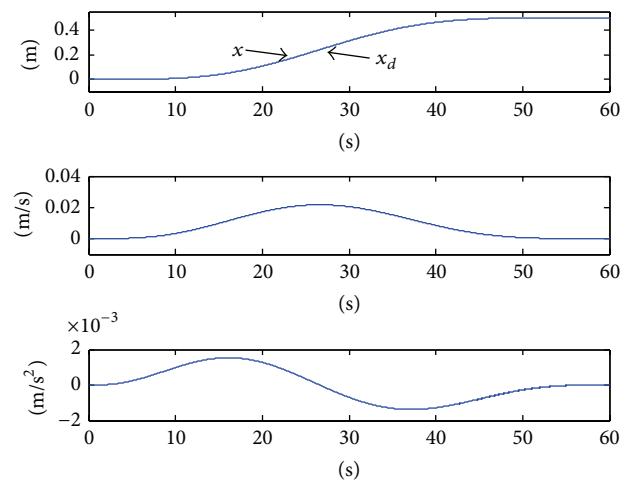


Figure 12: Graphic of LQR optimal controller results in the virtual prototype.

that the experiment was carried out with an acquisition card myRIO data of National Instruments; the measurement units were centimeters so that they should be taken into account in the interpretation of Figures 14, 15, and 16.

It is verified by the results the correct operation for achieving desired trajectories and references.

On the x-axis is the number of samples; to determine the time it must be multiplied by the sampling period, that is, 0.1 s.

7. Conclusions

By applying control techniques and optimal LQR and based on differential flatness it can be seen that the results for the trajectory tracking are highly achievable in both cases, with regard to the performance indices of each controller for the optimal controller; a great disadvantage exists since it is necessary that the values of the weighting matrix R be too low to achieve a speed of response adapted to the needs of the plant to controlling; this means that the gains are higher further input values, such as the voltage that must be adjusted

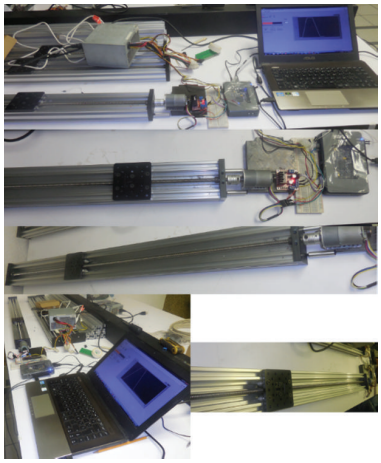


Figure 13: Experiment with physical linear slide.

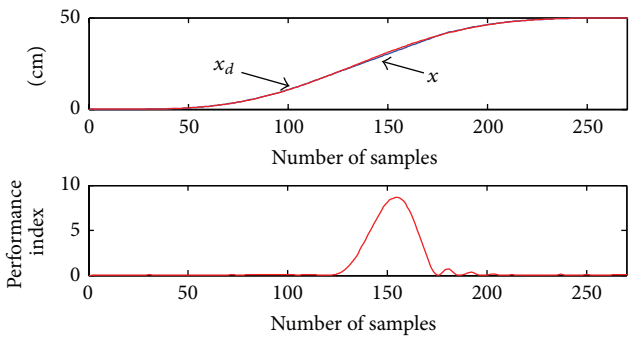


Figure 14: Controller based on differential flatness.

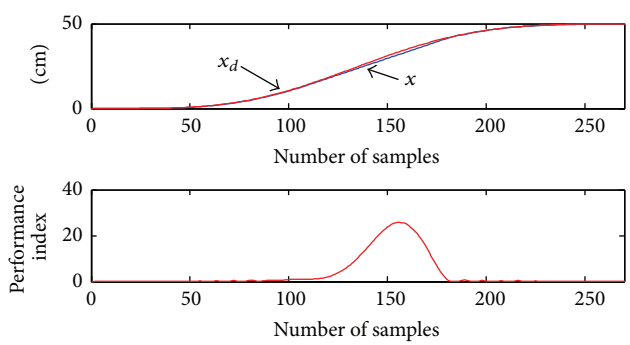


Figure 15: Performance Index LQR optimal controller with $R = 0.01$.

by trial and error of said matrix with the purpose of respecting the voltages of the actuators.

Figures 12 and 13 show that the values of the performance index are minimizing in the measure that the value of the weighting matrix, R, is minor compared to the index controller performance differential flatness, if it is minor, but this is not guaranteed to be optimal because the values of the control force are lower than those required for the good functioning of DC motors.

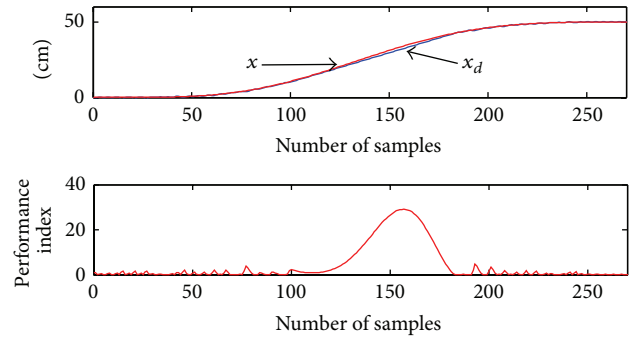


Figure 16: Performance Index LQR optimal controller with $R = 1e-10$.

In conclusion it follows that the use of control laws differential flatness is by far better than the use of laws of optimal control; the concept of optimality, in this particular case, is lost due to the randomness representing search for the value of the R matrix to obtain appropriate response speed without sacrificing system actuators.

Notation

n: Speed ratio
P: Force opposite to the movement of m
b_2: Coefficient of viscous friction
p: Pitch of the screw thread power
m: Mass to be displaced
J: Moment of inertia
V: Voltage
k_b: Constant emf
R: Resistance of DC motor
L: Motor inductance
k_{tao}: Constant torque
θ_1: Angular position of DC motor
$\dot{\theta}_1$: Angular velocity of DC motor
θ_2: Angular position of power screw
$\dot{\theta}_2$: Angular velocity of power screw
τ_1: Torque delivered by the DC motor
τ_2: Torque delivered by the speed reducer
x: Displacement of the mass
$\dot{x}$: Velocity of the mass
F: Force of displacement of the mass.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] A. Blanco Ortega, F. A. Gómez-Becerra, L. G. Vela Valdés, and R. O. Delgado Arcega, "A generalized proportional integral controller for an ankle rehabilitation machine based on an XY table," in *Proceedings of the IEEE International Conference on Mechatronics, Electronics and Automotive Engineering (ICMEAE '13)*, pp. 152–157, IEEE, Morelos, Mexico, November 2013.

[2] C. H. G. Valdivia, J. L. C. Escobedo, A. B. Ortega, M. A. O. Salazar, and F. A. G. Becerra, "Diseño y control de un sistema interactivo para la rehabilitación de tobillo: TobiBot," *Ingeniería Mecánica. Tecnología y Desarrollo*, vol. 5, no. 1, pp. 255–264, 2014.

[3] A. Blanco Ortega, J. Isidro Godoy, E. Quintero Mármol, and L. G. Vela Valdés, "Robot paralelo para rehabilitación asistida de tobillo," in *X Congreso Internacional sobre Innovación y Desarrollo Tecnológico (CIINDET '13)*, Cuernavaca, Mexico, March 2013.

[4] A. Blanco Ortega, H. R. Azcaray Rivera, R. F. Vázquez Bautista, and L. J. Morales Mendoza, "Máquina de Rehabilitación de Tobillo: prototipo virtual y físico," in *Proceedings of the 10th Congreso Internacional sobre Innovación y Desarrollo Tecnológico (CIINDET '13)*, vol. 1, Cuernavaca Morelos, México, March 2013.

[5] M. Fliess, J. Lévine, P. Martin, and P. Rouchon, "Flatness and defect of non-linear systems: introductory theory and examples," *International Journal of Control*, vol. 61, no. 6, pp. 1327–1361, 1995.

[6] M. Fliess and R. Marquez, "Continuous-time linear predictive control and flatness: a module-theoretic setting with examples," *International Journal of Control*, vol. 73, no. 7, pp. 606–623, 2000.

[7] J. Linares-Flores and H. Sira-Ramírez, "DC motor velocity control through a DC-to-DC power converter," in *Proceedings of the 43rd IEEE Conference on Decision & Control*, vol. 5, pp. 5297–5302, Nassau, Bahamas, December 2004.

[8] H. Sira-Ramírez and S. K. Agrawal, *Diĕerentially Flat Systems*, Marcel Dekker, Inc, 2004.

[9] P. Thounthong, S. Pierfederici, J.-P. Martin, M. Hinaje, and B. Davat, "Modeling and control of fuel cell/supercapacitor hybrid source based on differential flatness control," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 6, pp. 2700–2710, 2010.

[10] M. Jörgl and H. Gattringer, "Dynamical modeling and flatness based control of a belt drive system," *Proceedings in Applied Mathematics and Mechanics*, vol. 14, no. 1, pp. 887–888, 2014.

[11] P. Thounthong and S. Pierfederici, "A new control law based on the differential flatness principle for multiphase interleaved DC–DC converter," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 57, no. 11, pp. 903–907, 2010.

[12] M. Fliess, J. Lévine, P. Martin, and P. Rouchon, "A Lie-Bäcklund approach to equivalence and flatness of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 44, no. 5, pp. 922–937, 1999.

[13] M. A. Danzer, J. Wilhelm, H. Aschemann, and E. P. Hofer, "Model-based control of cathode pressure and oxygen excess ratio of a PEM fuel cell system," *Journal of Power Sources*, vol. 176, no. 2, pp. 515–522, 2008.

[14] A. Gensior, H. Sira-Ramirez, J. Rudolph, and H. Guldner, "On some nonlinear current controllers for three-phase boost rectifiers," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 2, pp. 360–370, 2009.

[15] E. Song, A. F. Lynch, and V. Dinavahi, "Experimental validation of nonlinear control for a voltage source converter," *IEEE Transactions on Control Systems Technology*, vol. 17, no. 5, pp. 1135–1144, 2009.

[16] F. U. Syed, M. L. Kuang, M. Smith, S. Okubo, and H. Ying, "Fuzzy gain-scheduling proportional–integral control for improving engine power and speed behavior in a hybrid electric vehicle," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 1, pp. 69–84, 2009.

[17] T. Rabbani, S. Munier, D. Dorchies, P.-O. Malaterre, A. Bayen, and X. Litrico, "Flatness-based control of open-channel flow in an irrigation canal using SCADA," *IEEE Control Systems Magazine*, vol. 29, no. 5, pp. 22–30, 2009.

[18] S. K. Agrawal, K. Pathak, J. Franch, R. Lampariello, and G. Hirzinger, "A differentially flat open-chain space robot with arbitrarily oriented joint axes and two momentum wheels at the base," *IEEE Transactions on Automatic Control*, vol. 54, no. 9, pp. 2185–2191, 2009.

[19] K. Hassani and W.-S. Lee, "Optimal tuning of linear quadratic regulators using quantum particle swarm optimization," in *Proceedings of the International Conference on Control, Dynamic Systems, and Robotics (CDSR '14)*, Paper no. 59, Ottawa, Canada, May 2014.

[20] F. L. Lewis, *Optimal Feedback Control: Practical Performance and Design Algorithms for Industrial and Aerospace Systems*, UTA Research Institute The University of Texas at Arlington, USA, 2012.

[21] G.-R. Yu and R.-C. Hwang, "Optimal PID speed control of brush less DC motors using LQR approach," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC '04)*, pp. 473–478, Hague, The Netherlands, October 2004.

[22] M. Ruderman, J. Krettek, F. Hoffmann, and T. Bertram, "Optimal state space control of DC motor," in *Proceedings of the 17th IFAC World Congress*, International Federation of Automatic Control, Seoul, Republic of Korea, July 2008.

[23] L. B. Prasad, B. Tyagi, and H. O. Gupta, "Optimal control of nonlinear inverted pendulum system using PID controller and LQR: performance analysis without and with disturbance input," *International Journal of Automation and Computing*, vol. 11, no. 6, pp. 661–670, 2014.

[24] J.-J. E. Slotine and W. Li, *Applied Nonlinear Control*, Prentice Hall, 1991.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 510241, 11 pages
http://dx.doi.org/10.1155/2015/510241

Research Article

A Posteriori Error Estimate for Finite Volume Element Method of the Second-Order Hyperbolic Equations

Chuanjun Chen,[1] Xin Zhao,[2] and Yuanyuan Zhang[1]

[1] *School of Mathematics and Information Sciences, Yantai University, Yantai 264005, China*
[2] *School of Mathematics and Computational Science, Xiangtan University, Xiangtan 411105, China*

Correspondence should be addressed to Chuanjun Chen; cjchen@ytu.edu.cn

Received 6 September 2015; Accepted 18 November 2015

Academic Editor: Xinguang Zhang

We establish a posteriori error estimate for finite volume element method of a second-order hyperbolic equation. Residual-type a posteriori error estimator is derived. The computable upper and lower bounds on the error in the H^1-norm are established. Numerical experiments are provided to illustrate the performance of the proposed estimator.

1. Introduction

The finite volume element method is a class of important numerical tools for solving partial differential equations. Due to the local conservation property and some other attractive properties, it is wildly used in many engineering fields, such as heat and mass transfer, fluid mechanics, and petroleum engineering, especially for those arising from conservation laws including mass, momentum, and energy. For the second-order hyperbolic equations, Li et al. [1] have proved the optimal order of convergence in H^1-norm. In [2], Kumar et al. have proved optimal order of convergence in L^2 and H^1-norm for the semidiscrete scheme and quasi-optimal order of convergence in maximum norm.

Since the pioneering work of Babuvška and Rheinboldt [3], the adaptive finite element methods based on a posteriori error estimates have become a central theme in scientific and engineering computations. Adaptive algorithm is among the most important means to boost accuracy and efficiency of the finite element discretization. The main idea of adaptive algorithm is to use the error indicator as a guide which shows whether further refinement of meshes is necessary. A computable a posteriori error estimator plays a crucial role in an adaptive procedure. A posteriori error analysis for the finite volume element method has been studied in [4–6] for the second-order elliptic problem, in [7–9] for the convection-diffusion equations, in [10] for the parabolic problems, in [11] for a model distributed optimal problem governed by linear parabolic equations, in [12] for the Stokes problem in two dimensions, and in [13] for the second-order hyperbolic equations.

However, to the best of our knowledge, there are few works related to the a posteriori error estimates of the finite volume element method for the second-order hyperbolic problems. The aim of this paper is to establish residual-type a posteriori error estimator of the finite volume element method for the second-order hyperbolic equation. We first construct a computable a posteriori error estimator of the finite volume element method. Then we analyze the residual-type a posteriori error estimates and obtain the computable upper and lower bounds on the error in the H^1-norm.

The organization of this paper is stated as follows. In Section 2, we present the framework of the finite volume element method for the second-order hyperbolic equation. In Section 3, we establish the residual-type a posteriori error estimator of the finite volume element method and derive the upper and lower bounds on the error in the H^1-norm. We provide some numerical experiments to illustrate the performance of the error estimator in Section 4.

2. Finite Volume Element Formulation

We use standard notation for Sobolev spaces $W^{s,p}(\Omega)$ with the norm $\|u\|_{s,p,\Omega}$ [14]. In order to simplify the notation, we

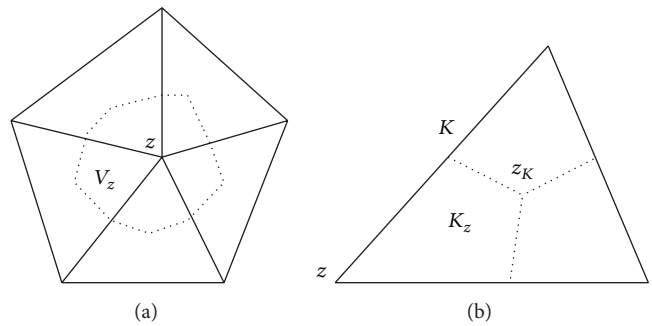


Figure 1: (a) The dotted line shows the boundary of the corresponding control volume V_z with z, a common vertex. (b) A triangle K is partitioned into three subregions K_z.

denote $W^{s,2}(\Omega)$ by $H^s(\Omega)$ and omit the index $p = 2$ and Ω whenever possible.

In this paper, we consider the following second-order hyperbolic problem:

$$\begin{aligned} u_{tt} - \nabla \cdot (a(x) \nabla u) &= f(x,t), \quad \text{in } \Omega \times (0,T], \\ u(x,t) &= 0, \quad \text{on } \partial\Omega \times (0,T], \\ u(x,0) &= u_0(x), \\ u_t(x,0) &= v_0(x), \\ &\text{in } \Omega, \end{aligned} \tag{1}$$

where $\Omega \subset \mathbb{R}^2$ is a polygonal bounded cross section, possessed with a Lipschitz boundary $\partial\Omega$. For simplicity, the right-hand side f is assumed to be measurable and square-integrable on $\Omega \times (0,T]$ and to be continuous with respect to time. The initial datum u_0 and v_0 are assumed to be measurable and square-integrable on Ω. $a(x,t) = (a_{ij}(x,t))_{i,j=1}^2$ is a real-valued smooth matrix function, uniformly symmetric, and positive definite in Ω.

The corresponding variational problem is to find $u \in H_0^1(\Omega)$, for $t > 0$, satisfying

$$(u_{tt}, v) + a(u,v) = (f,v), \quad \forall v \in H_0^1(\Omega), \tag{2}$$

where the bilinear form $a(\cdot,\cdot)$ is defined by

$$a(u,v) = \int_\Omega a(x) \nabla u \cdot \nabla v \, dx, \quad \forall u, v \in H_0^1(\Omega). \tag{3}$$

Denote by T_h the primal quasi-uniform triangulation of Ω with $h = \max h_K$, where h_K is the diameter of the triangle $K \in T_h$. Let $\mathscr{U}_h$ be the standard conforming finite element space of piecewise linear functions, defined on the triangulation T_h:

$$\mathscr{U}_h = \left\{ u \in C(\overline{\Omega}) : u|_K \text{ is linear and } u|_{\partial\Omega} = 0, \ \forall K \in T_h \right\}. \tag{4}$$

Denote by T_h^* the dual partition which is constructed in the same way as in [1, 15]. Let z_K be the barycenter of K. We connect z_K with the midpoints of the edges of K by straight line, thus partitioning K into three quadrilaterals K_z, $z \in Z_h(K)$, where $Z_h(K)$ are the vertices of K. Then with each vertex $z \in Z_h = \cup_{K \in T_h} Z_h(K)$, we associate a control volume V_z, which consists of the union of the subregions K_z, sharing the vertex z (see Figure 1). Finally, we obtain a group of control volumes covering the domain Ω, which is called the dual partition T_h^* of the triangulation T_h. Denote by Z_h^0 the set of interior vertices of Z_h and denote by $\mathscr{E}_h$ the set of all interior edges of T_h, respectively.

The partition T_h^* is regular or quasi-uniform, if there exists a positive constant $C > 0$ such that

$$C^{-1} h^2 \le \operatorname{meas}(V_z) \le C h^2, \quad \forall V_z \in T_h^*. \tag{5}$$

The dual partition T_h^* will also be quasi-uniform [5] if the finite element triangulation T_h is quasi-uniform. The test function space $\mathscr{V}_h$ is defined by

$$\mathscr{V}_h = \left\{ v \in L^2(\Omega) : v|_{V_z} \text{ is constant and } v|_{\partial\Omega} = 0 \ \forall V_z \in T_h^* \right\}. \tag{6}$$

For any $u_h \in \mathscr{U}_h$, we define an interpolation operator $\Pi_h : \mathscr{U}_h \to \mathscr{V}_h$, such that

$$\Pi_h u_h = \sum_{z \in Z_h^0} u_h(z) \Psi_z, \tag{7}$$

where Ψ_z is the characteristic function of the control volume V_z.

According to [16], for each $u_h \in \mathscr{U}_h$, there exists a positive constant C independent of h, such that Π_h satisfies the following inequality:

$$\|u_h - \Pi_h u_h\|_{0,K} \le C h_K |u_h|_{1,K}, \quad \forall K \in T_h. \tag{8}$$

Introduce the following adjoint elliptic problem:

$$-\nabla \cdot (a(x) \nabla u) = f \quad \text{in } \Omega, \text{ with } u = 0 \text{ on } \partial\Omega. \tag{9}$$

Denote by $\mathscr{T} : L^2(\Omega) \to H^2(\Omega) \bigcap H_0^1(\Omega)$ the solution operator of this problem, so that

$$a(\mathscr{T} f, \varphi) = (f, \varphi), \quad \forall \varphi \in H_0^1(\Omega). \tag{10}$$

Define negative norms by

$$\|v\|_{-s} = \sup\left\{\frac{(v,\varphi)}{\|\varphi\|_s};\ \varphi \in H^s(\Omega)\right\}, \quad \text{for } s \geq 0 \text{ integer.} \tag{11}$$

In fact, by Cauchy-Schwarz inequality, we obtain

$$\frac{(v,\varphi)}{\|\varphi\|_1} \leq \frac{\|v\|\,\|\varphi\|}{\|\varphi\|_1} \leq \frac{\|v\|\,\|\varphi\|_1}{\|\varphi\|_1} = \|v\|. \tag{12}$$

For our error analysis in the next section, it will be convenient to introduce such a norm defined by

$$|v|_{-s} = \left(\mathcal{T}^s v, v\right)^{1/2}, \quad \text{for } s \geq 0 \text{ integer.} \tag{13}$$

According to Thomée [17], we have the following lemma.

Lemma 1. *The norm $|v|_{-s}$ is equivalent to $\|v\|_{-s}$ and $(\mathcal{T} f, g) = (f, \mathcal{T} g)$, where s is a nonnegative integer. Particularly, $\|\mathcal{T} v\|_1$ is equivalent to $\|v\|_{-1}$ when $s = 1$.*

In order to get the fully discrete finite volume element method of (1), we give a partition of the time interval $[0,T]$: $0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$. Let $\tau_n = t_n - t_{n-1}$, $\tau = \max_{1\leq n\leq N}\tau_n$, $U_h^n = U_h(t_n)$, and $U_h^{n,1/2} = (U_h^{n+1} + U_h^{n-1})/2$. With the help of Π_h, we obtain the fully discrete finite volume element method of (1): to find $U_h^n \in \mathcal{U}_h$, for $1 \leq n \leq N$, such that

$$\begin{aligned} &\left(\partial_t\bar{\partial}U_h^n, \Pi_h\chi\right) + a\left(U_h^{n,1/2}, \Pi_h\chi\right) = \left(f^n, \Pi_h\chi\right), \\ &\qquad\qquad \forall\chi \in \mathcal{U}_h, \\ &U_h^0 = u_0, \\ &\bar{\partial}U_h^1 = v_0, \end{aligned} \tag{14}$$

where

$$\begin{aligned} \partial_t\bar{\partial}U_h^n &= \frac{\partial_t U_h^n - \partial_t U_h^{n-1}}{\tau_n} \\ &= \frac{\left(U_h^{n+1} - U_h^n\right)/\tau_{n+1} - \left(U_h^n - U_h^{n-1}\right)/\tau_n}{\tau_n}. \end{aligned} \tag{15}$$

By setting $v = \partial u/\partial t = u_t$ and $\mathcal{Y} = \binom{u}{v}$, the notation $\nabla\cdot(a(x)\nabla)\phi = \nabla\cdot(a(x)\nabla\phi)$, (1) can equivalently be written as

$$\mathcal{Y}_t - \begin{pmatrix} 0 & 1 \\ \nabla\cdot(a(x)\nabla) & 0 \end{pmatrix}\mathcal{Y} = F, \tag{16}$$

where $F = \binom{0}{f}$.

Let $V_h^n = \bar{\partial}U_h^{n+1}$; we define

$$\begin{aligned} U_\tau &= \frac{t - t_{n-1}}{\tau_n}U_h^n + \left(1 - \frac{t - t_{n-1}}{\tau_n}\right)U_h^{n-1}, \quad 1 \leq n \leq N, \\ V_\tau &= \frac{t - t_{n-1}}{\tau_n}V_h^n + \left(1 - \frac{t - t_{n-1}}{\tau_n}\right)V_h^{n-1}, \quad 1 \leq n \leq N. \end{aligned} \tag{17}$$

The residual system, with $Y_\tau = \binom{U_\tau}{V_\tau}$, is defined as follows:

$$\begin{aligned} &\left(\mathcal{Y} - Y_\tau\right)_t - \begin{pmatrix} 0 & 1 \\ \nabla\cdot(a(x)\nabla) & 0 \end{pmatrix}\left(\mathcal{Y} - Y_\tau\right) = \begin{pmatrix} P_u \\ P_v \end{pmatrix} \\ &\qquad\qquad \text{in } \Omega\times[0,T], \\ &u - U_\tau = 0 \\ &\qquad\qquad \text{on } \partial\Omega\times[0,T], \\ &\left(\mathcal{Y} - Y_\tau\right)(\cdot,0) = 0 \quad \text{in } \Omega, \end{aligned} \tag{18}$$

where the quantities P_u in $L^1(0,T;L^2(\Omega))$ and P_v in $L^1(0,T;H^{-1}(\Omega))$ are affine functions on each interval $[t_{n-1}, t_n]$, $1 \leq n \leq N$, that

$$P_u(\cdot,t) = \begin{cases} V_\tau - V_h^{n-1}, & 2 \leq n \leq N, \\ 0, & n = 1. \end{cases} \tag{19}$$

And the quantities P_v are defined as follows.

From the fully discrete algorithm (14), for any $\varphi \in H_0^1(\Omega)$, $v \in \mathcal{U}_h$, we have

$$\begin{aligned} &\left(\partial_t\bar{\partial}U_h^n, \varphi\right) + a\left(U_h^{n,1/2}, \varphi\right) \\ &\quad = -\left(f^n - \partial_t\bar{\partial}U_h^n, \varphi - \Pi_h v\right) + \left(f^n, \varphi\right) \\ &\qquad + a\left(U_h^{n,1/2}, \varphi\right) - a\left(U_h^{n,1/2}, \Pi_h v\right). \end{aligned} \tag{20}$$

Since $(V_\tau)_t = \partial_t\bar{\partial}U_h^n$, by (2) and (20), for $t \in (t^{n-1}, t^n]$, we get

$$\begin{aligned} &\left((v - V_\tau)_t, \varphi\right) + a\left(u - U_h^{n,1/2}, \varphi\right) \\ &\quad = \left(f^n - \partial_t\bar{\partial}U_h^n, \varphi - \Pi_h v\right) + \left(f - f^n, \varphi\right) \\ &\qquad - a\left(U_h^{n,1/2}, \varphi\right) + a\left(U_h^{n,1/2}, \Pi_h v\right). \end{aligned} \tag{21}$$

Adding the term $a(U_h^{n,1/2} - U_\tau, \varphi)$ into the two hand sides of (21), we get

$$\begin{aligned} &\left((v - V_\tau)_t, \varphi\right) + a\left(u - U_\tau, \varphi\right) \\ &\quad = \left(f^n - \partial_t\bar{\partial}U_h^n, \varphi - \Pi_h v\right) + \left(f - f^n, \varphi\right) \\ &\qquad - \left[a\left(U_h^{n,1/2}, \varphi\right) - a\left(U_h^{n,1/2}, \Pi_h v\right)\right] \\ &\qquad + a\left(U_h^{n,1/2} - U_\tau, \varphi\right). \end{aligned} \tag{22}$$

So on each interval $[t_{n-1}, t_n]$ $(2 \leq n \leq N)$, we have

$$\begin{aligned} \left(P_v, \varphi\right) &= \left(f^n - \partial_t\bar{\partial}U_h^n, \varphi - \Pi_h v\right) \\ &\quad - \left[a\left(U_h^{n,1/2}, \varphi\right) - a\left(U_h^{n,1/2}, \Pi_h v\right)\right] \\ &\quad + a\left(U_h^{n,1/2} - U_\tau, \varphi\right) + \left(f - f^n, \varphi\right), \\ &\qquad \forall\varphi \in H_0^1(\Omega),\ v \in \mathcal{U}_h. \end{aligned} \tag{23}$$

We define

$$(L^n,\varphi)=\left(f^n-\partial_t\bar{\partial}U_h^n,\varphi-\Pi_h v\right)-\left[a\left(U_h^{n,1/2},\varphi\right)-a\left(U_h^{n,1/2},\Pi_h v\right)\right]. \tag{24}$$

Then the term P_v on the interval $[t_{n-1},t_n]$ $(2\le n\le N)$ can be written as

$$(P_v,\varphi)=(L^n,\varphi)+a\left(U_h^{n,1/2}-U_\tau,\varphi\right)+(f-f^n,\varphi),\quad \forall\varphi\in H_0^1(\Omega),\ v\in\mathcal{U}_h. \tag{25}$$

When $t\in[0,t_1]$,

$$P_v(\cdot,t)=f(\cdot,t)+\nabla\cdot\left(a(x)\nabla\left(u_0+tv_0\right)\right). \tag{26}$$

3. Residual-Type A Posteriori Error Estimates

In this section, we will construct the residual-type a posteriori error estimates of the finite volume element method for (1). We introduce the jump of a vector-valued function across the edge $E\in\mathcal{E}_h$ which will be used in the residual-type a posteriori error estimates. Let E be an interior edge shared by elements K_+ and K_-. Define the unit normal vectors $\mathbf{n}_{K_+}$ and $\mathbf{n}_{K_-}$ on E pointing exterior to K_+ and K_-, respectively. Let $\mathbf{v}$ be a vector-valued function that is smooth inside each of the elements K_+ and K_-. $\mathbf{v}^+$ and $\mathbf{v}^-$ denote the traces of $\mathbf{v}$ on E taken from within the interior of K_+ and K_-, respectively. Then the jump of $\mathbf{v}$ on the edge E is defined by $[\mathbf{v}]_E=\mathbf{v}^+\cdot\mathbf{n}_{K_+}+\mathbf{v}^-\cdot\mathbf{n}_{K_-}$. We denote space refinement indicator by η_s^n defined by

$$\begin{aligned}\mathcal{R}_K^n&=f^n-\partial_t\bar{\partial}U_h^n+\nabla\cdot\left(a(x)\nabla U_h^{n,1/2}\right),\\ \mathcal{R}_E^n&=-\left[a(x)\nabla U_h^{n,1/2}\right]_E,\\ \eta_s^n&=\left(\sum_{K\in T_h}h_K^2\left\|\mathcal{R}_K^n\right\|_{0,K}^2+\sum_{E\in\mathcal{E}_h}h_E\left\|\mathcal{R}_E^n\right\|_{0,E}^2\right)^{1/2}.\end{aligned} \tag{27}$$

We define time refinement indicator η_t^n as

$$\eta_t^n=\tau_n\left\|U_h^n-U_h^{n-1}\right\|_1+\tau_n\left\|V_h^n-V_h^{n-1}\right\|. \tag{28}$$

3.1. Upper Bound. The Scott-Zhang interpolation function $\mathcal{I}_h:H_0^1(\Omega)\to\mathcal{U}_h$ is introduced in the following lemma [18].

Lemma 2. *For each $\varphi\in H_0^1(\Omega)$, a positive constant C is independent of h_K and h_E such that, for any $K\in T_h$, $E\in\mathcal{E}_h$*

$$\begin{aligned}\left\|\mathcal{I}_h\varphi\right\|_{1,\Omega}&\le C\|\varphi\|_{1,\Omega},\\ \left\|\varphi-\mathcal{I}_h\varphi\right\|_{0,K}&\le Ch_K\|\varphi\|_{1,\omega_K},\\ \left\|\varphi-\mathcal{I}_h\varphi\right\|_{0,E}&\le Ch_E^{1/2}\|\varphi\|_{1,\omega_E},\end{aligned} \tag{29}$$

where $\omega_K=\bigcup_{K'\cap K\ne\emptyset}K'$ and $\omega_E=\bigcup_{K\cap E\ne\emptyset}K$.

We also introduce the trace theorem [14].

Lemma 3 (trace theorem). *There exists a positive constant C independent of h_E such that*

$$\|\omega\|_{0,E}^2\le C\left(h_E^{-1}\|\omega\|_{0,K}^2+h_E\|\nabla\omega\|_{0,K}^2\right),\quad \forall\omega\in H^1(K),\ E\in\partial K,\ \forall K\in T_h. \tag{30}$$

Then we can get the following theorem for the upper bound of the error.

Theorem 4. *The following a posteriori error estimate holds between the solution u of (1) and the solution $(U_h^n)_{1\le n\le N}$ of (14), for $2\le m\le N$:*

$$\begin{aligned}&\left\|u^m-U_h^m\right\|+\left\|\sum_{k=1}^{k=m}\int_{t_{k-1}}^{t_k}(U_\tau-u)\,dt\right\|_1\\ &\le C\sum_{n=2}^{m}\left(\tau_n\left(\eta_t^n+\eta_s^n\right)\right)+C\sum_{n=2}^{m}\int_{t_{n-1}}^{t_n}\left\|f(\cdot,t)-f^n\right\|dt\\ &\quad+C\int_0^{t_1}\left\|f(\cdot,t)+\nabla\cdot\left(a(x)\nabla\left(u_0+tv_0\right)\right)\right\|dt.\end{aligned} \tag{31}$$

Proof. Taking the inner product of (18) with $\left(\begin{smallmatrix}u-U_\tau\\ \mathcal{T}(v-V_\tau)\end{smallmatrix}\right)$ and setting

$$Z(t)=\left(\left\|u-U_\tau\right\|^2+\left|v-V_\tau\right|_{-1}^2\right)^{1/2}, \tag{32}$$

we obtain, for $t\in[t_{n-1},t_n]$,

$$\begin{aligned}\frac{1}{2}\frac{dZ^2}{dt}&=\left(P_u,u-U_\tau\right)+\left(P_v,\mathcal{T}(v-V_\tau)\right)\le\|P_u\|\|u\\ &\quad-U_\tau\|+\left\|\nabla\cdot\left(a(x)\nabla\left(U_h^{n,1/2}-U_\tau\right)\right)\right\|_{-1}\\ &\quad\cdot\left\|\mathcal{T}(v-V_\tau)\right\|_1+\left\|L^n\right\|_{-1}\left\|\mathcal{T}(v-V_\tau)\right\|_1+\|f(\cdot,t)\\ &\quad-f^n\|\left\|\mathcal{T}(v-V_\tau)\right\|\le\|P_u\|\|u-U_\tau\|+C\left\|\nabla\right.\\ &\quad\cdot\left(a(x)\nabla\left(U_h^{n,1/2}-U_\tau\right)\right)\Big\|_{-1}\|v-V_\tau\|_{-1}+C\left\|L^n\right\|_{-1}\\ &\quad\cdot\|v-V_\tau\|_{-1}+C\|f(\cdot,t)-f^n\|\|v-V_\tau\|_{-1}\\ &\le C\left(\|P_u\|^2+\left\|L^n\right\|_{-1}^2+\|f(\cdot,t)-f^n\|^2\right.\\ &\quad\left.+\left\|\nabla\cdot\left(a(x)\nabla\left(U_h^{n,1/2}-U_\tau\right)\right)\right\|_{-1}^2\right)^{1/2}Z;\end{aligned} \tag{33}$$

hence,

$$\begin{aligned}\frac{dZ}{dt}&\le C\left(\|P_u\|^2+\left\|L^n\right\|_{-1}^2+\|f(\cdot,t)-f^n\|^2\right.\\ &\quad\left.+\left\|\nabla\cdot\left(a(x)\nabla\left(U_h^{n,1/2}-U_\tau\right)\right)\right\|_{-1}^2\right)^{1/2}\le C\left(\|P_u\|\right.\\ &\quad+\left\|L^n\right\|_{-1}+\|f(\cdot,t)-f^n\|\\ &\quad\left.+\left\|\nabla\cdot\left(a(x)\nabla\left(U_h^{n,1/2}-U_\tau\right)\right)\right\|_{-1}\right).\end{aligned} \tag{34}$$

Integrating the inequality from t_{n-1} to t_n $(2 \le n \le N)$, we have

$$\begin{aligned} Z(t_n) - Z(t_{n-1}) \le C \int_{t_{n-1}}^{t_n} \Big(& \|P_u\| + \|L^n\|_{-1} \\ & + \|f(\cdot,t) - f^n\| \\ & + \left\|\nabla \cdot \left(a(x) \nabla \left(U_h^{n,1/2} - U_\tau\right)\right)\right\|_{-1} \Big) dt. \end{aligned} \tag{35}$$

Using Lemma 1, we obtain

$$\begin{aligned} & \int_{t_{n-1}}^{t_n} \left\|\nabla \cdot \left(a(x) \nabla \left(U_h^{n,1/2} - U_\tau\right)\right)\right\|_{-1} dt \\ & = \int_{t_{n-1}}^{t_n} \left\|\nabla \cdot \left(a(x) \nabla \left(U_h^{n,1/2} - U_h^n + U_h^n - U_\tau\right)\right)\right\|_{-1} dt \\ & \le C \left\|U_h^n - U_h^{n-1}\right\|_1 \int_{t_{n-1}}^{t_n} \left(1 - \frac{t - t_{n-1}}{\tau_n}\right) dt \\ & \quad + C \frac{\tau_n}{2} \left\|U_h^{n+1} - U_h^n\right\|_1 + C \frac{\tau_n}{2} \left\|U_h^n - U_h^{n-1}\right\|_1 \\ & \le C\tau_n \left\|U_h^{n+1} - U_h^n\right\|_1 + C\tau_n \left\|U_h^n - U_h^{n-1}\right\|_1, \\ & \int_{t_{n-1}}^{t_n} \|P_u(\cdot,t)\| dt = \left\|V_h^n - V_h^{n-1}\right\| \int_{t_{n-1}}^{t_n} \frac{t - t_{n-1}}{\tau_n} dt \\ & = \frac{\tau_n}{2} \left\|V_h^n - V_h^{n-1}\right\|. \end{aligned} \tag{36}$$

By the definition of η_t^n, we get

$$\begin{aligned} & Z(t_n) - Z(t_{n-1}) \\ & \quad \le C \left(\tau_n \eta_t^n + \tau_n \|L^n\|_{-1} + \int_{t_{n-1}}^{t_n} \|f(\cdot,t) - f^n\| dt \right). \end{aligned} \tag{37}$$

In order to estimate $\|L^n\|_{-1}$, we choose $v = \mathcal{J}_h \varphi$ in (24); then

$$\begin{aligned} (L^n, \varphi) &= \left(f^n - \partial_t \bar{\partial} U_h^n, \varphi - \Pi_h v\right) \\ & \quad - \left[a\left(U_h^{n,1/2}, \varphi\right) - a\left(U_h^{n,1/2}, \Pi_h v\right)\right] \\ &= \left(f^n - \partial_t \bar{\partial} U_h^n, \varphi - v\right) \\ & \quad + \left(f^n - \partial_t \bar{\partial} U_h^n, v - \Pi_h v\right) \\ & \quad - \left[a\left(U_h^{n,1/2}, \varphi\right) - a\left(U_h^{n,1/2}, v\right)\right] \\ & \quad - \left[a\left(U_h^{n,1/2}, v\right) - a\left(U_h^{n,1/2}, \Pi_h v\right)\right] \\ &\triangleq \mathcal{J}_1 + \mathcal{J}_2 + \mathcal{J}_3 + \mathcal{J}_4. \end{aligned} \tag{38}$$

Using Green's formula, we have

$$\begin{aligned} \mathcal{J}_3 &= -\left(a(x) \nabla U_h^{n,1/2}, \nabla(\varphi - v)\right) \\ &= -\sum_{K \in T_h} \left(a(x) \nabla U_h^{n,1/2}, \nabla(\varphi - v)\right) \\ &= \sum_{K \in T_h} \left(\nabla \cdot \left(a(x) \nabla U_h^{n,1/2}\right), \varphi - v\right)_{0,K} \\ & \quad - \sum_{E \in \mathcal{E}_h} \left(\left[a(x) \nabla U_h^{n,1/2}\right]_E, \varphi - v\right)_{0,E}. \end{aligned} \tag{39}$$

By the definition of $\mathcal{R}_K^n, \mathcal{R}_E^n$, we get

$$\begin{aligned} & \mathcal{J}_1 + \mathcal{J}_3 \\ & = \sum_{K \in T_h} \left(f^n - \partial_t \bar{\partial} U_h^n + \nabla \cdot \left(a(x) \nabla U_h^{n,1/2}\right), \varphi - v\right)_{0,K} \\ & \quad - \sum_{E \in \mathcal{E}_h} \left(\left[a(x) \nabla U_h^{n,1/2}\right]_E, \varphi - v\right)_{0,E} \\ & = \sum_{K \in T_h} \left(\mathcal{R}_K^n, \varphi - v\right)_{0,K} + \sum_{E \in \mathcal{E}_h} \left(\mathcal{R}_E^n, \varphi - v\right)_{0,E}. \end{aligned} \tag{40}$$

From Cauchy-Schwarz inequality and Lemma 2, we can get

$$\begin{aligned} \left|\mathcal{J}_1 + \mathcal{J}_3\right| \le & C \sum_{K \in T_h} \left\{h_K \left\|\mathcal{R}_K^n\right\|_{0,K} \|\varphi\|_{1,\omega_K}\right\} \\ & + C \sum_{E \in \mathcal{E}_h} \left\{h_E^{1/2} \left\|\mathcal{R}_E^n\right\|_{0,E} \|\varphi\|_{1,\omega_E}\right\}. \end{aligned} \tag{41}$$

For $\mathcal{J}_4$, since $\Pi_h v$ is a constant in $K \bigcap K_z^*$, $z \in Z_h(K)$, $K_z^* \in T_h^*$, we have

$$\begin{aligned} & \int_K a(x) \nabla U_h^{n,1/2} \cdot \nabla v \, dx \\ & = \sum_{z \in Z_h(K)} \int_{K \bigcap K_z^*} a(x) \nabla U_h^{n,1/2} \cdot \nabla (v - \Pi_h v) \, dx \\ & = -\sum_{z \in Z_h(K)} \int_{K \bigcap K_z^*} \nabla \cdot \left(a(x) \nabla U_h^{n,1/2}\right) \cdot (v - \Pi_h v) \, dx \\ & \quad + \sum_{z \in Z_h(K)} \int_{\partial(K \bigcap K_z^*)} a(x) \nabla U_h^{n,1/2} \cdot \mathbf{n} (v - \Pi_h v) \, ds \\ & = -\int_K \nabla \cdot \left(a(x) \nabla U_h^{n,1/2}\right) \cdot (v - \Pi_h v) \, dx \\ & \quad + \int_{\partial K} a(x) \nabla U_h^{n,1/2} \cdot \mathbf{n} (v - \Pi_h v) \, ds \\ & \quad + \sum_{z \in Z_h(K)} \int_{K \bigcap \partial K_z^*} a(x) \nabla U_h^{n,1/2} \cdot \mathbf{n} (v - \Pi_h v) \, ds. \end{aligned} \tag{42}$$

Since $a(x)\nabla U_h^n$ and v are continuous inside each element $K \in T_h$, we have

$$\begin{aligned} &\sum_{z\in Z_h(K)} \int_{K\bigcap \partial K_z^*} a(x)\nabla U_h^n \cdot \mathbf{n} v\, ds = 0, \\ &\sum_{z\in Z_h(K)} \int_{K\bigcap \partial K_z^*} a(x)\nabla U_h^{n,1/2} \cdot \mathbf{n} v\, ds = 0. \end{aligned} \tag{43}$$

Thus,

$$\begin{aligned} \mathcal{J}_4 = &\sum_{K\in T_h} \left(\nabla\cdot\left(a(x)\nabla U_h^{n,1/2}\right), v-\Pi_h v\right)_{0,K} \\ &- \sum_{E\in\mathcal{E}_h} \left(\left[a(x)\nabla U_h^{n,1/2}\right]_E, v-\Pi_h v\right)_{0,E}. \end{aligned} \tag{44}$$

Then we get

$$\begin{aligned} \mathcal{J}_2 + \mathcal{J}_4 = &\sum_{K\in T_h} \left(f^n - \partial_t\bar{\partial} U_h^n + \nabla\cdot\left(a(x)\nabla U_h^{n,1/2}\right), v\right. \\ &\left.- \Pi_h v\right)_{0,K} - \sum_{E\in\mathcal{E}_h} \left(\left[a(x)\nabla U_h^{n,1/2}\right]_E, v-\Pi_h v\right)_{0,E} \\ = &\sum_{K\in T_h} \left(\mathcal{R}_K^n, v-\Pi_h v\right)_{0,K} + \sum_{E\in\mathcal{E}_h} \left(\mathcal{R}_E^n, v-\Pi_h v\right)_{0,E}. \end{aligned} \tag{45}$$

By (8) and Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} &\left|\sum_{K\in T_h} \left(\mathcal{R}_K^n, v-\Pi_h v\right)_{0,K}\right| \\ &\quad\le C\sum_{K\in T_h} \left\{h_K \left\|\mathcal{R}_K^n\right\|_{0,K} \|v\|_{1,K}\right\} \\ &\quad\le C\left(\sum_{K\in T_h} h_K^2 \left\|\mathcal{R}_K^n\right\|_{0,K}^2\right)^{1/2} \left(\sum_{K\in T_h} \|v\|_{1,K}^2\right)^{1/2} \\ &\quad= C\left(\sum_{K\in T_h} h_K^2 \left\|\mathcal{R}_K^n\right\|_{0,K}^2\right)^{1/2} \|\mathcal{J}_h\varphi\|_{1,\Omega} \\ &\quad\le C\left(\sum_{K\in T_h} h_K^2 \left\|\mathcal{R}_K^n\right\|_{0,K}^2\right)^{1/2} \|\varphi\|_1. \end{aligned} \tag{46}$$

$$\begin{aligned} &\left|\sum_{E\in\mathcal{E}_h} \left(\mathcal{R}_E^n, v-\Pi_h v\right)_{0,E}\right| \le \sum_{E\in\mathcal{E}_h} \left\|\mathcal{R}_E^n\right\|_{0,E} \|v-\Pi_h v\|_{0,E} \\ &\quad\le \left(\sum_{E\in\mathcal{E}_h} h_E \left\|\mathcal{R}_E^n\right\|_{0,E}^2\right)^{1/2} \\ &\quad\cdot \left(\sum_{E\in\mathcal{E}_h} h_E^{-1} \|v-\Pi_h v\|_{0,E}^2\right)^{1/2}. \end{aligned}$$

Since $\Pi_h v$ is a piecewise constant function, by Lemma 3 and (8), we get

$$\begin{aligned} &\sum_{E\in\mathcal{E}_h} h_E^{-1} \|v-\Pi_h v\|_{0,E}^2 \\ &\quad\le C\sum_{E\in\mathcal{E}_h} \left(h_E^{-2} \|v-\Pi_h v\|_{0,K}^2 + |v|_{1,K}^2\right) \le C\|v\|_1^2 \\ &\quad\le C\|\varphi\|_1^2. \end{aligned} \tag{47}$$

Substituting the estimate of $\mathcal{J}_1$–$\mathcal{J}_4$ into (38) and by the definition of η_s^n, we have

$$(L^n, \varphi) \le C\eta_s^n \|\varphi\|_1; \tag{48}$$

hence

$$\begin{aligned} &\frac{(L^n,\varphi)}{\|\varphi\|_1} \le C\eta_s^n, \\ &\|L^n\|_{-1} \le C\eta_s^n. \end{aligned} \tag{49}$$

Substituting the estimation of $\|L^n\|_{-1}$ into (37), we get

$$\begin{aligned} &Z(t_n) - Z(t_{n-1}) \\ &\quad\le C\left(\tau_n\eta_t^n + \tau_n\eta_s^n + \int_{t_{n-1}}^{t_n} \|f(\cdot,t) - f^n\|\, dt\right). \end{aligned} \tag{50}$$

Summing (50) from $n = 2$ to $n = m$, we obtain

$$\begin{aligned} Z(t_m) - Z(t_1) \le &C\sum_{n=2}^{m} \left(\tau_n\left(\eta_t^n + \eta_s^n\right)\right) \\ &+ C\sum_{n=2}^{m} \int_{t_{n-1}}^{t_n} \|f(\cdot,t) - f^n\|\, dt. \end{aligned} \tag{51}$$

For $n = 1$, we have

$$\begin{aligned} &Z(t_1) - Z(t_0) \\ &\quad\le C\int_0^{t_1} \|f(\cdot,t) + \nabla\cdot\left(a(x)\nabla\left(u_0 + tv_0\right)\right)\|\, dt. \end{aligned} \tag{52}$$

Noting that $Z(t_0) = Z(0) = 0$, then

$$\begin{aligned} &Z(t_m) \\ &\quad\le C\sum_{n=2}^{m} \left(\tau_n\left(\eta_t^n + \eta_s^n\right)\right) + C\sum_{n=2}^{m} \int_{t_{n-1}}^{t_n} \|f(\cdot,t) - f^n\|\, dt \\ &\qquad+ C\int_0^{t_1} \|f(\cdot,t) + \nabla\cdot\left(a(x)\nabla\left(u_0 + tv_0\right)\right)\|\, dt. \end{aligned} \tag{53}$$

By the fact that $(1/\sqrt{2})(a+b) \le \sqrt{a^2+b^2} \le a+b$ $(a, b > 0)$, we have

$$\begin{aligned} &\|u^m - U_h^m\| + \left|v^m - V_h^m\right|_{-1} \\ &\quad\le C\sum_{n=2}^{m} \left(\tau_n\left(\eta_t^n + \eta_s^n\right)\right) + C\sum_{n=2}^{m} \int_{t_{n-1}}^{t_n} \|f(\cdot,t) - f^n\|\, dt \\ &\qquad+ C\int_0^{t_1} \|f(\cdot,t) + \nabla\cdot\left(a(x)\nabla\left(u_0 + tv_0\right)\right)\|\, dt. \end{aligned} \tag{54}$$

In view of the definition of the operator $\mathcal{T}$, we have

$$\mathcal{T}\frac{\partial v}{\partial t}+u=\mathcal{T}f(\cdot,t), \tag{55}$$

$$\mathcal{T}\frac{\partial V_\tau}{\partial t}+U_h^{m,1/2}=\mathcal{T}f(\cdot,t^m),\quad t\in\left[t^{m-1},t^m\right]. \tag{56}$$

Subtracting (56) from (55), we get

$$\begin{aligned}&\mathcal{T}\frac{\partial(v-V_\tau)}{\partial t}\\&\quad+\left(u-U_h^{m,1/2}\right)=\mathcal{T}\left(f(\cdot,t)-f(\cdot,t^m)\right),\end{aligned} \tag{57}$$

$$\begin{aligned}&\mathcal{T}\frac{\partial(v-V_\tau)}{\partial t}+\mathcal{T}\left(f(\cdot,t^m)-f(\cdot,t)\right)\\&\quad+\left(U_\tau-U_h^{m,1/2}\right)=(U_\tau-u).\end{aligned} \tag{58}$$

Integrating (58) from t_{m-1} to t_m, we obtain

$$\begin{aligned}&\mathcal{T}\left(v^m-V_h^m\right)-\mathcal{T}\left(v^{m-1}-V_h^{m-1}\right)\\&\quad+\int_{t_{m-1}}^{t_m}\mathcal{T}\left(f(\cdot,t^m)-f(\cdot,t)\right)dt\\&\quad+\int_{t_{m-1}}^{t_m}\left(U_\tau-U_h^{m,1/2}\right)dt=\int_{t_{m-1}}^{t_m}(U_\tau-u)\,dt.\end{aligned} \tag{59}$$

Summing (59) from $k=1$ to $k=m$, we obtain

$$\begin{aligned}&\sum_{k=1}^{k=m}\int_{t_{k-1}}^{t_k}(U_\tau-u)\,dt\\&\quad=\mathcal{T}\left(v^m-V_h^m\right)\\&\qquad+\sum_{k=1}^{k=m}\int_{t_{k-1}}^{t_k}\mathcal{T}\left(f(\cdot,t^k)-f(\cdot,t)\right)dt\\&\qquad+\sum_{k=1}^{k=m}\int_{t_{k-1}}^{t_k}\left(U_\tau-U_h^{k,1/2}\right)dt.\end{aligned} \tag{60}$$

Thus, we have

$$\begin{aligned}&\left\|\sum_{k=1}^{k=m}\int_{t_{k-1}}^{t_k}(U_\tau-u)\,dt\right\|_1\\&\quad\le\left\|\mathcal{T}\left(v^m-V_h^m\right)\right\|_1\\&\qquad+\left\|\sum_{k=1}^{k=m}\int_{t_{k-1}}^{t_k}\mathcal{T}\left(f(\cdot,t^k)-f(\cdot,t)\right)dt\right\|_1\\&\qquad+\left\|\sum_{k=1}^{k=m}\int_{t_{k-1}}^{t_k}\left(U_\tau-U_h^{k,1/2}\right)dt\right\|_1\\&\quad\le C\left\|v^m-V_h^m\right\|_{-1}\\&\qquad+C\sum_{k=1}^{k=m}\int_{t_{k-1}}^{t_k}\left\|f(\cdot,t^k)-f(\cdot,t)\right\|_{-1}dt\\&\qquad+\sum_{k=1}^{k=m}\int_{t_{k-1}}^{t_k}\left(\left\|U_\tau-U_h^k\right\|_1+\left\|U_h^k-U_h^{k,1/2}\right\|_1\right)dt.\end{aligned} \tag{61}$$

Then,

$$\begin{aligned}&\left\|\sum_{k=1}^{k=m}\int_{t_{k-1}}^{t_k}(U_\tau-u)\,dt\right\|_1\\&\quad\le C\left\|v^m-V_h^m\right\|_{-1}\\&\qquad+C\sum_{k=1}^{k=m}\int_{t_{k-1}}^{t_k}\left\|f(\cdot,t^k)-f(\cdot,t)\right\|dt\\&\qquad+C\sum_{k=1}^{k=m}\tau_k\left\|U_h^k-U_h^{k-1}\right\|_1.\end{aligned} \tag{62}$$

By (62) and (54), we have

$$\begin{aligned}&\left\|u^m-U_h^m\right\|+\left\|\sum_{k=1}^{k=m}\int_{t_{k-1}}^{t_k}(U_\tau-u)\,dt\right\|_1\\&\quad\le C\sum_{n=2}^{m}\left(\tau_n\left(\eta_t^n+\eta_s^n\right)\right)+C\sum_{n=2}^{m}\int_{t_{n-1}}^{t_n}\left\|f(\cdot,t)-f^n\right\|dt\\&\qquad+C\int_0^{t_1}\left\|f(\cdot,t)+\nabla\cdot\left(a(x)\nabla\left(u_0+tv_0\right)\right)\right\|dt.\end{aligned} \tag{63}$$

□

3.2. Lower Bound. In order to derive the local lower bounds on the error, we will introduce some properties of the bubble functions. For each triangle $K\in T_h$, denote by $\lambda_{K,1},\lambda_{K,2},\lambda_{K,3}$ the barycentric coordinates. Define the element-bubble function ψ_K by

$$\begin{aligned}&\psi_K=27\lambda_{K,1}\lambda_{K,2}\lambda_{K,3},\quad\text{in } K;\\&\psi_K=0,\quad\text{in } \Omega\setminus K.\end{aligned} \tag{64}$$

Assume that K and K' share the edge $E\in\mathcal{E}_h$. Let the barycentric coordinates with respect to the end points of E be $\lambda_{E,1}$ and $\lambda_{E,2}$. Define the edge-bubble function ψ_E by

$$\begin{aligned}&\psi_E=4\lambda_{E,1}\lambda_{E,2},\quad\text{in } \omega_E=K\cup K';\\&\psi_E=0,\quad\text{in } \Omega\setminus\omega_E.\end{aligned} \tag{65}$$

For properties of the bubble functions, we have the following lemma [19].

Lemma 5. *For each of the elements $K \in T_h$ and $E \in \mathscr{E}_h$, functions ψ_K and ψ_E have the following properties:*

$$\begin{aligned}
&\operatorname{supp}\psi_K \subset K,\\
&\max_{x\in K}\psi_K = 1,\\
&\int_K \psi_K dx = \frac{9}{20}|K| \sim h_K^2,\\
&\|\nabla\psi_K\|_{0,K} \le Ch_K^{-1}\|\psi_K\|_{0,K}.\\
&\qquad\qquad \psi_K \in [0,1],\\
&\operatorname{supp}\psi_E \subset \omega_E,\\
&\max_{x\in\omega_E}\psi_E = 1,\\
&\int_E \psi_E ds = \frac{2}{3}h_E,\\
&\int_{\omega_E} \psi_E dx = \frac{1}{3}|\omega_E| \sim h_E^2,\\
&\|\nabla\psi_E\|_{0,\omega_E} \le Ch_E^{-1}\|\psi_E\|_{0,\omega_E},\\
&\qquad\qquad \psi_E \in [0,1].
\end{aligned} \tag{66}$$

We define the average of $\mathscr{R}_K^n$ on K ($\overline{\mathscr{R}_K^n}$) and the average of $\mathscr{R}_E^n$ on E ($\overline{\mathscr{R}_E^n}$) by

$$\begin{aligned}
\overline{\mathscr{R}_K^n} &= \frac{1}{|K|}\int_K \mathscr{R}_K^n dx,\\
\overline{\mathscr{R}_E^n} &= \frac{1}{h_E}\int_E \mathscr{R}_E^n ds.
\end{aligned} \tag{67}$$

Then we have the following local lower bounds.

Theorem 6. *For any $K \in T_h$, $E \in \mathscr{E}_h$, the following local posteriori lower bounds on the error $u^n - U_h^n$ hold for a positive constant C independent of h_K and h_E:*

$$\begin{aligned}
h_K\|\mathscr{R}_K^n\|_{0,K} \le C\Big(&\|u^n - U_h^n\|_{1,K} + h_K\|u_{tt}^n - \partial_t\bar{\partial}U_h^n\|_{0,K}\\
&+ \|U_h^{n+1} - U_h^n\|_{1,K} + \|U_h^n - U_h^{n-1}\|_{1,K}\\
&+ 2h_K\|\mathscr{R}_K^n - \overline{\mathscr{R}_K^n}\|_{0,K}\Big),
\end{aligned} \tag{68}$$

$$\begin{aligned}
h_E^{1/2}\|\mathscr{R}_E^n\|_{0,E} \le C\Big(&\|u^n - U_h^n\|_{1,\omega_E}\\
&+ h_E\|u_{tt}^n - \partial_t\bar{\partial}U_h^n\|_{0,\omega_E} + \|U_h^{n+1} - U_h^n\|_{1,\omega_E}\\
&+ \|U_h^n - U_h^{n-1}\|_{1,\omega_E} + h_E\|\mathscr{R}_K^n - \overline{\mathscr{R}_K^n}\|_{0,\omega_E}\\
&+ h_E^{1/2}\|\mathscr{R}_E^n - \overline{\mathscr{R}_E^n}\|_{0,E}\Big).
\end{aligned} \tag{69}$$

Proof. By triangle inequality, we have

$$\|\mathscr{R}_K^n\|_{0,K} \le \|\overline{\mathscr{R}_K^n}\|_{0,K} + \|\mathscr{R}_K^n - \overline{\mathscr{R}_K^n}\|_{0,K}. \tag{70}$$

By the properties of ψ_K, the definition of $\mathscr{R}_K^n$, and Green's formulation, we have

$$\begin{aligned}
\|\overline{\mathscr{R}_K^n}\|_{0,K}^2 &\sim \left(\overline{\mathscr{R}_K^n}, \psi_K\overline{\mathscr{R}_K^n}\right)\\
&= \left(\mathscr{R}_K^n, \psi_K\overline{\mathscr{R}_K^n}\right) - \left(\mathscr{R}_K^n - \overline{\mathscr{R}_K^n}, \psi_K\overline{\mathscr{R}_K^n}\right)\\
&= \left(f^n - \partial_t\bar{\partial}U_h^n + \nabla\cdot\left(a(x)\nabla U_h^{n,1/2}\right), \psi_K\overline{\mathscr{R}_K^n}\right)\\
&\quad - \left(\mathscr{R}_K^n - \overline{\mathscr{R}_K^n}, \psi_K\overline{\mathscr{R}_K^n}\right)\\
&= \left(f^n - u_{tt}^n, \psi_K\overline{\mathscr{R}_K^n}\right) + \left(u_{tt}^n - \partial_t\bar{\partial}U_h^n, \psi_K\overline{\mathscr{R}_K^n}\right)\\
&\quad + \left(\nabla\cdot\left(a(x)\nabla U_h^{n,1/2}\right), \psi_K\overline{\mathscr{R}_K^n}\right)\\
&\quad - \left(\mathscr{R}_K^n - \overline{\mathscr{R}_K^n}, \psi_K\overline{\mathscr{R}_K^n}\right)\\
&= a\left(u^n, \psi_K\overline{\mathscr{R}_K^n}\right)\\
&\quad - \int_K a(x)\nabla U_h^{n,1/2}\cdot\nabla\left(\psi_K\overline{\mathscr{R}_K^n}\right)dx\\
&\quad + \left(u_{tt}^n - \partial_t\bar{\partial}U_h^n, \psi_K\overline{\mathscr{R}_K^n}\right) - \left(\mathscr{R}_K^n - \overline{\mathscr{R}_K^n}, \psi_K\overline{\mathscr{R}_K^n}\right)\\
&= \int_K a(x)\nabla\left(u^n - U_h^{n,1/2}\right)\cdot\nabla\left(\psi_K\overline{\mathscr{R}_K^n}\right)dx\\
&\quad + \left(u_{tt}^n - \partial_t\bar{\partial}U_h^n, \psi_K\overline{\mathscr{R}_K^n}\right) - \left(\mathscr{R}_K^n - \overline{\mathscr{R}_K^n}, \psi_K\overline{\mathscr{R}_K^n}\right)\\
&\equiv \mathscr{P}_1 + \mathscr{P}_2 + \mathscr{P}_3.
\end{aligned} \tag{71}$$

For $\mathscr{P}_1$, with Cauchy-Schwarz inequality and Lemma 5, we get

$$\begin{aligned}
|\mathscr{P}_1| &\le C\left|u^n - U_h^{n,1/2}\right|_{1,K}\left\|\nabla\left(\psi_K\overline{\mathscr{R}_K^n}\right)\right\|_{0,K}\\
&= C\left|u^n - U_h^{n,1/2}\right|_{1,K}\|\nabla\psi_K\|_{0,K}\left|\overline{\mathscr{R}_K^n}\right|\\
&\le Ch_K^{-1}\left|u^n - U_h^{n,1/2}\right|_{1,K}\|\psi_K\|_{0,K}\left|\overline{\mathscr{R}_K^n}\right|\\
&= Ch_K^{-1}\left|u^n - U_h^{n,1/2}\right|_{1,K}\left\|\psi_K\overline{\mathscr{R}_K^n}\right\|_{0,K}\\
&\le Ch_K^{-1}\left|u^n - U_h^{n,1/2}\right|_{1,K}\left\|\overline{\mathscr{R}_K^n}\right\|_{0,K}.
\end{aligned} \tag{72}$$

By Cauchy-Schwarz inequality and Lemma 5, we obtain

$$\begin{aligned}
|\mathscr{P}_2| &\le \|u_{tt}^n - \partial_t\bar{\partial}U_h^n\|_{0,K}\left\|\psi_K\overline{\mathscr{R}_K^n}\right\|_{0,K}\\
&\le C\|u_{tt}^n - \partial_t\bar{\partial}U_h^n\|_{0,K}\left\|\overline{\mathscr{R}_K^n}\right\|_{0,K},\\
|\mathscr{P}_3| &\le \|\mathscr{R}_K^n - \overline{\mathscr{R}_K^n}\|_{0,K}\left\|\psi_K\overline{\mathscr{R}_K^n}\right\|_{0,K}\\
&\le \|\mathscr{R}_K^n - \overline{\mathscr{R}_K^n}\|_{0,K}\left\|\overline{\mathscr{R}_K^n}\right\|_{0,K}.
\end{aligned} \tag{73}$$

Combining (71)–(73), we obtain

$$\begin{aligned} h_K \left\|\mathscr{R}_K^n\right\|_{0,K} &\le C\Big(\left\|u^n - U_h^{n,1/2}\right\|_{1,K} \\ &\quad + h_K \left\|u_{tt}^n - \partial_t\bar{\partial}U_h^n\right\|_{0,K} + 2h_K \left\|\mathscr{R}_K^n - \overline{\mathscr{R}_K^n}\right\|_{0,K}\Big) \\ &\le C\Big(\left\|u^n - U_h^n\right\|_{1,K} + h_K \left\|u_{tt}^n - \partial_t\bar{\partial}U_h^n\right\|_{0,K} \\ &\quad + \left\|U_h^{n+1} - U_h^n\right\|_{1,K} + \left\|U_h^n - U_h^{n-1}\right\|_{1,K} \\ &\quad + 2h_K \left\|\mathscr{R}_K^n - \overline{\mathscr{R}_K^n}\right\|_{0,K}\Big). \end{aligned} \tag{74}$$

For (69), by triangle inequality, similarly we have

$$h_E^{1/2} \left\|\mathscr{R}_E^n\right\|_{0,E} \le h_E^{1/2} \left\|\overline{\mathscr{R}_E^n}\right\|_{0,E} + h_E^{1/2} \left\|\mathscr{R}_E^n - \overline{\mathscr{R}_E^n}\right\|_{0,E}. \tag{75}$$

By Lemma 5 and Green's formulation, we get

$$\begin{aligned} \left\|\overline{\mathscr{R}_E^n}\right\|_{0,E}^2 &\sim \left(\overline{\mathscr{R}_E^n}, \psi_E\overline{\mathscr{R}_E^n}\right)_{0,E} \\ &= \left(\mathscr{R}_E^n, \psi_E\overline{\mathscr{R}_E^n}\right)_{0,E} + \left(\overline{\mathscr{R}_E^n} - \mathscr{R}_E^n, \psi_E\overline{\mathscr{R}_E^n}\right)_{0,E} \\ &= \left(a(x)\nabla U_h^{n,1/2}, \nabla\left(\psi_E\overline{\mathscr{R}_E^n}\right)\right)_{0,\omega_E} \\ &\quad + \left(\nabla\cdot\left(a(x)\nabla U_h^{n,1/2}\right), \psi_E\overline{\mathscr{R}_E^n}\right)_{0,\omega_E} \\ &\quad + \left(\overline{\mathscr{R}_E^n} - \mathscr{R}_E^n, \psi_E\overline{\mathscr{R}_E^n}\right)_{0,E} \\ &= \int_{\omega_E} a(x)\nabla U_h^{n,1/2}\cdot\nabla\left(\psi_E\overline{\mathscr{R}_E^n}\right)dx \\ &\quad - \int_{\omega_E} a(x)\nabla u^n\cdot\nabla\left(\psi_E\overline{\mathscr{R}_E^n}\right)dx \\ &\quad + \int_{\omega_E} a(x)\nabla u^n\cdot\nabla\left(\psi_E\overline{\mathscr{R}_E^n}\right)dx \\ &\quad + \left(\nabla\cdot\left(a(x)\nabla U_h^{n,1/2}\right), \psi_E\overline{\mathscr{R}_E^n}\right)_{0,\omega_E} \\ &\quad + \left(\overline{\mathscr{R}_E^n} - \mathscr{R}_E^n, \psi_E\overline{\mathscr{R}_E^n}\right)_{0,E} \\ &= \int_{\omega_E} a(x)\nabla\left(U_h^{n,1/2} - u^n\right)\cdot\nabla\left(\psi_E\overline{\mathscr{R}_E^n}\right)dx \\ &\quad + \left(\nabla\cdot\left(a(x)\nabla U_h^{n,1/2}\right), \psi_E\overline{\mathscr{R}_E^n}\right)_{0,\omega_E} \\ &\quad + \int_{\omega_E}\left(f^n - \partial_t\bar{\partial}U_h^n\right)\left(\psi_E\overline{\mathscr{R}_E^n}\right)dx \\ &\quad + \int_{\omega_E}\left(\partial_t\bar{\partial}U_h^n - u_{tt}^n\right)\left(\psi_E\overline{\mathscr{R}_E^n}\right)dx \\ &\quad + \left(\overline{\mathscr{R}_E^n} - \mathscr{R}_E^n, \psi_E\overline{\mathscr{R}_E^n}\right)_{0,E} \\ &= \int_{\omega_E} a(x)\nabla\left(U_h^{n,1/2} - u^n\right)\cdot\nabla\left(\psi_E\overline{\mathscr{R}_E^n}\right)dx \\ &\quad + \left(\mathscr{R}_K^n, \psi_E\overline{\mathscr{R}_E^n}\right)_{0,\omega_E} \\ &\quad + \left(\partial_t\bar{\partial}U_h^n - u_{tt}^n, \psi_E\overline{\mathscr{R}_E^n}\right) \\ &\quad + \left(\overline{\mathscr{R}_E^n} - \mathscr{R}_E^n, \psi_E\overline{\mathscr{R}_E^n}\right)_{0,E} \\ &\equiv \mathscr{O}_1 + \mathscr{O}_2 + \mathscr{O}_3 + \mathscr{O}_4. \end{aligned} \tag{76}$$

Now we will estimate the right-hand terms of (76). By Lemma 5 and the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} |\mathscr{O}_1| &\le C\left|U_h^{n,1/2} - u^n\right|_{1,\omega_E}\left\|\nabla\left(\psi_E\overline{\mathscr{R}_E^n}\right)\right\|_{0,\omega_E} \\ &= C\left|U_h^{n,1/2} - u^n\right|_{1,\omega_E}\left\|\nabla\psi_E\right\|_{0,\omega_E}\left|\overline{\mathscr{R}_E^n}\right| \\ &\le Ch_E^{-1}\left|U_h^{n,1/2} - u^n\right|_{1,\omega_E}\left\|\psi_E\right\|_{0,\omega_E}\left|\overline{\mathscr{R}_E^n}\right| \\ &\le Ch_E^{-1/2}\left|U_h^{n,1/2} - u^n\right|_{1,\omega_E}\left\|\overline{\mathscr{R}_E^n}\right\|_{0,E}, \\ |\mathscr{O}_2| &\le \left\|\mathscr{R}_K^n\right\|_{0,\omega_E}\left\|\psi_E\overline{\mathscr{R}_E^n}\right\|_{0,\omega_E} \\ &= \left\|\mathscr{R}_K^n\right\|_{0,\omega_E}\left\|\psi_E\right\|_{0,\omega_E}\left|\overline{\mathscr{R}_E^n}\right| \le Ch_E\left\|\mathscr{R}_K^n\right\|_{0,\omega_E}\left|\overline{\mathscr{R}_E^n}\right| \\ &\le Ch_E^{1/2}\left\|\mathscr{R}_K^n\right\|_{0,\omega_E}\left\|\overline{\mathscr{R}_E^n}\right\|_{0,E}, \\ |\mathscr{O}_3| &\le \left\|\partial_t\bar{\partial}U_h^n - u_{tt}^n\right\|_{0,\omega_E}\left\|\psi_E\overline{\mathscr{R}_E^n}\right\|_{0,\omega_E} \\ &\le Ch_E^{1/2}\left\|u_{tt}^n - \partial_t\bar{\partial}U_h^n\right\|_{0,\omega_E}\left\|\overline{\mathscr{R}_E^n}\right\|_{0,E}, \\ |\mathscr{O}_4| &\le \left\|\overline{\mathscr{R}_E^n} - \mathscr{R}_E^n\right\|_{0,E}\left\|\psi_E\overline{\mathscr{R}_E^n}\right\|_{0,E} \\ &\le \left\|\overline{\mathscr{R}_E^n} - \mathscr{R}_E^n\right\|_{0,E}\left\|\overline{\mathscr{R}_E^n}\right\|_{0,E}. \end{aligned} \tag{77}$$

Combining (77) with (76), we get

$$\begin{aligned} \left\|\overline{\mathscr{R}_E^n}\right\|_{0,E} &\le Ch_E^{-1/2}\left|U_h^{n,1/2} - u^n\right|_{1,\omega_E} + Ch_E^{1/2}\left\|\mathscr{R}_K^n\right\|_{0,\omega_E} \\ &\quad + Ch_E^{1/2}\left\|u_{tt}^n - \partial_t\bar{\partial}U_h^n\right\|_{0,\omega_E} \\ &\quad + \left\|\overline{\mathscr{R}_E^n} - \mathscr{R}_E^n\right\|_{0,E}. \end{aligned} \tag{78}$$

With (74), we obtain

$$\begin{aligned} h_E^{1/2}\left\|\mathscr{R}_E^n\right\|_{0,E} &\le C\left\|u^n - U_h^{n,1/2}\right\|_{1,\omega_E} + Ch_E\left\|u_{tt}^n\right. \\ &\quad \left. - \partial_t\bar{\partial}U_h^n\right\|_{0,\omega_E} + Ch_E\left\|\mathscr{R}_K^n - \overline{\mathscr{R}_K^n}\right\|_{0,\omega_E} + h_E^{1/2}\left\|\mathscr{R}_E^n\right. \\ &\quad \left. - \overline{\mathscr{R}_E^n}\right\|_{0,E} \le C\Big(\left\|u^n - U_h^n + U_h^n - U_h^{n,1/2}\right\|_{1,\omega_E} \\ &\quad + h_E\left\|u_{tt}^n - \partial_t\bar{\partial}U_h^n\right\|_{0,\omega_E} + h_E\left\|\mathscr{R}_K^n - \overline{\mathscr{R}_K^n}\right\|_{0,\omega_E} \end{aligned}$$

Table 1: Error estimates for Case 1.

h	$\left\|u^N - U_h^N\right\|_0$	Rate	$\left\|u^N - U_h^N\right\|_1$	Rate	$\mathfrak{D}^N$	$\mathfrak{N}^N$	$\mathscr{R}$
$1/2^2$	$1.3839e-02$	—	$1.6047e-01$	—	0.4589	21.3382	46.4952
$1/2^3$	$4.0831e-03$	1.7610	$8.2097e-02$	0.9669	0.4417	21.1967	47.9920
$1/2^4$	$9.3149e-04$	2.1321	$4.1279e-02$	0.9919	0.4307	21.0618	48.9005
$1/2^5$	$2.0574e-04$	2.1787	$2.0670e-02$	0.9979	0.4247	20.9680	49.3734
$1/2^6$	$4.6977e-05$	2.1308	$1.0338e-02$	0.9996	0.4215	20.9138	49.6128

Table 2: Error estimates for Case 2.

h	$\left\|u^N - U_h^N\right\|_0$	Rate	$\left\|u^N - U_h^N\right\|_1$	Rate	$\mathfrak{D}^N$	$\mathfrak{N}^N$	$\mathscr{R}$
$1/2^2$	$1.2513e-01$	—	2.3532	—	6.6633	66.7292	10.0144
$1/2^3$	$3.2994e-02$	1.9232	1.1822	0.9931	6.3537	64.6603	10.1768
$1/2^4$	$8.1195e-03$	2.0227	$5.9245e-01$	0.9967	6.1806	63.5799	10.2870
$1/2^5$	$1.8443e-03$	2.1383	$2.9641e-01$	0.9991	6.0904	63.0179	10.3471
$1/2^6$	$4.0436e-04$	2.1894	$1.4822e-01$	0.9999	6.0444	62.7297	10.3782

$$
\begin{aligned}
&+ h_E^{1/2} \left\|\mathscr{R}_E^n - \overline{\mathscr{R}_E^n}\right\|_{0,E}\Big) \le C\Big(\left\|u^n - U_h^n\right\|_{1,\omega_E} \\
&+ h_E \left\|u_{tt}^n - \partial_t\bar{\partial}U_h^n\right\|_{0,\omega_E} + \left\|U_h^{n+1} - U_h^n\right\|_{1,\omega_E} \\
&+ \left\|U_h^n - U_h^{n-1}\right\|_{1,\omega_E} + h_E \left\|\mathscr{R}_K^n - \overline{\mathscr{R}_K^n}\right\|_{0,\omega_E} \\
&+ h_E^{1/2} \left\|\mathscr{R}_E^n - \overline{\mathscr{R}_E^n}\right\|_{0,E}\Big).
\end{aligned} \tag{79}
$$

□

4. Numerical Examples

Now we present some numerical examples to show the performance of the proposed error estimator. We consider problem (1) in $\Omega \times [0,T] = [0,1;0,1] \times [0,1]$. We discretize Ω into N number of rectangles in each direction and then each rectangle is divided into two triangles, resulting in a mesh with size $h = \sqrt{2}/N$. Discretize time by taking time step $\tau_n = \Delta t = h$. We consider the following two cases.

Case 1. Consider

$$
\begin{aligned}
a(x,y) &= 1 + \sin\left(\frac{\pi}{4}x\right) + \sin\left(\frac{\pi}{4}y\right) + e^{2x} + e^{2y}, \\
u(x,y,t) &= x(1-x)\,y(1-y)\,e^t.
\end{aligned} \tag{80}
$$

Case 2. Consider

$$
\begin{aligned}
a(x,y) &= e^{(x+y)/2}, \\
u(x,y,t) &= \sin(\pi x)\sin(\pi y)\,e^t.
\end{aligned} \tag{81}
$$

Define

$$
\begin{aligned}
\mathfrak{D}^m &= \sum_{n=2}^{m} \left\|u^n - U_h^n\right\|_1, \\
\mathfrak{N}^m &= \sum_{n=2}^{m} \left(\eta_t^n + \eta_s^n\right), \\
\mathscr{R} &= \frac{\mathfrak{N}^m}{\mathfrak{D}^m}.
\end{aligned} \tag{82}
$$

We present the results of the above cases when $m = N$ at Tables 1 and 2.

From Tables 1 and 2 we can see that the global a posteriori error estimator can predict the exact global error. The error estimator is reliable as evidenced by the ratio $\mathscr{R}$ listed on the tables. This list shows that the ratio $\mathscr{R}$ is converging to a constant when the mesh size is decreased by half. This shows that the proposed global a posteriori error estimator is robust for predicting the error in the finite volume element method.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is supported by National Natural Science Foundation of China (nos. 11301456, 11426193, and 11571297) and Shandong Province Natural Science Foundation (nos. ZR2014AP003 and ZR2014AM003).

References

[1] R. Li, Z. Chen, and W. Wu, *Generalized Difference Methods for Differential Equations Numerical Analysis of Finite Volume Methods*, Marcel Dekker, New York, NY, USA, 2000.

[2] S. Kumar, N. Nataraj, and A. K. Pani, "Finite volume element method for second order hyperbolic equations," *International Journal of Numerical Analysis and Modeling*, vol. 5, no. 1, pp. 132–151, 2008.

[3] I. Babuvška and W. C. Rheinboldt, "Error estimates for adaptive finite element computations," *SIAM Journal on Numerical Analysis*, vol. 15, no. 4, pp. 736–754, 1978.

[4] M. Afif, A. Bergam, Z. Mghazli, and R. Verfürth, "A posteriori estimators for the finite volume discretization of an elliptic problem," *Numerical Algorithms*, vol. 34, no. 2–4, pp. 127–136, 2003.

[5] C. Bi and V. Ginting, "A residual-type a posteriori error estimate of finite volume element method for a quasi-linear elliptic problem," *Numerische Mathematik*, vol. 114, no. 1, pp. 107–132, 2009.

[6] X. Ye, "A posterior error estimate for finite volume methods of the second order elliptic problem," *Numerical Methods for Partial Differential Equations*, vol. 27, no. 5, pp. 1165–1178, 2011.

[7] D. Kröner and M. Ohlberger, "A posteriori error estimates for upwind finite volume schemes for nonlinear conservation laws in multidimensions," *Mathematics of Computation*, vol. 69, no. 229, pp. 25–39, 2000.

[8] R. Lazarov and S. Tomov, "A posteriori error estimates for finite volume element approximations of convection-diffusion-reaction equations," *Computational Geosciences*, vol. 6, no. 3-4, pp. 483–503, 2002.

[9] M. Ohlberger, "A posteriori error estimate for finite volume approximations to singularly perturbed nonlinear convection-diffusion equations," *Numerische Mathematik*, vol. 87, no. 4, pp. 737–761, 2001.

[10] Z. Chen and J. Feng, "An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems," *Mathematics of Computation*, vol. 73, no. 247, pp. 1167–1193, 2004.

[11] W. Liu and N. Yan, "A posteriori error estimates for optimal control problems governed by parabolic equations," *Numerische Mathematik*, vol. 93, no. 3, pp. 497–521, 2003.

[12] P. Chatzipantelidis, C. Makridakis, and M. Plexousakis, "A posteriori error estimates for a finite volume method for the Stokes problem in two dimensions," *Applied Numerical Mathematics*, vol. 46, no. 1, pp. 45–58, 2003.

[13] C. Bernardi and E. Süli, "Time and space adaptivity for the second-order wave equation," *Mathematical Models and Methods in Applied Sciences*, vol. 15, no. 2, pp. 199–225, 2005.

[14] S. Brenner and L. Scott, *The Mathematical Theory of Finite Element Methods*, Springer, New York, NY, USA, 2008.

[15] R. E. Ewing, T. Lin, and Y. P. Lin, "On the accuracy of the finite volume element method based on piecewise linear polynomials," *SIAM Journal on Numerical Analysis*, vol. 39, no. 6, pp. 1865–1888, 2002.

[16] P. Chatzipantelidis, R. D. Lazarov, and V. Thomée, "Error estimates for a finite volume element method for parabolic equations in convex polygonal domains," *Numerical Methods for Partial Differential Equations*, vol. 20, pp. 650–674, 2004.

[17] V. Thomée, *Galerkin Finite Element Method for Parabolic Equations*, Springer, Berlin, Germany, 2006.

[18] L. R. Scott and S. Zhang, "Finite element interpolation of non-smooth functions satisfying boundary condition," *Mathematics of Computation*, vol. 54, no. 190, pp. 483–493, 1990.

[19] R. Verfürth, *A Review of a Posteriori Error Estimation and Adaptive Mesh-refinement Techniques*, John Wiley & Sons, New York, NY, USA, 1996.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 198371, 18 pages
http://dx.doi.org/10.1155/2015/198371

Research Article

Synchronization of Discrete-Time Chaotic Fuzzy Systems by means of Fuzzy Output Regulation Using Genetic Algorithm

Tonatiuh Hernández Cortés,[1] A. Verónica Curtidor López,[1] Jorge Rodríguez Valdez,[1] Jesús A. Meda Campaña,[1] Ricardo Tapia Herrera,[1] and José de Jesús Rubio[2]

[1] *Instituto Politécnico Nacional, SEPI-ESIME Zacatenco, Avenida IPN S/N, 07738 México, DF, Mexico*
[2] *Instituto Politécnico Nacional, SEPI-ESIME Azcapotzalco, Avenida de las Granjas No. 682, 02250 Azcapotzalco, México, DF, Mexico*

Correspondence should be addressed to Tonatiuh Hernández Cortés; tona_hernand@yahoo.com.mx

Received 7 August 2015; Accepted 4 November 2015

Academic Editor: Rongwei Guo

The synchronization of chaotic systems, described by discrete-time T-S fuzzy models, is treated by means of fuzzy output regulation theory. The conditions for designing a discrete-time output regulator are given in this paper. Besides, when the system does not fulfill the conditions for exact tracking, a new regulator based on genetic algorithms is considered. The genetic algorithms are used to approximate the adequate membership functions, which allow the adequate combination of local regulators. As a result, the tracking error is significantly reduced. Both the Complete Synchronization and the Generalized Synchronization problem are studied. Some numerical examples are used to illustrate the effectiveness of the proposed approach.

1. Introduction

A special nonlinear dynamical phenomenon, known as chaos, emerged in mid-1960s and reached applicable technology in the late 1990s and was considered as one of the three monumental discoveries of the twentieth century. On the other hand, fuzzy logic, a set theory and then an infinite-valued logic, gets a wide applicability in many industrial, commercial, and technical fields, ranging from control, automation, and artificial intelligence, just to name a few. Fuzzy logic and chaos had been considered by many researches and engineers as fundamental concepts and theories and their broad applicability in technology as well. The interaction between fuzzy logic and chaos has been developed for the last 20 years leading to research topics as fuzzy modeling of chaotic systems using Takagi-Sugeno models, linguistic descriptions of chaotic systems, fuzzy control of chaos, synchronization, and a combination of fuzzy chaos for engineering applications [1, 2].

In the 1960s, Rechenberg [3] introduced "evolution strategies," a method to optimize real-valued parameters for devices such as airfoils. This idea was further developed by Schwefel in [4]. Genetic algorithms (GAs) were initially developed by Bremermann [5] in 1958 but popularized and developed by Holland in the 1960s. In contrast with evolution strategies and evolutionary programming, Holland's idea was not to design algorithms to solve specific problems but rather to formally study the phenomenon of adaptation, as it occurs in nature, and develop ways in which the mechanisms of natural adaptation might be transferred into computer systems [6]. The genetic algorithm is presented as abstraction of biological evolution and theoretical framework for adaptation for moving from one population of "chromosomes" (e.g., strings of ones and zeros, or "bits") to a new population by using a kind of "natural selection" together with the genetics-inspired operators of crossover, mutation, and inversion. Each chromosome consists of "genes" (e.g., bits); each gene is being an instance of a particular "allele" (e.g., 0 or 1). The operator selection chooses those chromosomes in the population that will be allowed to reproduce and those adjusted chromosomes produce more offspring than the less ones [7].

According to Fogel and Aderson [8], Bremermann was the first to implement real-coded genetic algorithms as well as providing a mathematic model of GA known as the one-max function. In contrast to genetic algorithms, Evolutionary Strategies were initially developed for the purpose of parameter optimization. The idea was to imitate the principles of organic evolution in experimental parameter optimization for applications such as pipe bending or PID control for a nonlinear system [9].

Synchronization of chaotic systems is one of the more exiting problems in control science and can be referred at least to Huygen's observations [10]; it is understood as one of the trajectories of two autonomous chaotic systems, starting from nearly initial conditions and converging to the other, and remains as $t \to \infty$; in [11] it was reported that some kind of chaotic systems possesses a self-synchronization property. However, not all chaotic systems can be decomposed in two separate responses subsystems and be able to synchronize the drive system. The ideas of these works have led to improvement in many fields, such as communications [12], encrypted systems, the complex information processing within the human brain, coupled biochemical reactors, and earthquake engineering [13].

Synchronization can be classified as follows: Complete Synchronization: it is when two identical chaos oscillators are mutually coupled and one drives to the other; Generalized Synchronization: it differs from the previous case by the fact that there are different chaos oscillators and the states of one are completely defined by the other; Phase Synchronization: it occurs when the coupled oscillators are not identical and have different amplitude that is still unsynchronized, while the phases of oscillators evolve to be synchronized [14]. It is worth mentioning that studies in synchronization of nonlinear systems have been reformulated based on the previous results from classical control theory such as [15–18].

In this paper, the fuzzy output regulation theory and Takagi-Sugeno (T-S) fuzzy models are used to solve the Complete and Generalized Synchronization by using linear local regulators. Isidori and Byrnes [19] showed that the output regulation established by Francis could be extended for a nonlinear sector as a general case, resulting in a set of nonlinear partial differential equations called Francis-Isidori-Byrnes (FIB). Unfortunately these equations in many cases are too difficult to solve in a practical manner. For this reason in [20] the approach based on the weighted summation of local linear regulators is presented and in [21] the new membership functions in the regulator are approximated by soft computing techniques.

So, the main contribution of the present work is to find a control law for synchronizing of chaotic systems described by discrete-time Takagi-Sugeno fuzzy models, first when the system fulfills the following: (1) the input matrix for all subsystems is the same and (2) the local regulators share the same zero error manifold $\pi(w_k)$. In this way, the results given in [20] are extended to the discrete-time domain. On the other hand, when the system *master-slave* does not fulfill the aforementioned conditions, new membership functions are computed in order to enhance the performance of the fuzzy regulator. Such proposed membership functions are different from those given in the plant or exosystem and are tuned by using the GA. The tuning of the new membership functions, which is as generalized bell-shaped function, is given by optimization of the form parameter.

The rest of the paper is organized as follows. In Section 2 the discrete-time output regulation problem formulation is given with a brief review of the Takagi-Sugeno models and the discrete-time fuzzy regulation problem. In Section 3 the tuning of membership functions by means of GAs is thoroughly discussed. In Section 4 Complete and Generalized Synchronization with some examples are presented and finally, in Section 5, some conclusions are drawn.

2. The Discrete-Time Fuzzy Output Regulation Problem

Consider a nonlinear discrete-time system defined by

$$x_{k+1} = f(x_k, \omega_k, u_k), \tag{1}$$

$$y_k = c(x_k), \tag{2}$$

$$\omega_{k+1} = s(\omega_k), \tag{3}$$

$$y_{\mathrm{ref},k} = q(\omega_k), \tag{4}$$

$$e_k = h(x_k, \omega_k), \tag{5}$$

where $x_k \in \mathbb{R}^n$ is the state vector of the plant, $w_k \in W \subset \mathbb{R}^s$ is the state vector of the exosystem, which generates the reference and/or the perturbation signals, and $u_k \in \mathbb{R}^m$ is the input signal. Equation (5) refers to difference between output system of the plant ($y_k \in \mathbb{R}^m$) and the reference signal ($y_{\mathrm{ref},k} \in \mathbb{R}^m$); that is, $h(x_k, \omega_k) = y_k - y_{\mathrm{ref},k} = c(x_k) - q(w_k)$ and take into account that $m \leq n$. Besides, it is assumed that $f(x_k, u_k, w_k)$, $h(x_k, w_k)$, and $s(w_k)$ are analytical functions and also that $f(0,0,0) = 0$, $s(0) = 0$, and $h(0,0) = 0$ [22].

Clearly, by linearizing (1)–(5) around $x_k = 0$, one gets

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + Pw_k, \\ y_k &= Cx_k, \\ w_{k+1} &= Sw_k, \\ y_{\mathrm{ref},k} &= Qw_k, \\ e_k &= Cx_k - Qw_k. \end{aligned} \tag{6}$$

Thus, the Nonlinear Regulator Problem [19, 23] consists of finding a controller $u_k = \alpha(x_k, w_k)$, such that the closed-loop system $x_{k+1} = Ax_k + B\alpha(x_k, 0)$ has an asymptotically stable equilibrium point, and the solution of system (6) satisfies $\lim_{k\to\infty} e_k = 0$.

So, by defining $\pi(w_k)$ as the steady-state zero error manifold and $\gamma(w_k)$ as the steady-state input, the following theorem gives the conditions for the solution of nonlinear regulation problem.

Theorem 1. *Suppose that $w_{k+1} = s(w_k)$ is Poisson stable and there exists a gain K such that the matrix $A + BK$ is stable and*

there exist mappings $x_{ss}(t) = \pi(w_k)$ *and* $u_{ss} = \gamma(w_k)$ *with* $\pi(0) = 0$ *and* $\gamma(0) = 0$ *satisfying*

$$\begin{aligned} \pi(s(w_k)) &= f(\pi(w_k), w_k, \gamma(w_k)), \\ 0 &= h(\pi(w_k), w_k). \end{aligned} \tag{7}$$

Then the control signal for the nonlinear regulation is given by

$$u_k = K(x_k - \pi(w_k)) + \gamma(w_k). \tag{8}$$

Proof. See [22–24]. □

The equation set (7) is known as Discrete-Time Francis-Isidori-Byrnes (DTFIB) equations and linear counterpart is obtained when the mappings $x_{ss,k} = \pi(w_k)$ and $u_{ss,k} = \gamma(w_k)$ transform into $x_{ss,k} = \Pi w_k$ and $u_{ss,k} = \Gamma w_k$, respectively. Thus, the problem is reduced to solve linear matrix equations [25] given by

$$\begin{aligned} \Pi S &= A\Pi + B\Gamma + P, \\ 0 &= C\Pi - Q. \end{aligned} \tag{9}$$

2.1. The Discrete-Time Output Fuzzy Regulation Problem. Takagi and Sugeno proposed a fuzzy model composed of a set of linear subsystems with IF-THEN rules capable of relating physical knowledge, linguistic characteristics, and properties of the system. Such a model successfully represents a nonlinear system at least in a predefined region of phase space [15]. The T-S model for the plant and exosystem can be described as follows [26]:

Plant Model

Rule i:

IF $z_{1,1,k}$ is $M_{1,1}^i$ and ... and $z_{1,p_1,k}$ is M_{1,p_1}^i,

$$\textbf{THEN} \begin{cases} x_{k+1} = A_i x_k + B_i u_k + P_i w_k, \\ y_k = C_i x_k, \end{cases} \quad i = 1, 2, \ldots, r_1, \tag{10}$$

where r_1 is the number of rules in the model of the plant and the sets $M_{1,j}^i$ are the fuzzy sets defined based on the previous dynamic knowledge of the system.

Exosystem Model

Rule i:

IF $z_{2,1,k}$ is $M_{2,1}^i$ and ... and $z_{2,p_2,k}$ is M_{2,p_2}^i,

$$\textbf{THEN} \begin{cases} w_{k+1} = S_i w_k, \\ y_{\text{ref},k} = Q_i w_k, \end{cases} \quad i = 1, 2, \ldots, r_2, \tag{11}$$

where r_2 is the number of rules in the model of the exosystem and $M_{2,j}^i$ are the fuzzy sets.

Then, the regulation problem defined by (1)–(5) can be represented through the T-S discrete-time fuzzy model; that is, [20]

$$\begin{aligned} x_{k+1} &= \sum_{i=1}^{r_1} h_{1,i}(z_{1,k}) \{A_i x_k + B_i u_k + P_i w_k\}, \\ w_{k+1} &= \sum_{i=1}^{r_2} h_{2,i}(z_{2,k}) S_i w_k, \\ e_k &= \sum_{i=1}^{r_1} h_{1,i}(z_{1,k}) C_i x_k - \sum_{i=1}^{r_2} h_{2,i}(z_{2,k}) Q_i w_k, \end{aligned} \tag{12}$$

where $x_k \in \mathbb{R}^n$ is the state vector of the plant, $w_k \in \mathbb{R}^s$ is the state vector of the exosystem, $u_k \in \mathbb{R}^m$ is the input signal, $e_k \in \mathbb{R}^m$, and $h_{*,i}(z)$ is the normalized weight of each rule, 1 for the plant and 2 for the exosystem, which depends on the membership function for the premise variable $z_{*,k}$ in $M_{*,j}^i$; that is,

$$\begin{aligned} \overline{\omega}_{*,i}(z_{*,k}) &= \prod_{j=1}^{p_*} M_{*,j}^i(z_{*,j,k}), \\ h_{*,i}(z_{*,k}) &= \frac{\overline{\omega}_{*,i}(z_{*,k})}{\sum_{i=1}^{r_*} \overline{\omega}_{*,i}(z_{*,k})}, \\ \sum_{i=1}^{r_*} h_{*,i}(z_{*,k}) &= 1, \\ h_{*,i}(z_{*,k}) &\geq 0 \end{aligned} \tag{13}$$

with $z_{*,k} = [z_{*,1,k} \;\; z_{*,2,k} \;\; \cdots \;\; z_{*,p_*,k}]$ as a function of x_k and/or w_k, $i = 1, \ldots, r_*$ and $j = 1, \ldots, p$.

The *discrete-time fuzzy output regulation problem* consists of finding a controller $u_k = \alpha(x_k, w_k)$, such that the closed-loop system with no external signal

$$x_{k+1} = \sum_{i=1}^{r_1} h_{1,i}(z_{1,k}) \{A_i x_k + B_i \alpha(x_k, 0)\} \tag{14}$$

has an asymptotically stable equilibrium point.

The solution of system (12) satisfies

$$\lim_{k \to \infty} e_k = 0. \tag{15}$$

In order to achieve the synchronization of chaotic systems described by a T-S discrete-time fuzzy model it is necessary to fulfill (7) [27, 28]. Then

$$\begin{aligned} &\pi(s(w_k)) \\ &\quad = \sum_{i=1}^{r_1} h_{1,i}(z_{1,k}) \{A_i \pi(w_k) + B_i \gamma(w_k) + P_i w_k\}, \end{aligned} \tag{16}$$

$$0 = \sum_{i=1}^{r_1} h_{1,i}(z_{1,k}) C_i \pi(w_k) - \sum_{i=1}^{r_2} h_{2,i}(z_{2,k}) Q_i w_k, \tag{17}$$

where $\pi(w_k)$ is the zero error steady-state manifold which becomes invariant by the effect of the steady-state input $\gamma(w_k)$.

Assuming the mappings $\pi(w_k)$ and $\gamma(w_k)$ as

$$\tilde{\pi}(w_k) = \sum_{i=1}^{r_1} h_{1,i}(z_{1,k}) \sum_{j=1}^{r_2} h_{2,i}(z_{2,k}) \Pi_{i,j} w_k, \tag{18}$$

$$\tilde{\gamma}(w_k) = \sum_{i=1}^{r_1} h_{1,i}(z_{1,k}) \sum_{j=1}^{r_2} h_{2,i}(z_{2,k}) \Gamma_{i,j} w_k, \tag{19}$$

respectively, with $\Pi_{i,j}$ and $\Gamma_{i,j}$ as a solution of $r_1 \cdot r_2$ lineal local problems,

$$\Pi_{ij} S_j = A_i \Pi_{ij} + B_i \Gamma_{ij} + P_i, \tag{20}$$

$$0 = C_i \Pi_{i,j} - Q_i, \tag{21}$$

for all $i = 1, \ldots, r_1$ and $j = 1, \ldots, r_2$, the following control law can be obtained [20, 22, 23]:

$$\begin{aligned} u_k = &\sum_{h=1}^{r_1} h_{1,i}(z_{1,k}) \\ &\cdot K_i \left[x_k - \sum_{i=1}^{r_1} h_{1,i}(z_{1,k}) \sum_{j=1}^{r_2} h_{2,j}(z_{2,k}) \Pi_{ij} w_k \right] \\ &+ \sum_{i=1}^{r_1} h_{1,i}(z_{1,k}) \sum_{j=1}^{r_2} h_{2,j}(z_{2,k}) \Gamma_{ij} w_k. \end{aligned} \tag{22}$$

However, by substitution of $\pi(w_k)$ and $\gamma(w_k)$ in (16) and (17) and considering

(1) the steady-state zero error manifold $\pi(w_k) = \Pi w_k$, that is, $\Pi_{ij} = \Pi$,

(2) the input matrices $B_i = B$ and/or $\Gamma_{ij} = \Gamma$,

for all $i = 1, \ldots, r_1$ and $j = 1, \ldots, r_2$, the following control signal u_k emerges:

$$\begin{aligned} u_k = &\sum_{h=1}^{r_1} h_{1,i}(z_{1,k}) K_i [x_k - \Pi w_k] \\ &+ \sum_{i=1}^{r_1} h_{1,i}(z_{1,k}) \sum_{j=1}^{r_2} h_{2,j}(z_{2,k}) \Gamma_{ij} w_k. \end{aligned} \tag{23}$$

On the other hand, the existence of a fuzzy stabilizer of the form $u = \sum_{i=1}^{r_1} h_{1,i}(z_{1,k}) K_i x_k$, ensuring that the tracking error converges asymptotically to zero, can be obtained from the Parallel Distributed Compensator (PDC) [29, 30] or another stability analysis for T-S fuzzy models such as [31].

Remark 2. The control signal in (22) is given by the substitution of (18) and (19) in (16) and (17); the proposed controller provides the following advantages:

(1) All parameters included in the controller are known; this includes the membership functions of the plant and exosystem, which are well defined in the T-S fuzzy model.

On the other hand, Π_{ij} and Γ_{ij} come directly from solving of $r_1 * r_2$ local linear problems equivalent to solving the Francis equations; such problems can be easily solved by using programs like Matlab or Mathematica.

(2) In the case when the $r_1 * r_2$ local linear problems lead to $\Pi_{ij} \neq \Pi$, then at least a bounded error is ensured.

(3) It is clear that when $\Pi_{ij} = \Pi$, the term $\sum_{i=1}^{r_1} h_{1,i}(z_{1,k}) \sum_{j=1}^{r_2} h_{2,j}(z_{2,k}) \Pi_{ij} w_k$ changes to Πw_k leading to controller defined in (23).

(4) The following condition: the input matrices $B_i = B$ and/or $\Gamma_{ij} = \Gamma$ avoids the crossed terms in the solutions of (16) and (17), allowing the exact fuzzy output regulation.

(5) The proposed controller can be seen as a simple substitution of the aforementioned elements.

On the other hand, the following disadvantages can appear:

(1) If the condition is that the steady-state zero error manifold $\pi(w_k) \neq \Pi w_k$, that is, $\Pi_{ij} \neq \Pi$, then, it will be necessary to adjust the local regulator by means of new membership functions. Please refer to Section 3.

(2) As expected, the complexity of the controller increases according to the number of local subsystems.

The following theorem provides the conditions for the existence of the exact fuzzy output regulator for a discrete-time T-S fuzzy models.

Theorem 3. *The exact fuzzy output regulation with full information of systems defined as (12) is solvable if (a) there exists the same zero error steady-state manifold $\pi(w_k) = \Pi$; (b) there exist $u_k = \sum_{i=1}^{r_1} h_{i,k}(z_{1,k}) K_i x_k$ for the fuzzy system; (c) the exosystem $\omega_{k+1} = s(\omega_k)$ is Poisson stable, and the input matrices for all subsystems B_i are equal. Moreover, the Exact Output Fuzzy Regulation Problem is solvable by the controller*

$$\begin{aligned} u_k = &\sum_{h=1}^{r_1} h_{1,i}(z_{1,k}) K_i [x_k - \Pi w_k] \\ &+ \sum_{i=1}^{r_1} h_{1,i}(z_{1,k}) \sum_{j=1}^{r_2} h_{2,j}(z_{2,k}) \Gamma_{ij} w_k. \end{aligned} \tag{24}$$

Proof. From the previous analysis, the existence of mappings $\pi(w_k) = \Pi w_k$ and $\gamma(w_k) = \sum_{i=1}^{r_1} h_{1,i}(z_{1,k}) \sum_{j=1}^{r_2} h_{2,j}(z_{2,k}) \Gamma_{ij} w_k$ is guaranteed when the input matrices for all subsystems B_i are equal, and the solution of $r_1 \cdot r_2$ lineal local problems is

$$\begin{aligned} \Pi_{ij} S_j &= A_i \Pi_{ij} + B_i \Gamma_{ij} + P_i, \\ 0 &= C_i \Pi_{i,j} - Q_i, \end{aligned} \tag{25}$$

leading to $\Pi_{ij} = \Pi$ for all $i = 1, \ldots, r_1$ and $j = 1, \ldots, r_2$.

On the other hand, the inclusion of condition (b) has been thoroughly discussed in [22, 23, 27, 28, 31–33], and it implies the existence of a fuzzy stabilizer.

Condition (c) ensures the nonexistence of crossed terms in the local Francis equations. Finally, condition (d) is introduced to avoid the fact that the reference signal converges to zero, which would turn the regulation problem into a simple stability problem. The rest of the proof follows directly from the previous analysis. □

3. The Output Regulator by means of Local Regulators and Tuning of New Membership Functions

In this section, a discrete-time T-S fuzzy model is considered to solve the exact output regulation on the basis of linear local controllers. So, the main goal is to find a complete regulator based on the fuzzy summation of local regulators using adequate membership functions, such that the result given in Theorem 3 can be relaxed [34]. These membership functions are not necessarily the same included in the fuzzy plant, as is described in (19). Thus, the steady-state input $\gamma(w_k)$ can be defined as

$$\gamma(w_k) = \sum_{i=1}^{r_1} \mu_{1,i}(z_{1,k}) \sum_{j=1}^{r_2} \mu_{2,j}(z_{2,k}) \Gamma_{ij} w_k, \tag{26}$$

where $\mu_{1,i}(z_{1,k})$ and $\mu_{2,j}(z_{2,k})$ are new membership functions, such that the fuzzy output regulator obtained from local regulators provides the exact fuzzy output regulation. This approach requires the computation of the linear local controllers and the computation of the new membership functions. In this work, such functions are represented by the following expression:

$$\mu_{1,i}(x_k) = \frac{1}{1 + \left|(x_k - c_i)/a_i\right|^{2b_i}}, \quad \forall i = 1, \ldots, r_1, \tag{27}$$

$$\mu_{2,j}(w_k) = \frac{1}{1 + \left|(w_k - c_j)/a_j\right|^{2b_j}}, \quad \forall j = 1, \ldots, r_2. \tag{28}$$

$\mu_{1,i}$ and $\mu_{2,j}$ are well known as *generalized bell-shaped membership functions* and the parameters a_i, a_j, b_i, b_j, c_i, and c_j determine the form, center, and amplitude, respectively. Therefore, from (22), the input can be defined by

$$\begin{aligned} u_k = &\sum_{i=1}^{r_1} h_{1,i}(x_k) k_i \{x_k - \pi(w_k)\} \\ &+ \sum_{i=1}^{r_1} \mu_{1,i}(w_k) \sum_{j=1}^{r_2} \mu_{2,j}(w_k) \Gamma_{ij} w_k, \end{aligned} \tag{29}$$

because $z_{*,k}$ is a function of x_k and in steady-state $x_k = \pi(w_k)$.

Then, for tuning the membership functions (27) and (28), the parameters a_i and a_j will be optimized by means of genetic algorithms, ensuring the correct interpolation between the local linear regulators. The foregoing can be summarized in the control scheme depicted in Figure 1.

To this end, the following algorithm should be carried out.

Algorithm 4. Main steps to take into account to solve the tuning membership functions problem are as follows [35]:

(1) Start with a randomly generated population of nl-bit chromosomes (candidate solutions to a problem). The traditional representation is binary as in Figure 2.

A binary string is called "chromosome." Each position therein is called "gene" and the value in this position is named "allele."

(2) Calculate the fitness $f(x)$ of each chromosome x in the population.

(3) Repeat the following steps until n offspring have been created:

(i) Select a pair of parent chromosomes from the current population, the probability of selection being an increasing function of fitness. Selection is done "with replacement," meaning that the same chromosome can be selected more than once to become a parent.

(ii) With probability p_c (the "crossover probability" or "crossover rate"), crossover the pair at a randomly chosen point (chosen with uniform probability) to form two offspring. If no crossover takes place, form two offspring that are exact copies of their respective parents. (Note that here the crossover rate is defined to be the probability that two parents will cross over in a single point.)

(iii) Mutate the two offspring at each locus with probability p_m (the mutation probability or mutation rate), and place the resulting chromosomes in the new population.

(4) Replace the current population with the new population.

(5) Go to step (2).

This process ends when the fitness function reaches a value less than or equal to a predefined bound, or when the maximum number of iterations is reached. Each iteration of this process is called a generation.

In order to apply a genetic algorithm it requires the following five basic components:

(i) A representation of the potential solutions to the problem.

(ii) One way to create an initial population of possible solutions (typically a random process).

(iii) An evaluation function to play the role of the environment, classifying the solutions in terms of its "fitness."

(iv) Genetic operators that alter the composition of the chromosomes that will be produced for generations.

(v) Values for the various parameters used by the genetic algorithm (population size, crossover probability, mutation probability, maximum number of generations, etc.).

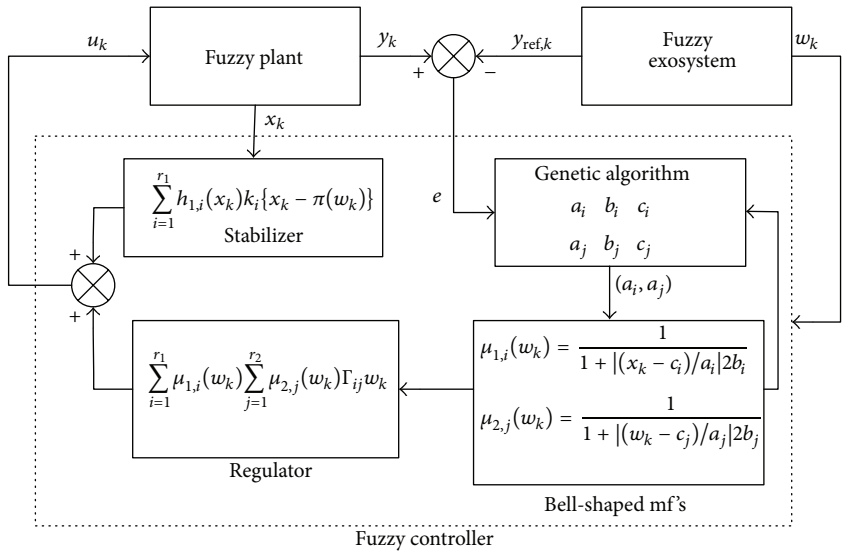


Figure 1: Control scheme for fuzzy output regulation and genetic algorithms.

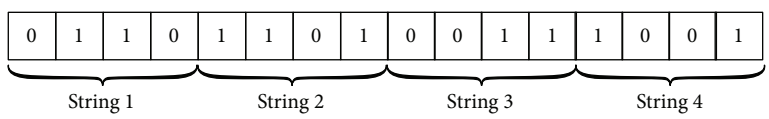


Figure 2: Binary string commonly used in genetic algorithms.

The aforementioned can be summarized by the flowchart depicted in Figure 3.

4. Synchronization of T-S Discrete-Time Fuzzy Systems

The stabilization and synchronization of chaotic systems are two of the most challenging and stimulating problems due to their capabilities of describing a great variety of very interesting phenomena in physics, biology, chemistry, and engineering, to name a few. In this section, the regulation theory is used to synchronize chaotic systems described by T-S discrete-time fuzzy models. Both the drive system and the response system are modeled by the same attractor (Rössler's equation) with the difference that response system can be influenced by an input signal. This type of synchronization is known as *Complete Synchronization* (CS) [36].

Considering two Rössler chaotic oscillators as $\dot{w} = f(w)$ (drive system) and $\dot{x} = f(x, w, u)$ (response system), the ordinary differential equations of these systems are

$$\begin{gathered}\dot{w}_1 = -(w_2 + w_3),\\ \dot{w}_2 = w_1 + aw_2,\\ \dot{w}_3 = bw_1 - (c - w_1)w_3,\\ \text{Rössler attractor}\\ \text{Drive system,}\\ \dot{x}_1 = -(x_2 + x_3),\\ \dot{x}_2 = x_1 + ax_2,\\ \dot{x}_3 = bx_1 - (c - x_1)x_3 + u,\\ \text{Rössler attractor}\\ \text{Response system.}\end{gathered} \tag{30}$$

According to [15], these systems can be exactly represented by means of the following two-rule T-S fuzzy models when a, b and c are constants and $x_1 \in [c - d, c + d]$ with $d > 0$. Thus

$$\begin{aligned}\dot{x}(t) &= \sum_{i=1}^{2} h_{1,i}(x_1(t))\left\{\widetilde{A}_i x(t) + \widetilde{B}_i u(t)\right\},\\ \dot{w}(t) &= \sum_{i=1}^{2} h_{2,i}(w_1(t))\widetilde{S}_i w(t),\\ e(t) &= \sum_{i=1}^{2} h_{1,i}(x_1(t))\widetilde{C}_i x(t)\\ &\quad - \sum_{i=1}^{2} h_{2,i}(w_1(t))\widetilde{Q}_i w(t),\end{aligned} \tag{31}$$

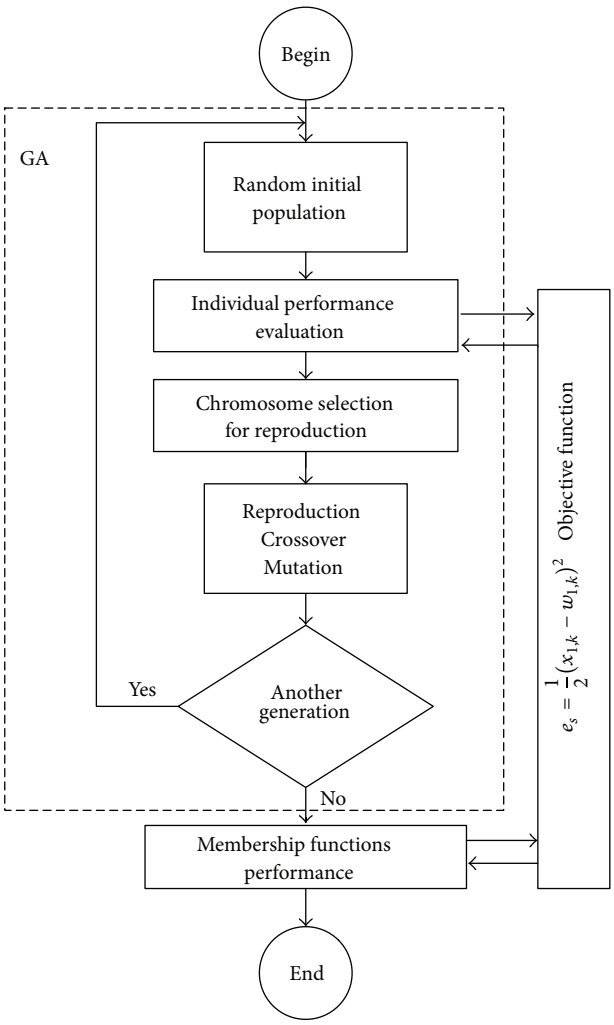


Figure 3: Flowchart of genetic algorithms.

where

$$\widetilde{A}_1 = \begin{bmatrix} 0 & -1 & -1 \\ 1 & a & 0 \\ b & 0 & -d \end{bmatrix},$$
$$\widetilde{S}_1 = \begin{bmatrix} 0 & -1 & -1 \\ 1 & a_w & 0 \\ b_w & 0 & -d_w \end{bmatrix},$$
$$\widetilde{A}_2 = \begin{bmatrix} 0 & -1 & -1 \\ 1 & a & 0 \\ b & 0 & d \end{bmatrix}, \tag{32}$$
$$\widetilde{S}_2 = \begin{bmatrix} 0 & -1 & -1 \\ 1 & a_w & 0 \\ b_w & 0 & d_w \end{bmatrix},$$
$$\widetilde{B}_1 = \widetilde{B}_2 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T,$$
$$\widetilde{C}_i = \widetilde{Q}_i = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}.$$

Besides, the membership functions for such systems are

$$\begin{aligned} h_{1,1}(x_1) &= \frac{1}{2}\left(1 + \frac{c - x_1}{d}\right), \\ h_{1,2}(x_1) &= \frac{1}{2}\left(1 - \frac{c - x_1}{d}\right), \\ h_{2,1}(w_1) &= \frac{1}{2}\left(1 + \frac{c - w_1}{d}\right), \\ h_{2,2}(w_1) &= \frac{1}{2}\left(1 - \frac{c - w_1}{d}\right). \end{aligned} \tag{33}$$

Now, the continuous-time T-S fuzzy model can be converted to the following discrete counterpart by using [1]

$$\begin{aligned} G_i &= \exp(A_i T_s) = I + A_i T_s + A_i^2 \frac{T_s^2}{2!} + \cdots, \\ H_i &= \int_0^{T_s} \exp(A_i \tau) B_i d\tau = (G_i - I) A_i^{-1} B_i. \end{aligned} \tag{34}$$

Therefore, the discrete-time T-S fuzzy model is given by

$$\begin{aligned} x_{k+1} &= \sum_{i=1}^{2} h_{1,i}(x_{1,k}) \{A_i x_k + B_i u_k + P_i w_k\}, \\ w_{k+1} &= \sum_{i=1}^{2} h_{2,i}(w_{1,k}) S_i w_k, \\ e_k &= \sum_{i=1}^{2} h_{1,i}(x_{1,k}) C_i x_k - \sum_{i=1}^{2} h_{2,i}(w_{1,k}) Q_i w_k, \end{aligned} \tag{35}$$

where

$$A_1 = \begin{bmatrix} 1 & -T_s & -T_s \\ T_s & aT_s + 1 & 0 \\ bT_s & 0 & 1 - dT_s \end{bmatrix},$$
$$S_1 = \begin{bmatrix} 1 & -T_s & -T_s \\ T_s & a_w T_s + 1 & 0 \\ b_w T_s & 0 & 1 - d_w T_s \end{bmatrix},$$
$$A_2 = \begin{bmatrix} 1 & -T_s & -T_s \\ T_s & aT_s + 1 & 0 \\ bT_s & 0 & T_s d + 1 \end{bmatrix}, \tag{36}$$
$$S_2 = \begin{bmatrix} 1 & -T_s & -T_s \\ T_s & a_w T_s + 1 & 0 \\ b_w T_s & 0 & d_w T_s + 1 \end{bmatrix},$$
$$B_1 = B_2 = \begin{bmatrix} 0 & 0 & T_s \end{bmatrix}^T,$$
$$C_i = Q_i = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}.$$

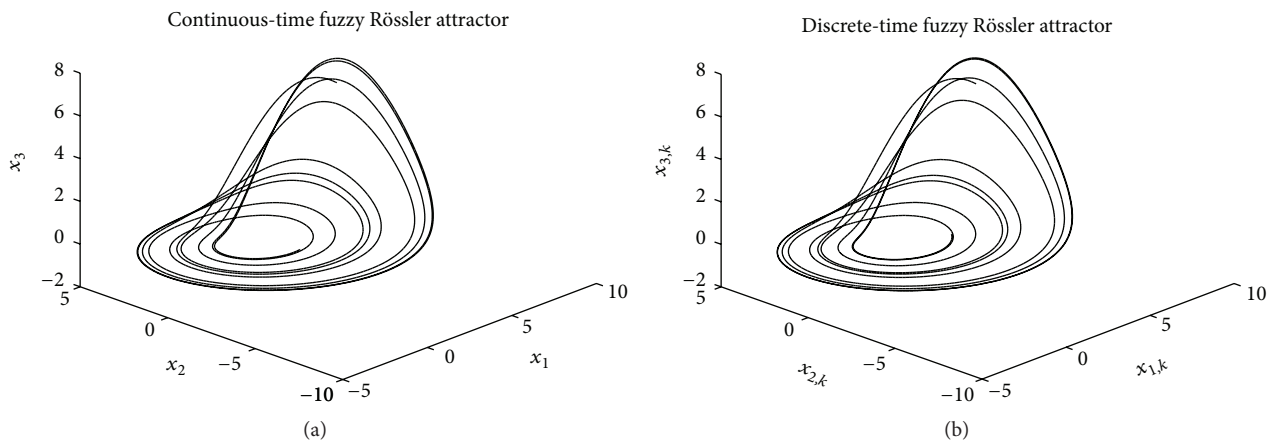


Figure 4: (a) Continuous T-S fuzzy model for a Rössler attractor and (b) discrete T-S fuzzy model for a Rössler attractor.

Besides, the membership functions for the T-S discrete-time fuzzy model are

$$\begin{aligned} h_{1,1}(x_{1,k}) &= \frac{1}{2}\left(1+\frac{c-x_{1,k}}{d}\right), \\ h_{1,2}(x_{1,k}) &= \frac{1}{2}\left(1+\frac{c-x_{1,k}}{d}\right), \\ h_{1,2}(w_{1,k}) &= \frac{1}{2}\left(1+\frac{c_w-w_{1,k}}{d_w}\right), \\ h_{2,2}(w_{1,k}) &= \frac{1}{2}\left(1+\frac{c_w-w_{1,k}}{d_w}\right), \end{aligned} \tag{37}$$

with $a = 0.34$, $b = 0.4$, $c = 4.5$, $d = 10$, $a_w = 0.34$, $b_w = 0.4$, $c_w = 4.5$, $d_w = 10$, and T_s as the sampling time.

Figure 4(b) shows the trajectory of the discrete-time version of the continuous-time T-S fuzzy Rössler model, with $T_s = 0.00357$. It can be seen that the overall shape of the trajectory is similar to that in Figure 4(a).

Then, by using the approach derived in this work and from (20) and (17), the zero error steady-state manifold $\pi(w_k) = \Pi$ is

$$\Pi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{38}$$

Γ_{ij} are

$$\begin{aligned} \Gamma_{1,1} &= [0 \;\; 0 \;\; 0], \\ \Gamma_{1,2} &= [0 \;\; 0 \;\; 20], \\ \Gamma_{2,1} &= [0 \;\; 0 \;\; -20], \\ \Gamma_{2,2} &= [0 \;\; 0 \;\; 0]. \end{aligned} \tag{39}$$

On the other hand, the fuzzy stabilizer for this system is computed by means of *Ackermann's formula*, and by locating the eigenvalues at

(1) subsystem $[0.9980+0.0062i \;\; 0.9980-0.0062i \;\; 0.9982]$,

(2) subsystem $[0.9980+0.0062i \;\; 0.9980-0.0062i \;\; 0.9982]$,

the following gains are obtained:

$$\begin{aligned} K_1 &= [3.1629 \;\; 0.9257 \;\; 8.0354], \\ K_2 &= [3.1629 \;\; 0.9257 \;\; -11.9646]. \end{aligned} \tag{40}$$

Remark 5. It is important to verify that the fuzzy feedback stabilizer is valid for the overall T-S fuzzy model, by checking that the eigenvalues of the interpolation regions are inside the unit circle also [31].

Then, by setting the initial conditions as $x_{1,k} = 5$, $x_{2,k} = 0$, $x_{3,k} = 6$, $w_{1,k} = 1$, $w_{2,k} = 0$, and $w_{1,k} = 0$ and by applying the controller (23), the results depicted in Figures 5 and 6 are obtained.

The tracking for the drive states and response states is drawn in Figure 7.

4.1. Generalized Synchronization. The discrete-time fuzzy synchronization problem is solvable when the conditions of Theorem 3 are fulfilled. However, when two chaotic systems are different, fulfilling these conditions is not so common. This is because the local regulators have different zero error steady-state manifolds, in general [20]. From the regulation point of view, the problem can be seen as finding, if it is possible, a transformation $\pi : w_k \to x_k$ regarding mapping the trajectories of the drive system into the ones of the response systems; that is, $x_k = \pi(w_k)$; this is known as Generalized Synchronization [37, 38] and satisfies

$$\lim_{k\to\infty} \left\| x_k(x(0)) - w_k(w(0)) \right\| = 0, \tag{41}$$

with $x(0)$ and $w(0)$ as initial conditions.

Now, consider the chaotic drive system as $\dot{w} = f(w)$ and $\dot{x} = f(x, w, u)$ as the response system.

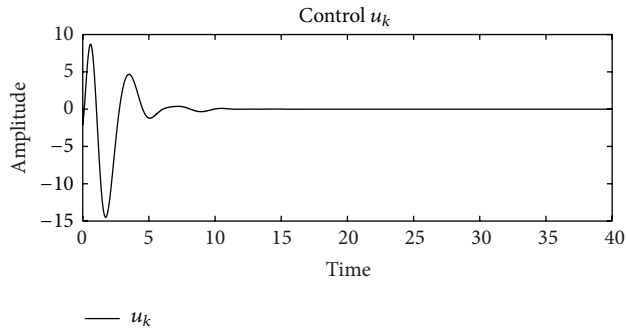


Figure 5: Control signal for the Complete Synchronization of two discrete-time T-S fuzzy Rössler systems.

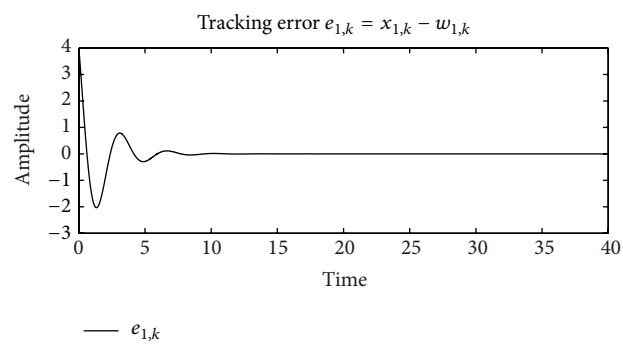


Figure 6: Tracking error signal for the Complete Synchronization of two discrete-time T-S fuzzy Rössler systems.

The equations for the aforementioned systems are described as follows:

$$\begin{gathered}
\dot{w}_1 = \alpha\left(w_2 - w_1\right), \\
\dot{w}_2 = r w_1 - w_2 - w_1 w_3, \\
\dot{w}_3 = w_1 w_2 - \beta w_3, \\
\text{Lorenz Attractor} \\
\text{Drive system} \\
\dot{x}_1 = a\left(x_2 - x_1\right), \\
\dot{x}_2 = (c-a)\, x_1 - x_1 x_3 + c x_2 + u, \\
\dot{x}_3 = x_1 x_2 - b x_3, \\
\text{Chen's attractor} \\
\text{Response system}
\end{gathered} \tag{42}$$

As before these systems can be exactly represented by means of the following two-rule continuous-time T-S fuzzy models when $x_1 \in [X_{\min}, X_{\max}]$ and $w_1 \in [M_1, M_2]$. Therefore,

$$\begin{aligned}
\dot{x}(t) &= \sum_{i=1}^{2} h_{1,i}\left(x_1(t)\right)\left\{\widetilde{A}_i x(t) + \widetilde{B}_i u(t)\right\}, \\
\dot{w}(t) &= \sum_{i=1}^{2} h_{2,i}\left(w_1(t)\right)\widetilde{S}_i w(t), \\
e(t) &= \sum_{i=1}^{2} h_{1,i}\left(x_1(t)\right)\widetilde{C}_i x(t) \\
&\quad - \sum_{i=1}^{2} h_{2,i}\left(w_1(t)\right)\widetilde{Q}_i w(t),
\end{aligned} \tag{43}$$

where

$$\begin{aligned}
\widetilde{A}_1 &= \begin{bmatrix} -a & a & 0 \\ c-a & c & -X_{\min} \\ 0 & X_{\min} & -b \end{bmatrix}, \\
\widetilde{S}_1 &= \begin{bmatrix} -\alpha & \alpha & 0 \\ r & -1 & -M_1 \\ 0 & M_1 & -\beta \end{bmatrix}, \\
\widetilde{A}_2 &= \begin{bmatrix} -a & a & 0 \\ c-a & c & -X_{\max} \\ 0 & X_{\max} & -b \end{bmatrix}, \\
\widetilde{S}_2 &= \begin{bmatrix} -\alpha & \alpha & 0 \\ r & -1 & -M_2 \\ 0 & M_2 & -\beta \end{bmatrix}, \\
\widetilde{B}_1 &= \widetilde{B}_2 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T, \\
\widetilde{C}_i &= \widetilde{Q}_i = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}.
\end{aligned} \tag{44}$$

Then, by using (34), the discrete counterpart is obtained:

$$\begin{aligned}
x_{k+1} &= \sum_{i=1}^{2} h_{1,i}\left(x_{1,k}\right)\left\{A_i x_k + B_i u_k + P_i w_k\right\}, \\
w_{k+1} &= \sum_{i=1}^{2} h_{2,i}\left(w_{1,k}\right) S_i w_k, \\
e_k &= \sum_{i=1}^{2} h_{1,i}\left(x_{1,k}\right) C_i x_k - \sum_{i=1}^{2} h_{2,i}\left(w_{1,k}\right) Q_i w_k,
\end{aligned} \tag{45}$$

where

$$\begin{aligned}
A_1 &= \begin{bmatrix} 1 - aT_s & aT_s & 0 \\ -(a-c)\,T_s & cT_s + 1 & -X_{\min} T_s \\ 0 & T_s X_{\min} & 1 - bT_s \end{bmatrix}, \\
S_1 &= \begin{bmatrix} 1 - a_w T_s & a_w T_s & 0 \\ c_w T_s & 1 - T_s & -M_1 T_s \\ 0 & M_1 T_s & 1 - b_w T_s \end{bmatrix}, \\
A_2 &= \begin{bmatrix} 1 - aT_s & aT_s & 0 \\ -(a-c)\,T_s & cT_s + 1 & -X_{\max} T_s \\ 0 & T_s X_{\max} & 1 + bT_s \end{bmatrix}, \\
S_2 &= \begin{bmatrix} 1 - a_w T_s & a_w T_s & 0 \\ c_w T_s & 1 - T_s & -M_2 T_s \\ 0 & M_2 T_s & 1 - b_w T_s \end{bmatrix},
\end{aligned}$$

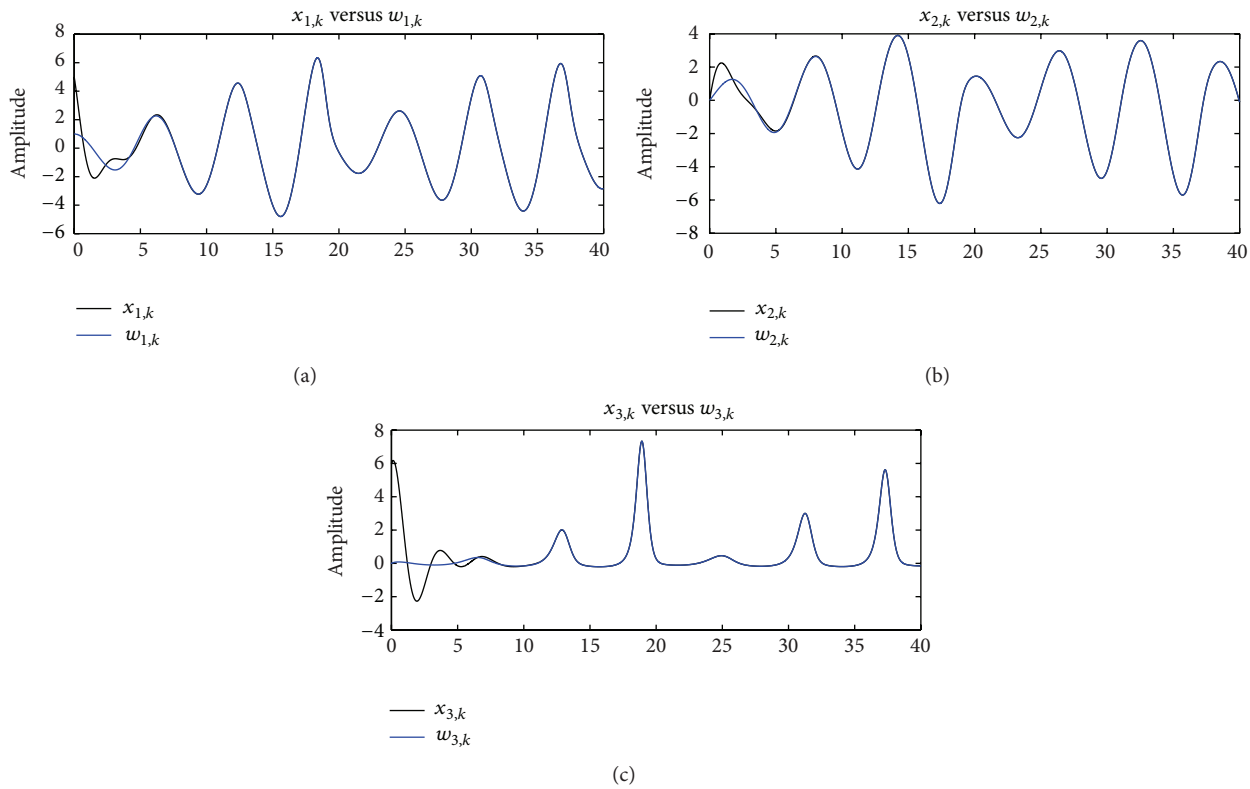


Figure 7: States of drive and response tracking for a Complete Synchronization.

$$B_1 = B_2 = \begin{bmatrix} 0 & T_s & 0 \end{bmatrix}^T,$$
$$C_i = Q_i = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \tag{46}$$

with $a = 35$, $b = 3$, $c = 28$, $a_w = 10$, $b_w = 8/3$, and $c_w = 28$ and T_s as the sampling time. Notice that, in this case, the Generalized Synchronization problem consists of the tracking of w_1 by x_1. This can be inferred by the form of C_i and Q_i. The membership functions for this system are defined as follows:

Plant Membership Functions

$$h_{1,1}(x_{1,k}) = \frac{-x_{1,k} + X_{\max}}{X_{\max} - X_{\min}},$$
$$h_{1,2}(x_{1,k}) = \frac{x_{1,k} - X_{\min}}{X_{\max} - X_{\min}}. \tag{47}$$

Exosystem Membership Functions

$$h_{2,1}(w_{1,k}) = \frac{-w_{1,k} + M_1}{M_2 - M_1},$$
$$h_{2,2}(w_{1,k}) = \frac{w_{1,k} - M_2}{M_2 - M_1}. \tag{48}$$

They are depicted in Figure 8.

Figures 9(a) and 9(b) show the behavior of the two discrete-time T-S fuzzy models, with $x(0) = [1 \;\; 1 \;\; 1]^T$ and $w(0) = [1 \;\; 0 \;\; 0]^T$. Then, by using the approach derived in this work and from (20) and (17), the zero error steady-state manifold for each subsystem is

$$\Pi_{1,1} = \begin{bmatrix} 1 & 0 & 0 \\ 0.7143 & 0.2857 & 0 \\ 3.0041 & -0.0143 & 1.2861 \end{bmatrix},$$
$$\Pi_{1,2} = \begin{bmatrix} 1 & 0 & 0 \\ 0.7143 & 0.2857 & 0 \\ 3.0041 & -0.0143 & -1.2861 \end{bmatrix},$$
$$\Pi_{2,1} = \begin{bmatrix} 1 & 0 & 0 \\ 0.7143 & 0.2857 & 0 \\ -3.0041 & 0.0143 & -1.2861 \end{bmatrix},$$
$$\Pi_{2,2} = \begin{bmatrix} 1 & 0 & 0 \\ 0.7143 & 0.2857 & 0 \\ -3.0041 & 0.0143 & 1.2861 \end{bmatrix}. \tag{49}$$

Γ_{ij} are

$$\Gamma_{1,1} = \begin{bmatrix} -102.2648 & -0.7142 & -30.0121 \end{bmatrix},$$
$$\Gamma_{1,2} = \begin{bmatrix} -102.2648 & -0.7142 & 30.0121 \end{bmatrix},$$
$$\Gamma_{2,1} = \begin{bmatrix} -102.2648 & -0.7142 & -30.0121 \end{bmatrix},$$
$$\Gamma_{2,2} = \begin{bmatrix} -102.2648 & -0.7142 & 30.0121 \end{bmatrix}. \tag{50}$$

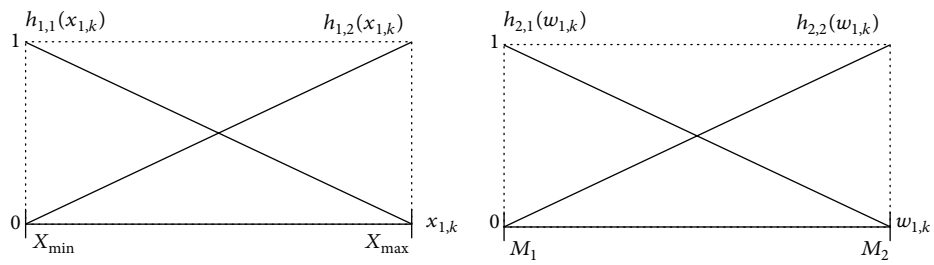


Figure 8: Membership functions for Generalized Synchronization.

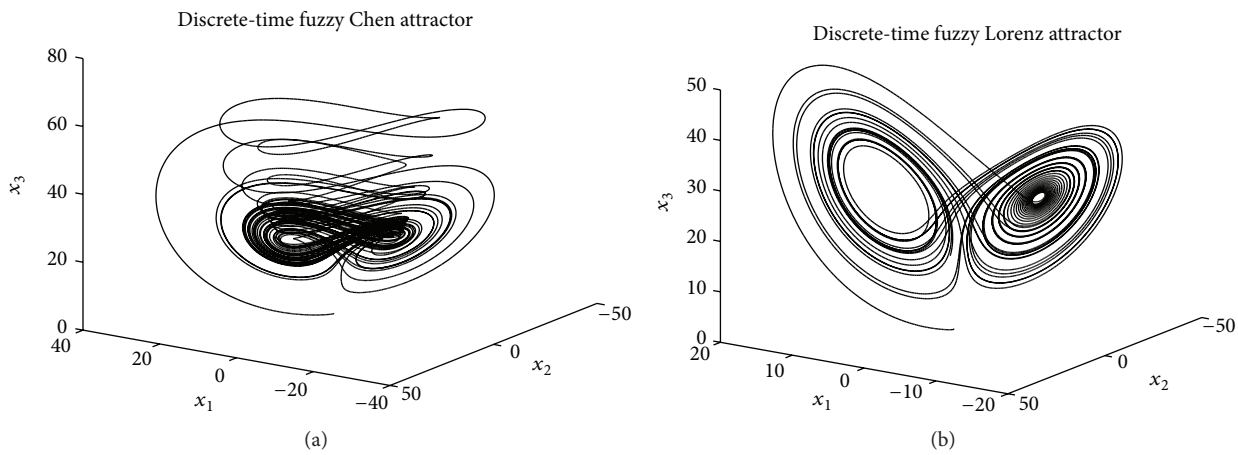


Figure 9: (a) Discrete-time T-S fuzzy model for a Chen attractor and (b) discrete T-S fuzzy model for a Lorenz attractor.

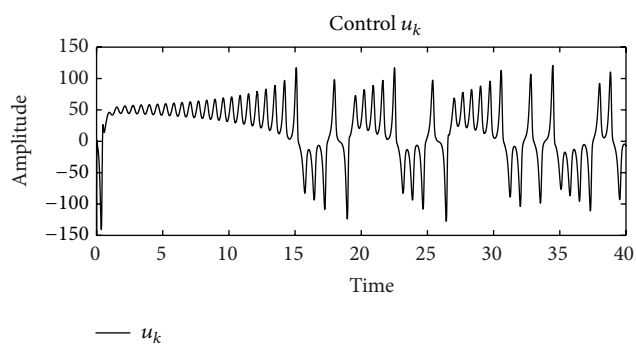


Figure 10: Control signal for the Generalized Synchronization of Chen-Lorenz discrete-time T-S fuzzy systems.

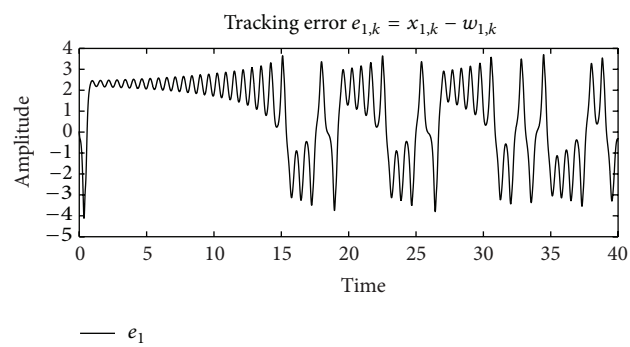


Figure 11: Tracking error $e_{1,k}$ signal for the Generalized Synchronization of Chen-Lorenz discrete-time T-S fuzzy systems.

On the other hand, the fuzzy stabilizer for this system is computed by means of *Ackermann's formula*, and by locating the eigenvalues at

(1) subsystem $[0.9596 \quad 0.9955 \quad 0.82]$,

(2) subsystem $[0.9596 \quad 0.9955 \quad 0.82]$,

the following gains are obtained:

$$K_1 = [-14.3228 \quad -214.9000 \quad -19.6566],$$
$$K_2 = [-14.3228 \quad -214.9000 \quad 19.6566]. \tag{51}$$

Notice that the stability is ensured in the fuzzy interpolation region by these gains, for all $t \geq 0$.

As can be readily observed, conditions of Theorem 3 are not fulfilled because the zero error steady-state manifold is not the same for the local regulators. However, the tracking error can be bounded by using the controller defined in (22). Figures 10–12 depicted the behavior of the controller (22) with $x(0) = [1 \quad 1 \quad 1]^T$ and $w(0) = [1 \quad 0 \quad 0]^T$ as the initial conditions.

4.2. Generalized Synchronization by Using Genetic Algorithms. The main objective of integrating genetic algorithms is to obtain values of the parameters (a, b, and c) of the new membership functions of the regulator and reducing, in this

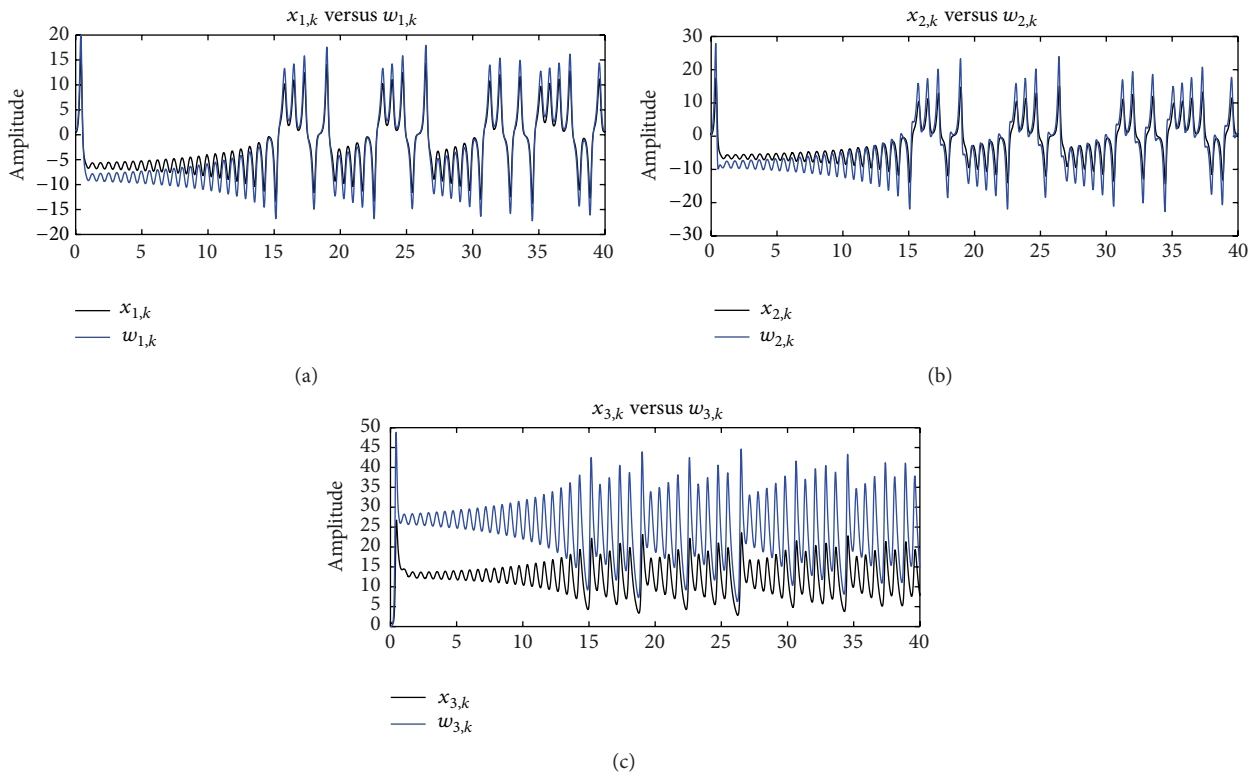


Figure 12: States of drive and response system for a Generalized Synchronization.

way, the tracking error. To this end, it is important to consider the following:

(i) Consider the representation of the potential solutions of the problem; the common representation is binary.

(ii) An initial population of possible solutions, in a random process, is selected (population size affects the efficiency and performance of GA). For this case, a population size equal to 20 chromosomes is chosen.

(iii) A fitness function, which indicates how good or bad a certain solution is, is defined. In this case the fitness function is the mean square tracking error, given by the following expression:

$$e_s = \frac{1}{2}\left(x_{1,k} - w_{1,k}\right)^2. \tag{52}$$

For the representation of possible solutions it is necessary to know the variables to optimize; for this case it has a_i and a_j parameters for $i = 1, 2$ and $j = 1, 2$ defined in (27) and (28); besides, b_i, b_j, c_i, and c_j will be fixed with a constant value; that is, $b_1 = 3$, $c_1 = 30$ for (27) considering $\mu_{1,2} = 1 - \mu_{1,1}$; $b_1 = 3$, $c_1 = 20$ for (28) also considering $\mu_{2,2} = 1 - \mu_{2,1}$. The abovementioned are represented in Figure 13.

It is also important to know the interval value in which the variable a will be operating; for this case, such interval is $x_{1,k} \in [-30 \;\; 30]$. Thus, the size of each variable, in bits, and the chromosome length can be computed by using the following expression [39]:

$$S_{\text{var}} = (\text{int} \uparrow) \log_2 \left(\lim_{\text{upper}} - \lim_{\text{lower}}\right) 10^{\text{accurate}}, \tag{53}$$

where S_{var} is the size of each variable in bits, the term (int $\uparrow$) is the decimal integer value, and $\lim_{\text{upper}} = 30$ and $\lim_{\text{lower}} = -30$ with an accuracy in 3. Therefore, the variable size is equal to 16 bits, but there are two variables, a_1 and a_2; thus the chromosomes length is 32 bits. Therefore, the chromosomes are represented in Figure 14.

The initial population is random binary. To assess each individual in the objective function is needed to decode, in this case, a real number. To this end, the following expression is used:

$$\text{Var} = \left[\text{Dec}_{\text{val}}\left(\frac{\lim_{\text{upper}} - \lim_{\text{lower}}}{2^{\text{length(bit)}} - 1}\right) + \lim_{\text{lower}}\right]. \tag{54}$$

Var is a decimal value (phenotype) for each binary chromosome (genotype), Dec_{val} is the decimal value for each binary string, and length(bit) is equal to 16 bits.

In order to compute the individual aptitude is necessary to introduce the value of each parameter in the bell-shaped membership functions μ_1 and μ_2. The tracking error function (52) disregards the nonoptimal solution and allows the optimal performance by considering that $x_{1,k}$ tracks $w_{1,k}$. In Table 1 the individual aptitude is depicted.

Table 1: Individual aptitude.

Chromosome	Binary string	Decoded integer	Fitness function
1	010110100010001000101110101011 01	$f(-7.691493, -16.518837)$	0.971109
2	110010111011111111110101000000111	$f(15.386557, 21.537530)$	6.840302
3	101000110000100011101111010101 11	$f(7.116228, 22.616648)$	9.404465
4	011011101111001000111100010110 00	$f(-3.463890, -13.742489)$	0.096779
5	011111101100100010001110100000000	$f(-0.247166, 2.945754)$	0.017387
6	011111101100100010001110100000000	$f(-2.282414, -24.456703)$	12.438373
7	011101001100001100000111100110 01	$f(-13.451286, 16.387121)$	0.761323
8	001111011100011111010000101011 00	$f(4.320043, -25.338247)$	13.686271
9	100101010100010000000011010000 10	$f(19.709377, 4.057404)$	0.209690
10	100000100010111100010001011101 11	$f(0.443946, -22.452399)$	9.159322
11	001111001010001110111100111000 01	$f(-13.682979, 12.366613)$	0.000746
12	000011000111110111001010001011 00	$f(-23.463279, 15.066789)$	0.000895
13	100101100000100010010100011111 01	$f(4.475563, 4.162142)$	0.017576
14	101100110101101101001001011110 10	$f(10.432136, -11.074846)$	0.001021
15	010100111010101011111010110011 01	$f(-9.005478, 24.944686)$	13.030471
16	001000110110011000110000000001 011	$f(-18.809583, -16.241123)$	0.399280
17	100100011001000011111100101111 01	$f(3.567834, 25.439811)$	13.825474
18	100010110010101101001110111100 01	$f(2.268925, -9.964782)$	0.003840
19	011000000110111011000000001001 011	$f(-6.412421, 13.060105)$	0.039420
20	000010111100010100010100110001 01	$f(-23.609277, -21.781125)$	5.130460

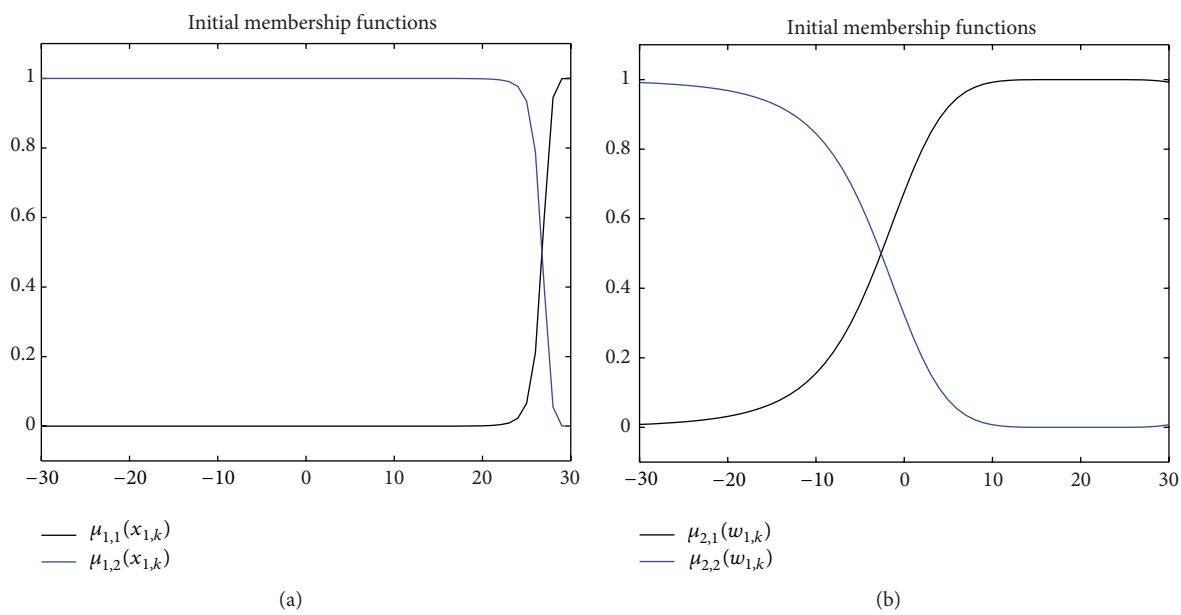


Figure 13: Initial membership functions for the proposed fuzzy regulator tuned by GAs.

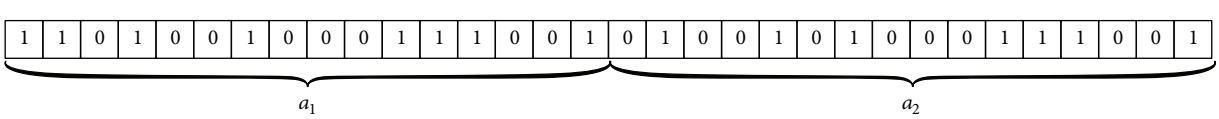


Figure 14: Individual binary representation.

Each fitness function value is converted into a set point or fitness value. To this end, the following expression is used:

$$\text{fitness}_{\text{val}} = F(\star) - F_{\min}, \tag{55}$$

where $F(\star)$ is the fitness function value and $F_{\min}$ is the minimum fitness function value.

The two chromosomes with better fitness (elite chromosomes) can live and produce offspring in the next generation. There are many methods to select a pair of chromosomes; the most popular one is named *proportional selection method*:

(a) Calculate the total fitness value

$$\text{fitness}_T = \sum_{i=1}^{20} \text{fitness}(i). \tag{56}$$

(b) Compute the probability P_i for each chromosome

$$P_i = \frac{\text{fitness}(i)}{\text{fitness}_T}. \tag{57}$$

(c) Calculate the cumulative probability for each chromosome

$$Q_k = \sum_{k=0}^{i} P_k. \tag{58}$$

The total fitness value fitness_T is equal to 86.03420. The probabilities P_i and Q_k are shown in Table 2.

(d) Generate a random number r in the range $[0, 1]$.

(e) If $Q_i - 1 < r \leq Q_i$, then select the chromosome to be the one of the parents.

(f) Repeat (d) and (e) to obtain the other parent.

Apply crossover operation on the selected pair, if they have been chosen for crossover (based on probability of crossover $p_c = 1.0$). The most applied crossover operation is single point crossover. Based on the probability of bit mutation $p_m = 0.01$, flip the correspondent bit if selected for mutation. At this point, the process of producing a pair of offspring from two selected parents is finished.

The elite chromosomes of the previous population are not subject to mutation.

In the following simulations the states $x_{1,k}$, control signal u_k, and tracking error $e_{1,k}$ depicted in Figures 16–18, respectively, are obtained after applying (29) and by replacing the new membership functions adjusted by GAs. The final membership functions after tuning the form parameter by means of GA are depicted in Figure 15. It can be readily observed that these new membership functions are different from the original ones (see Figure 13). Notice that, in this example, the tracking error (Figure 18) is less than the error obtained by using the approach discussed in Section 4. This is due to the new membership functions which allow reducing considerably the tracking error. This suggests that the approach presented in this work may be improved by tuning the parameters of center and amplitude in the new membership functions; however this study is not completed yet.

Table 2: Individual aptitude.

Chromosome	P_i	Q_k
1	0.011	0.011
2	0.080	0.091
3	0.109	0.200
4	0.001	0.201
5	0.000	0.201
6	0.145	0.346
7	0.009	0.355
8	0.159	0.514
9	0.002	0.516
10	0.106	0.623
11	0.000	0.623
12	0.000	0.623
13	0.000	0.623
14	0.000	0.623
15	0.151	0.775
16	0.005	0.779
17	0.161	0.940
18	0.000	0.940
19	0.000	0.940
20	0.060	1.000

4.3. Pseudofuzzy Generalized Synchronization by Using GAs. In this section the Generalized Synchronization by using GAs and linear local regulator design will be addressed. However, the crossed terms within the local regulators are arbitrarily removed from the design process. Thus, the steady-state input $\gamma(w_k)$ can be defined as

$$\gamma(w_k) = \sum_{i=1}^{r_1} \mu_{1,i}(z_{1,k}) \Gamma_{ii} w_k, \tag{59}$$

where $\mu_{1,i}(z_{1,k})$ are new membership functions, such that the fuzzy output regulator obtained from local regulators provides the exact fuzzy output regulation. This approach requires the computation of the linear local controllers and the computation of the new membership functions; such functions are represented by (27) and (28). Therefore the final control system is defined by

$$u_k = \sum_{i=1}^{r_1} h_{1,i}(x_k) k_i \{x_k - \pi(w_k)\} + \sum_{i=1}^{r_1} \mu_{1,i}(w_k) \Gamma_{ii} w_k, \tag{60}$$

because $z_{*,k}$ is a function of x_k and in steady-state $x_k = \pi(w_k)$.

For this case the initial membership functions proposed are shown in Figure 19. As it can be seen, these membership functions are not properly fuzzy because they do not fulfill the convex sum; that is, $\sum_{i=1}^{r_1} \mu_i(w_k) \neq 0$. Even so, these functions will be adjusted by GAs in order to ensure the output regulation by using input (60). The following simulation provides a better behavior in the Lorenz-Chen discrete-time fuzzy system by using different membership function and tuning by GAs. Such results are depicted in Figures 20–22.

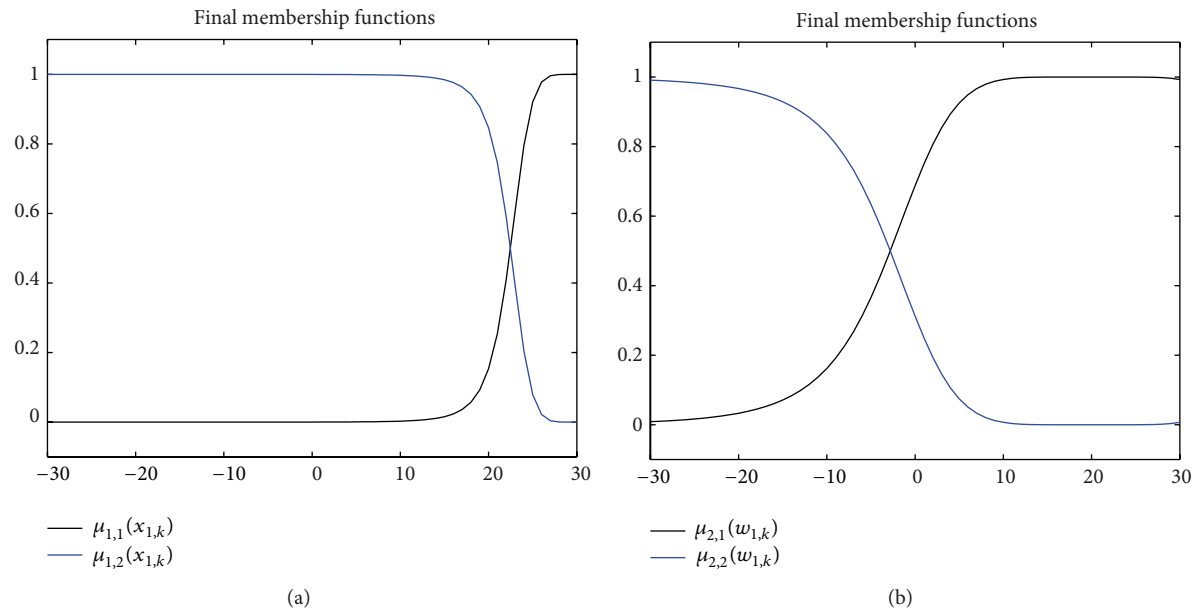


Figure 15: Final membership functions for the proposed fuzzy regulator tuned by GAs.

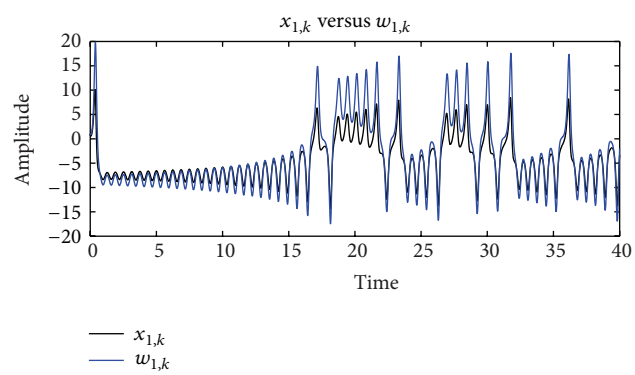


Figure 16: States $x_{1,k}$ and $w_{1,k}$ for the Generalized Synchronization of Chen-Lorenz discrete-time T-S fuzzy systems.

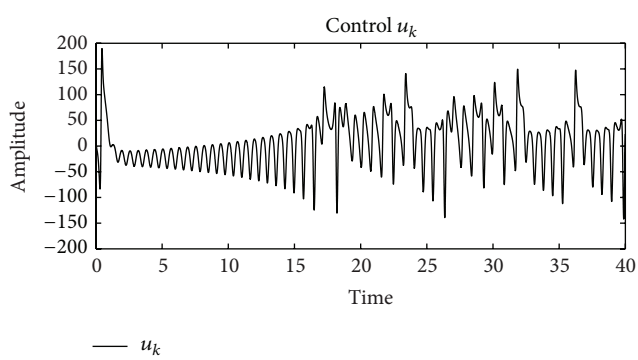


Figure 17: Control signal for the Generalized Synchronization of Chen-Lorenz discrete-time T-S fuzzy systems.

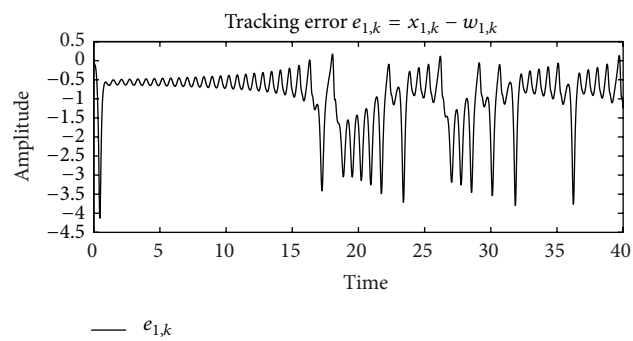


Figure 18: Tracking error $e_{1,k}$ signal for the Generalized Synchronization of Chen-Lorenz discrete-time T-S fuzzy systems.

Notice that the final membership function depicted in Figure 23 is different from the initial proposed in Figure 19 and also notice that the sum of these new interpolation functions is not equal to one. For that reason this approach is called Pseudofuzzy Generalized Synchronization.

5. Conclusions

A fuzzy output regulator for discrete-time systems, based on the combination of linear regulators combined by different membership functions, has been presented. Synchronization of discrete-time chaos attractors can be possible; by means of fuzzy output regulation, sufficient conditions for the controller are given. However, when the conditions can not be fulfilled, new membership functions in the regulator are included; these ones are optimized by genetic algorithms. The main advantage is that membership functions, which allow the proper combination of the local regulators, can be easily obtained by means of GAs. As a consequence, the presented result allows a very precise synchronization of chaotic systems described by T-S discrete-time fuzzy models on the basis of local regulators.

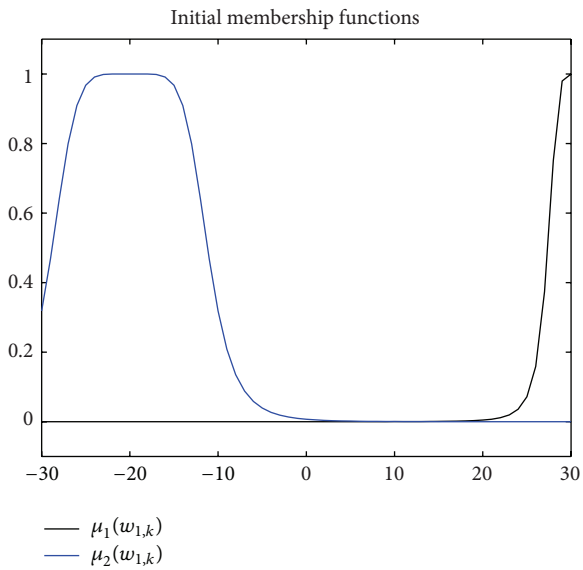


Figure 19: Initial membership functions for the pseudofuzzy regulator tuned by GAs.

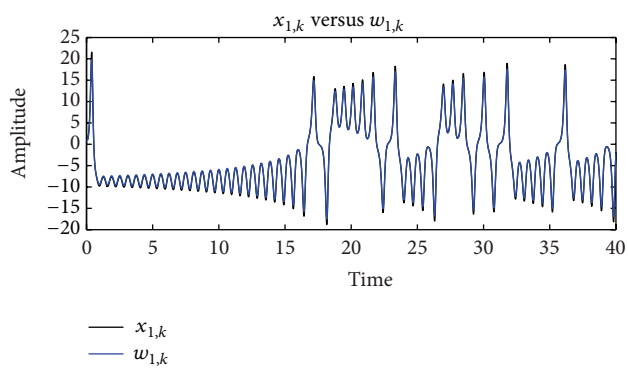


Figure 20: States $x_{1,k}$ and $w_{1,k}$ for the Generalized Synchronization of Chen-Lorenz discrete-time T-S fuzzy systems.

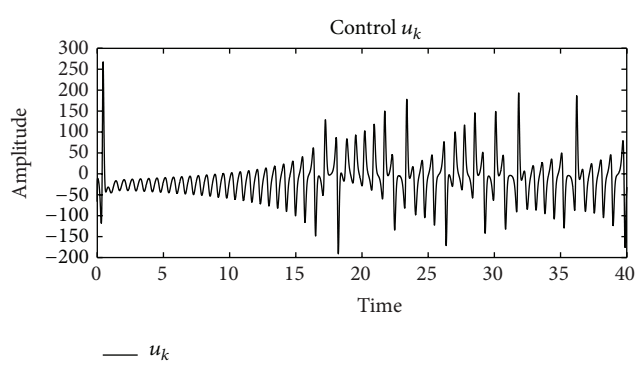


Figure 21: Control signal for the Generalized Synchronization of Chen-Lorenz discrete-time T-S fuzzy systems.

Complete and Generalized Synchronization are used to illustrate the applicability of the proposed approach. Besides, the method proposed in this work avoids the disadvantage of constructing an exact fuzzy regulator based on overall T-S fuzzy system which may result to be very large. Instead, the given approach offers a simple way to design the complete regulator based on local regulators but with membership functions optimized by soft computing.

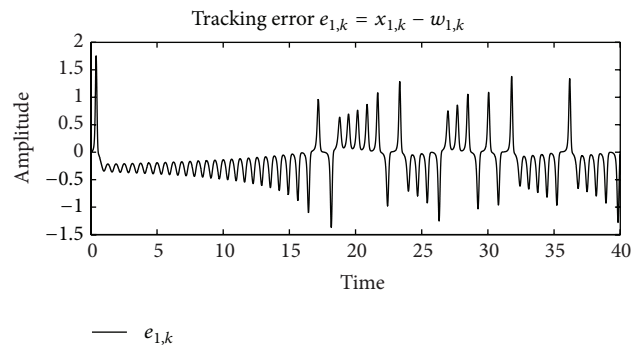


Figure 22: Tracking error $e_{1,k}$ for the Generalized Synchronization of Chen-Lorenz discrete-time T-S fuzzy systems.

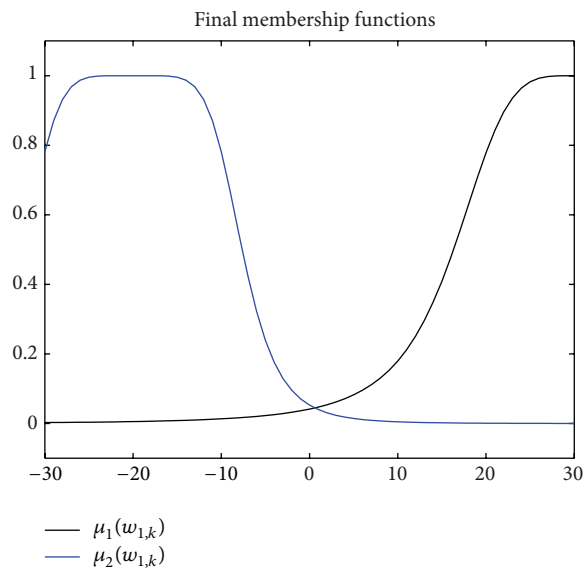


Figure 23: Final interpolation functions for the pseudofuzzy regulator tuned by GAs.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is partially supported by Consejo Nacional de Ciencia y Tecnología (CONACYT) through Scholarship SNI and by Instituto Politécnico Nacional (IPN) through research Projects 20150487 and 20150794 and Scholarships COFAA, EDI, and BEIFI.

References

[1] Z. Li and G. Chen, Eds., *Integration of Fuzzy Logic and Chaos Theory*, vol. 187, Springer, Berlin, Germany, 2006.

[2] B. Wang, H. Cao, Y. Wang, and D. Zhu, "Linear matrix inequality based fuzzy synchronization for fractional order chaos," *Mathematical Problems in Engineering*, vol. 2015, Article ID 128580, 14 pages, 2015.

[3] I. Rechenberg, *Optimization of technical systems after the principles of biological evolution [Ph.D. thesis]*, Fromman-Holzboog, Stuttgart, Germany, 1971.

[4] H.-P. Schwefel, *Numerical Optimization of Computer Models*, John Wiley & Sons, New York, NY, USA, 1981.

[5] H. Bremermann, *The Evolution of Intelligence: The Nervous System as a Model of Its Environment*, University of Washington, Department of Mathematics, 1958.

[6] J. H. Holland, "Outline for a logical theory of adaptive systems," *Journal of the ACM*, vol. 9, no. 3, pp. 297–314, 1962.

[7] J. H. Holland, *Adaptation in Natural and Artificial Systems*, MIT Press, Cambridge, Mass, USA, 1992.

[8] D. Fogel and R. Anderson, "Revisiting Bremermann's genetic algorithm. I. Simultaneous mutation of all parameters," in *Proceedings of the Congress on Evolutionary Computation*, vol. 2, pp. 1204–1209, La Jolla, Calif, USA, 2000.

[9] I. Rechenberg, "Cybernetic solution path of an experimental problem," in *Royal Aircraft Establishment Translation No. 1122*, B. F. Toms, Ed., Ministry of Aviation, Royal Aircraft Establishment, Farnborough Hants, UK, 1965.

[10] C. Hugenii, "Horologium oscillatorium (parisiis, France, 1973)," Tech. Rep., Iowa State University Press, Ames, Iowa, 1986, (English translation: The pendulum clock).

[11] L. M. Pecora and T. L. Carroll, "Synchronization in chaotic systems," *Physical Review Letters*, vol. 64, no. 8, pp. 821–824, 1990.

[12] K. M. Cuomo and A. V. Oppenheim, "Circuit implementation of synchronized chaos with applications to communications," *Physical Review Letters*, vol. 71, no. 1, pp. 65–68, 1993.

[13] M.-L. Hung and H.-T. Yau, "Circuit implementation and synchronization control of chaotic horizontal platform systems by wireless sensors," *Mathematical Problems in Engineering*, vol. 2013, Article ID 903584, 6 pages, 2013.

[14] I. Zelinka, S. Celikovský, H. Richter, and G. Chen, *Evolutionary Algorithms and Chaotic Systems*, Studies in Computational Intelligence, Springer, Berlin, Germany, 2010.

[15] K. Tanaka and H. O. Wang, *Fuzzy Control Systems Design and Analysis. A Linear Matrix Inequality Approach*, John Wiley & Sons, New York, NY, USA, 2001.

[16] E. Solak, Ö. Morgül, and U. Ersoy, "Observer-based control of a class of chaotic systems," *Physics Letters A*, vol. 279, no. 1-2, pp. 47–55, 2001.

[17] Ö. Morgül and E. Solak, "On the synchronization of chaos systems by using state observers," *International Journal of Bifurcation and Chaos*, vol. 7, no. 6, pp. 1307–1322, 1997.

[18] P. Sangapate, "Adaptive control and synchronization of the shallow water model," *Mathematical Problems in Engineering*, vol. 2012, Article ID 529251, 9 pages, 2012.

[19] A. Isidori and C. I. Byrnes, "Output regulation of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 35, no. 2, pp. 131–140, 1990.

[20] J. A. Meda-Campaña, B. Castillo-Toledo, and G. Chen, "Synchronization of chaotic systems from a fuzzy regulation approach," *Fuzzy Sets and Systems*, vol. 160, no. 19, pp. 2860–2875, 2009.

[21] R. Tapia-Herrera, J. A. Meda-Campaña, S. Alcántara-Montes, T. Hernández-Cortés, and L. Salgado-Conrado, "Tuning of a TS fuzzy output regulator using the steepest descent approach and ANFIS," *Mathematical Problems in Engineering*, vol. 2013, Article ID 873430, 14 pages, 2013.

[22] C. I. Byrnes, F. Delli Priscoli, and A. Isidori, *Output Regulation of Uncertain Nonlinear Systems*, Systems & Control: Foundations & Applications, Birkhäuser, Boston, Mass, USA, 1997.

[23] A. Isidori, *Nonlinear Control Systems*, Springer, Berlin, Germany, 1995.

[24] B. Castillo-Toledo and J. A. Meda-Campaña, "The fuzzy discrete-time robust regulation problem: an LMI approach," *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 3, pp. 360–367, 2004.

[25] B. A. Francis, "The linear multivariable regulator problem," *SIAM Journal on Control and Optimization*, vol. 15, no. 3, pp. 486–505, 1977.

[26] L.-X. Wang, *A Course in Fuzzy Systems and Control*, Prentice Hall PTR, Upper Saddle River, NJ, USA, 1997.

[27] B. Castillo-Toledo, J. A. Meda-Campaña, and A. Titli, "A fuzzy output regulator for takagi-sugeno fuzzy models," in *Proceedings of the IEEE International Symposium on Intelligent Control*, vol. 2, pp. 310–315, Houston, Tex, USA, December 2003.

[28] J. A. Meda-Campaña and B. Castillo-Toledo, "On the output regulation for TS fuzzy models using sliding modes," in *Proceedings of the American Control Conference (ACC '05)*, pp. 4062–4067, Portland, Ore, USA, June 2005.

[29] H. O. Wang, K. Tanaka, and M. Griffin, "Parallel distributed compensation of nonlinear systems by Takagi-Sugeno fuzzy model," in *Proceedings of the International Joint Conference of the 4th IEEE International Conference on Fuzzy Systems and the Second International Fuzzy Engineering Symposium*, vol. 2, pp. 531–538, Yokohama, Japan, March 1995.

[30] K. Tanaka and M. Sugeno, "Stability analysis and design of fuzzy control systems," *Fuzzy Sets and Systems*, vol. 45, no. 2, pp. 135–156, 1992.

[31] J. A. Meda-Campana, J. Rodriguez-Valdez, T. Hernandez-Cortes, R. Tapia-Herrera, and V. Nosov, "Analysis of the fuzzy controllability property and stabilization for a class of T–S fuzzy models," *IEEE Transactions on Fuzzy Systems*, vol. 23, no. 2, pp. 291–301, 2015.

[32] I. Abdelmalek, N. Goléa, and M. Laid Hadjili, "A new fuzzy Lyapunov approach to non-quadratic stabilization of Takagi-Sugeno fuzzy models," *International Journal of Applied Mathematics and Computer Science*, vol. 17, no. 1, pp. 39–51, 2007.

[33] B. Castillo-Toledo and J. A. Meda-Campaña, "The fuzzy discrete-time robust regulation problem: a LMI approach," in *Proceedings of the 41st IEEE Conference on Decision and Control*, vol. 2, pp. 2159–2164, IEEE, Las Vegas, Nev, USA, December 2002.

[34] T. Hernández Cortés, J. Meda Campaña, L. Páramo Carranza, and J. Gómez Mancilla, "A simplified output regulator for a class of takagi-sugeno fuzzy models," *Mathematical Problems in Engineering*, vol. 2015, Article ID 148173, 18 pages, 2015.

[35] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley Longman Publishing, Boston, Mass, USA, 1st edition, 1989.

[36] K. Pyragas, "Weak and strong synchronization of chaos," *Physical Review E*, vol. 54, no. 5, pp. R4508–R4511, 1996.

[37] L. Kocarev and U. Parlitz, "Generalized synchronization, predictability, and equivalence of unidirectionally coupled dynamical systems," *Physical Review Letters*, vol. 76, no. 11, pp. 1816–1819, 1996.

[38] R. He and P. G. Vaidya, "Analysis and synthesis of synchronous periodic and chaotic systems," *Physical Review A*, vol. 46, no. 12, pp. 7387–7392, 1992.

[39] A. H. Wright, "Genetic algorithms for real parameter optimization," in *Foundations of Genetic Algorithms*, pp. 205–218, Morgan Kaufmann, Boston, Mass, USA, 1991.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 651841, 10 pages
http://dx.doi.org/10.1155/2015/651841

Research Article

Prognostics and Health Management of an Automated Machining Process

Cheng He,[1,2] Jiaming Li,[3] and George Vachtsevanos[3]

[1] *School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China*
[2] *School of Intelligent Manufacturing and Control Engineering, Shanghai Second Polytechnic University, Shanghai 201209, China*
[3] *School of Electrical & Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250, USA*

Correspondence should be addressed to George Vachtsevanos; giv@gatech.edu

Received 21 September 2015; Accepted 4 November 2015

Academic Editor: Uchechukwu E. Vincent

Machine failure modes are presenting a major burden to the operator, the plant, and the enterprise causing significant downtime, labor cost, and reduced revenue. New technologies are emerging over the past years to monitor the machine's performance, detect and isolate incipient failures or faults, and take appropriate actions to mitigate such detrimental events. This paper addresses the development and application of novel Prognostics and Health Management (PHM) technologies to a prototype machining process (a screw-tightening machine). The enabling technologies are built upon a series of tasks starting with failure analysis, testing, and data processing aimed to extract useful features or condition indicators from raw data, a symbolic regression modeling framework, and a Bayesian estimation method called particle filtering to predict the feature state estimate accurately. The detection scheme declares the fault of a machine critical component with user specified accuracy or confidence and given false alarm rate while the prediction algorithm estimates accurately the remaining useful life of the failing component. Simulation results support the efficacy of the approach and match well the experimental data.

1. Introduction

Prognosis and Health Management (PHM) has emerged over recent years as significant technologies that are making an impact on both military and commercial maintenance practices. Automated machining processes are employed extensively in manufacturing. Screw-tightening machines are critical assets of an automated machining process. This study focuses on such a machining process with novel features for automatic screw tightening in crucial manufacturing, assembly, and other operations. Figure 1 shows a picture of the equipment. Screw tightening is usually carried out manually resulting in wasted time, operator mistakes, and inconsistent applied torques. The automatic assembly line requires, therefore, for improved performance an automatic screw-tightening machine.

The machine consists mainly of seven parts: feeder, screw falling device, screw separating device, screwdriver, guiding device, 3-axis motion platform, and a clamping device. The function of the feeder is to arrange the screws in a line. The operator places thousands of screws in the center of the feeder. The screws are arranged in a row along the track of the feeder via a vibration mechanism and arranged in a line at the end of the track. The screws move along the track of the feeder and, via gravity, they drop into the pipe one by one and arranged in a line in the pipe. A screw separating device picks one screw at a time and sent it to the head of the screwdriver. The 3-axis motion platform moves the screw hole of the parts behind the head of the screwdriver. Then, the screwdriver moves and tightens the screw.

2. Failure Analysis

Machine critical failure modes include the following: the screws may pile up on the trail of the feeder, thus preventing the screw separating device from separating the screws and arranging them one at a time. The 3-axis motion platform does not move to the precise location. The screwdriver is unable to tighten the screw. Among several other failure

Figure 1: Automated screw-tightening machine.

modes, the 3-axis motive platform not moving to its precise location is the most serious one. In this case, the screw cannot be tightened in the screw hole and the parts cannot be connected correctly. The major reason for the 3-axis motive platform not being able to move precisely is the presence of a fault condition in bearings. Bearing fault/failure is deemed to be critical, severe, and frequent and must be addressed if the automated process will perform properly and expeditiously.

The overall architecture for the bearing health management system is shown in Figure 2. We will describe in the next section the function of each module of the architecture.

3. The Test Platform

In order to obtain test data, under normal and faulty conditions, for the targeted bearings a test platform was designed and built. Accelerated life testing is performed on two rolling element bearings. The experimental data was processed and appropriate features or condition indicators (Cis) were extracted. The processed data combined with a suitable model of the fault growth and a novel estimation algorithm, called particle filtering, are used to implement the diagnostic and prognostic routines.

3.1. Test Platform Design. The test platform consists of a single-axis servomotor-driven module and a data acquisition and analysis system. Figure 3 is the single-axis servomotor-driven module. Figure 4 is the data acquisition system.

3.1.1. Single-Axis Servomotor-Driven Module. The single-axis servomotor-driven module includes a 400 w servo motor, a coupling device, two bearings, two bearing-bases, a ball screw, and a suitable electrical control system. The servomotor is connected to the ball screw via the coupling. Both ends of the ball screw are installed in the bearings. The bearings are fixed in the bearing-housing.

3.1.2. Data Acquisition System. The data acquisition system is composed of a 16-channel dynamic signal data acquisition instrument (model: DH5902), a charge adapter, and two

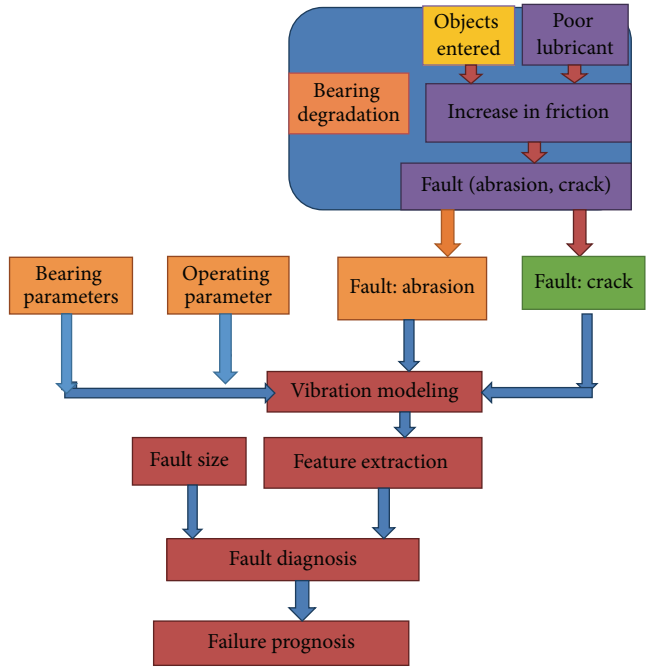


Figure 2: Prognosis and health management architecture for bearing fault modes.

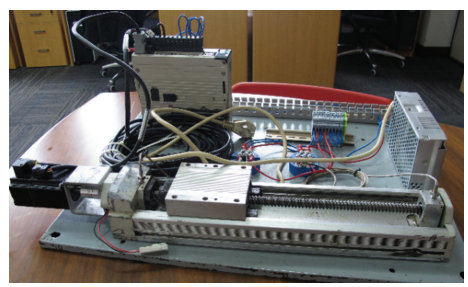


Figure 3: Single-axis servomotor-driven module.

vibration sensors (model: PCB356A32). The vibration sensors are installed in a rack with a magnetic base. They acquire the vibration signal and send it to the 16-channel dynamic signal data acquisition instrument via the charge adapter. The Industrial Control Computer (IPC) receives the data from the 16-channel dynamic signal data acquisition instrument via wireless transmission.

3.1.3. Experimental Parameters. The servomotor has three speeds: 500 r/min, 1000 r/min, and 1500 r/min. The sampling frequency is 2.5 Kz, and the sampling period is 10 seconds. Two bearings were tested. Each bearing is tested along the Z direction resulting in two test points for each speed and each test point providing three entries; that is, there are 6 data entries for each speed setting.

3.2. Fault Analysis. Rolling bearings are typically damaged due to various causes, such as improper assembly, poor lubrication, and moisture. Corrosion and overload may also lead to premature bearing damage. If lubrication and maintenance

Figure 4: Data acquisition system.

procedures are not performed regularly over time, fatigue spall and bearing wear may occur. The main failure modes are fatigue spall, wear, plastic deformation, corrosion, fracture, gluing, cage damage, and so forth.

3.2.1. Fault Conditions. Among all the bearing faults, abrasion is one of the most common. Abrasion implies that material is removed from the metal surface. The characteristic feature for abrasion is a shallow trench, with a bright surface. It is produced in rolling contact surface or guide surface. With abrasion present, the bearing clearance increases.

3.2.2. Root-Cause Analysis. Bearing abrasion is due primarily to two causes: insufficient bearing lubrication and small particle penetration into the bearing resulting in material removal from the metallic surface via sliding friction. Surface wear is the result of the relative motion between rolling track and rolling body when dust and other objects enter the bearing surface resulting in increased bearing surface roughness and reduced motion accuracy. The motion accuracy of the machine is also reduced increasing the machine's vibration and noise. For precision bearings, the amount of wear determines the life of the bearing. On the other side, there is also a kind of wear with slight vibration.

3.3. Fault Injection. In testing, faults are injected periodically until the bearing reaches a failure condition and fault data are recorded. The fault injection consists of the following arrangement: a hole is drilled on the side of the bearing base and a screw is tightened in the hole step by step. The screw generates friction with the bearing surface resulting in an abrasion fault mode.

4. Feature Extraction and Selection

As a bearing defect evolves, due to poor lubrication for example, the fault mode excites a specific frequency associated with the particular type of defect. The amplitude and time duration of the defect frequency are generally good indicators of defect severity. The increase of defect size usually increases the ratio of impulse force to operational noise. The increase of defect length also increases the impulse force duration. If the defect area is large along the raceway turning direction, harmonics of the frequency will also imply an indication of the severity of the defect present. Acquired vibration data are used to extract appropriate features.

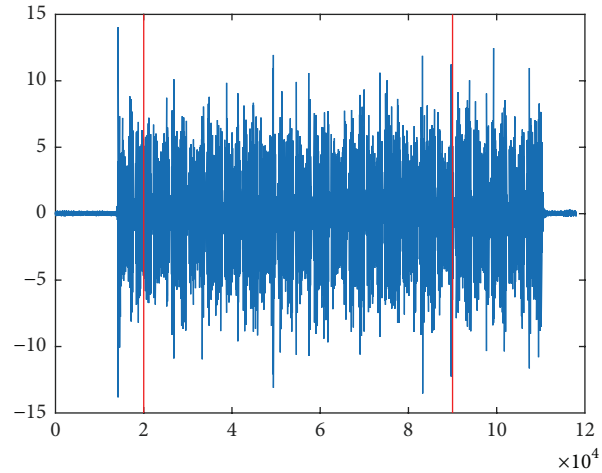


Figure 5: Section signals in time domain.

Table 1: Four special features in frequency domain. (66.25 Hz has the maximum energy.) [1].

	50 Hz	100 Hz	150 Hz	66.25 Hz
Amplitude	f_1	f_3	f_5	f_7
Spectrum density	f_2	f_4	f_6	f_8

4.1. Feature Extraction. The vibration signal is converted from the time to the frequency domain. Analysis shows that the signal changes rapidly between 1700 Hz and 1900 Hz with features extracted in this area. Several vibration signals from bearings with different fault conditions are illustrated in Figure 5. For each data set we cut the period with signal as shown in Figure 5, the section between these two lines.

Frequency domain studies indicate when a bearing defect exists with the defect exhibiting a signature in the frequency spectra of the vibration signals. The signal is digitized in a frequency region around the resonances of the structure and such features as the energy, amplitude, and so forth are extracted. The frequency spectra are shown in Figure 6.

We choose four frequency domain features, as is shown in Table 1.

Time domain features include mean value (f_9) and standard deviation (f_{10}). We calculate the correlation coefficient for each feature. Table 2 shows the results.

4.2. Feature Fusion. Within an automated health management system, there are many areas where fusion technologies play a contributing role. At the lowest level, data fusion can be used to combine information from a multisensory data array to validate signals and create features. At a higher level, fusion may be used to combine features in intelligent ways so as to obtain the best possible diagnosis information. Finally, knowledge fusion is used to incorporate experience-based information such as legacy failure rates or physical model predictions with signal-based information.

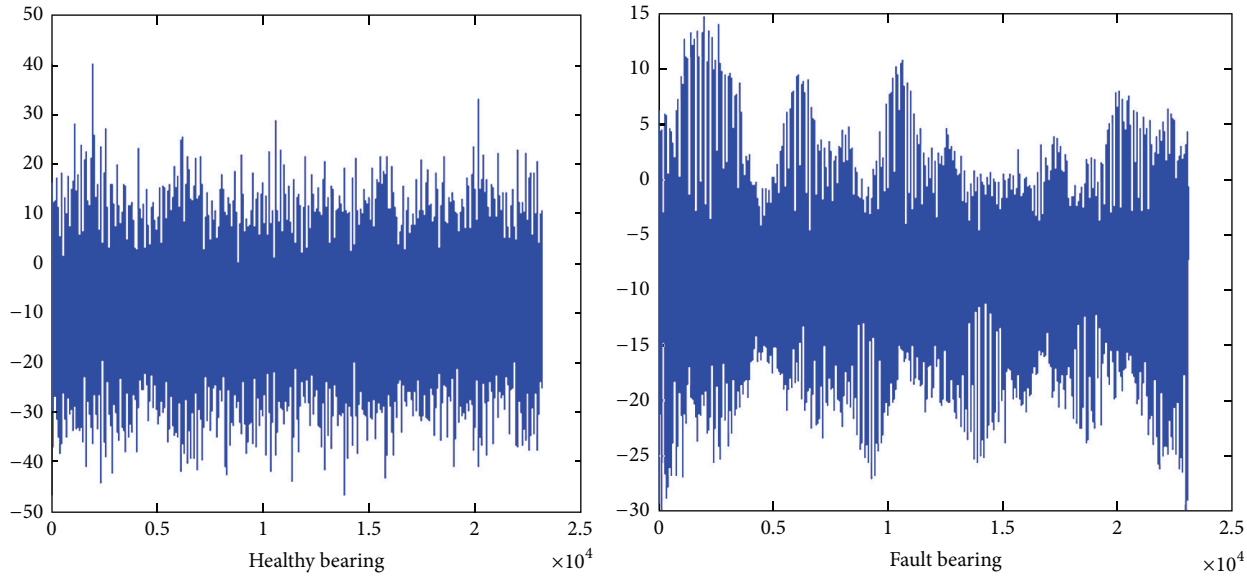


Figure 6: Bearing vibration signals in frequency domain.

Table 2: Correlation coefficient.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}
CC	**0.7972**	0.3797	**0.7867**	0.5902	0.2729	0.0327	0.0097	0.2957	0.2368	**0.9347**

Three features with the highest correlation coefficient are picked up: f_1, f_3, f_{10}.

Table 3: Correlation coefficient.

Feature	f_1	f_3	f_{10}	f_{combined}
CC	0.79	0.78	0.93	0.95

The combined feature has the best highest coefficient and performs better than any individual feature.

Thus, we combine the features and get the combined feature:

$$F_{\text{combined}} = f_1^{0.29} \cdot f_3^{0.18} \cdot f_{10}^{2.01}. \quad (1)$$

The combined correlation coefficient is 0.95. The combined feature performance is shown in Figure 7.

4.3. Performance Comparison. We apply correlation coefficient as the performance metric and come up with Table 3.

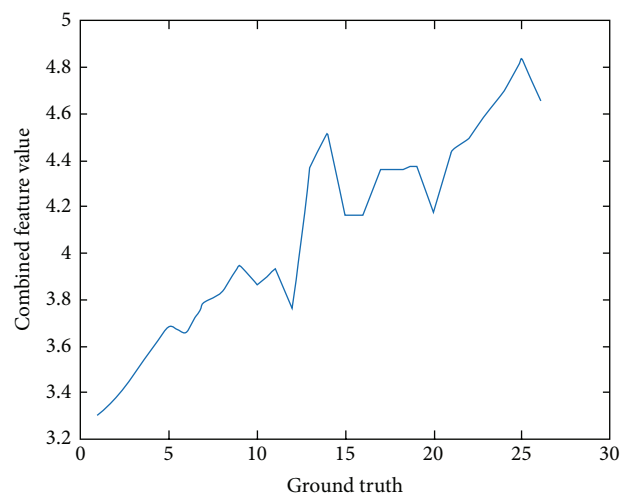


Figure 7: Combined feature performance.

5. Modeling

Reliable, high-fidelity fault growth models form the foundation for accurate and robust detection and failure prediction. A suitable modeling framework assists in the development, testing, and evaluation of detection and prediction algorithms. It may be employed to generate data for data-driven methods to diagnostics/prognostics and test and validate routines for data processing tool development, among others. The flexibility provided by a simulation platform, housing appropriate detection and progression models, is a unique attribute in the study of how fault processes are initiated and propagating so that corrective action can be taken before a catastrophic event occurs. The objective of the modeling effort is to develop, test, and evaluate novel fault initiation and progression models that will assist in the design and implementation of "smart" sensors and sensing modalities for critical machine systems.

5.1. Modeling-Symbolic Regression. Fault detection and prediction algorithms rely on data, a model of the degradation process and an estimation method that, given the current

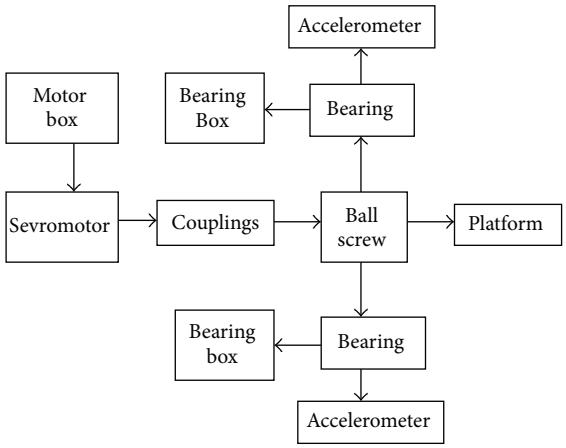


Figure 8: Structure model of the platform.

state of the system, predicts its evolution over the next time step. Such models are typically based on first principles while others are built on the basis of data. We exploit in this effort a modeling framework called symbolic regression. It is a type of regression analysis that searches the space of mathematical expressions to find the model that best fits a given dataset, in terms of both accuracy and simplicity. No particular model is provided as a starting point to the algorithm. Instead, initial expressions are formed by randomly combining mathematical building blocks such as mathematical operators, analytic functions, constants, and state variables. (Usually, a subset of these primitives will be specified by the operator, but that is not a requirement of the technique.) New equations are then formed by combining previous ones, using genetic programming. In linear regression, the dependent variable is a linear combination of the parameters (but need not be linear in the independent variables). Nonlinear symbolic regression and other regression techniques incorporating uncertainty are based on similar principles. The structure model of the platform is shown in Figure 8.

The functional model of the platform is shown in Figure 9.

During real time operation, model parameters are tuned as new data is streaming in and an error between the estimated state and the detected state is exploited in an optimization scheme to arrive at the "best" model for further prediction purposes. Eureqa is a useful tool to implement the symbolic regression routine. Exploiting the test data, Eureqa searches to find the model (both model structure and model parameters) that best fit a given data set. We arrive at the following parametric model:

$$\text{Feature} = 8.11 \times 10^3 + 0.262x^3. \tag{2}$$

x implies cycle times; here we use x as time, t.

Figure 10(a) shows the result from the application of Eureqa. Figure 10(b) shows a plot of $(8.11 \times 10^3 + 0.262x^3)$ and the error as the process proceeds. Figure 10(c) shows the accuracy versus complexity plot. By trading off accuracy for complexity, the best model from the list generated by Eureqa is $(8.11 \times 10^3 + 0.262x^3)$.

A sufficient data set is obtained by using the following formula to enrich the data:

$$\text{Feature} = g(x[k]) + n(0, S), \tag{3}$$

where $\delta(x)$ is the model fitted by regression, S is the standard deviation, N is the number of samples, and x is a variable:

$$S = \sqrt{\frac{\sum_{i=1}^{N} [1/i - \delta(x_i)]^2}{N-1}}. \tag{4}$$

6. The Particle Filtering Framework for Fault Diagnosis and Failure Prognosis

Particle filtering is an emerging and powerful methodology for sequential signal processing based on the concepts of Bayesian theory and Sequential Importance Sampling (SIS). Particle filtering is very suitable for nonlinear systems or in the presence of non-Gaussian process/observation noise. For this approach, both diagnosis and prognosis rely upon estimating the current value of a fault/degradation dimension, as well as other important parameters, and use a set of observations (or measurements) for this purpose. This research team has pioneered the introduction of particle filtering techniques into fault diagnosis and failure prognosis [1, 2]. The team has also demonstrated how these techniques can be combined with traditional artificial intelligence methods in a synergistic way [3]. The success of this novel model-based approach has been demonstrated in a number of diverse application domains from rotorcraft critical components to electrical systems, environmental control systems, and high power amplifiers [4]. Figure 11 depicts the flow of the component modules comprising the health management architecture.

The underlying principle of the methodology is the approximation of the conditional state probability distribution $p(z_k/x_k)$ using a swarm of points called particles and a set of weights associated with them representing the discrete probability masses. Particles can be generated and recursively updated easily given a nonlinear process model (which describes the evolution in time of the system under analysis), a measurement model, a set of available measurements $z_{1,k} = (z_1, \ldots, z_k)$, and an initial estimation for the state PDF $p(x_0)$, as shown in the following equations:

$$\begin{aligned} x_k &= f_k(x_{k-1}, \omega_k) \longleftrightarrow p(x_k \mid x_{k-1}), \\ z_k &= h_k(x_k, v_k) \longleftrightarrow p(z_k \mid x_k). \end{aligned} \tag{5}$$

As in every Bayesian estimation problem, the estimation process can be achieved into two steps, namely, prediction and filtering. On the one hand, prediction uses both knowledge of the previous state estimation and the process model to generate the a priori state PDF estimation for the next time instant, as is shown in the following expression:

$$\begin{aligned} &p(x_k \mid z_{1:k-1}) \\ &\quad = \int p(x_k \mid x_{k-1})\, p(x_{k-1} \mid z_{1:k-1})\, dx_{k-1}. \end{aligned} \tag{6}$$

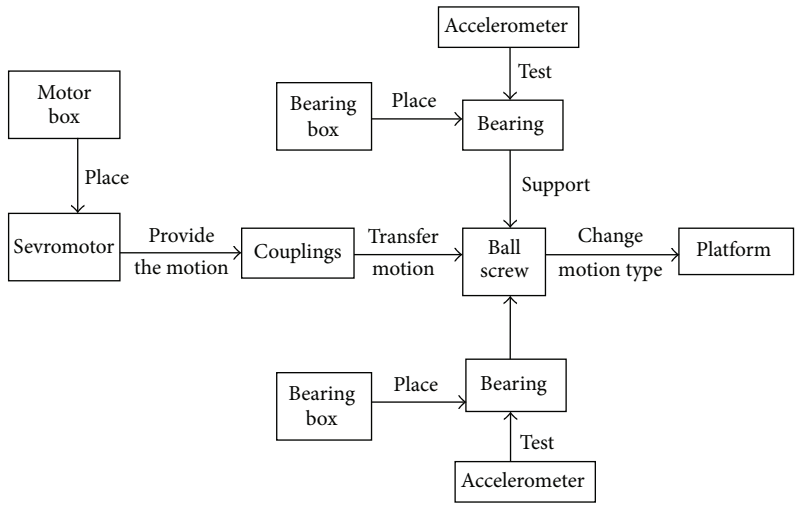


Figure 9: Functional model of the platform.

On the other hand, the filtering step considers the current observation z_k and the a priori state PDF to generate the a posteriori state PDF by using Bayes' formula:

$$p(x_k \mid z_{1:k}) = \frac{p(z_k \mid x_k)\, p(x_k \mid z_{1:k-1})}{p(z_k \mid z_{1:k-1})}. \tag{7}$$

The actual distributions then would be approximated by a set of samples and the corresponding normalized importance weights $\widetilde{w}_k^i = \widetilde{w}_k(x_{0:k}^i)$ for the ith sample:

$$p(x_k \mid z_{1:k}) \approx \sum_{i=1}^{N} \widetilde{w}_k\left(x_{0:k}^i\right) \delta\left(x_{0:k} - x_{0:k}^i\right), \tag{8}$$

where the update for the importance weights is given by

$$w_k = w_{k-1} \frac{p(z_k \mid x_k)\, P(x_k \mid x_{k-1})}{P(x_k \mid x_{0:k-1}, z_{1:k})}. \tag{9}$$

Fault Diagnosis. The proposed particle-filter-based diagnosis framework aims to accomplish the tasks of fault detection and identification, under general assumptions of non-Gaussian noise structures and nonlinearities in process dynamic models, using a reduced particle population to represent the state pdf [5]. A compromise between model-based and data-driven techniques is accomplished by the use of a particle filter-based module built upon the nonlinear dynamic state model:

$$\begin{aligned} x_d(t+1) &= f_b(x_d(t), n(t)), \\ x_c(t+1) &= f_t(x_d(t), x_c(t), w(t)), \\ \text{Features}(t) &= h_t(x_d(t), x_c(t), v(t)), \end{aligned} \tag{10}$$

where f_b, f_t and h_t are nonlinear mappings, $x_d(t)$ is a collection of Boolean states associated with the presence of a particular operating condition in the system (normal operation, fault type #1, #2, etc.), $x_c(t)$ is a set of continuous-valued states that describe the evolution of the system given those operating conditions, and $w(t)$ and $v(t)$ are non-Gaussian distributions that characterize the process and feature noise signals, respectively. At any given instant of time, this framework provides an estimate of the probability masses associated with each fault mode, as well as a pdf estimate for meaningful physical variables in the system. Once this information is available within the diagnostic module, it is conveniently processed to generate proper fault alarms and to inform about the statistical confidence of the detection routine. Customer specifications are translated into acceptable margins for types I and II errors in the detection routine. The algorithm itself will indicate when type II error (false negatives) has decreased to the desired level. Typical results of the diagnostic algorithm are shown in Figure 12.

Failure Prognosis. The evolution in time of the fault dimension may be described through the state equation:

$$\begin{aligned} x_1(t+1) &= x_1(t) + x_2(t) \cdot F(x(t), t, U), w_1(t), \\ x_2(t+1) &= x_2(t) + w_2(t), \end{aligned} \tag{11}$$

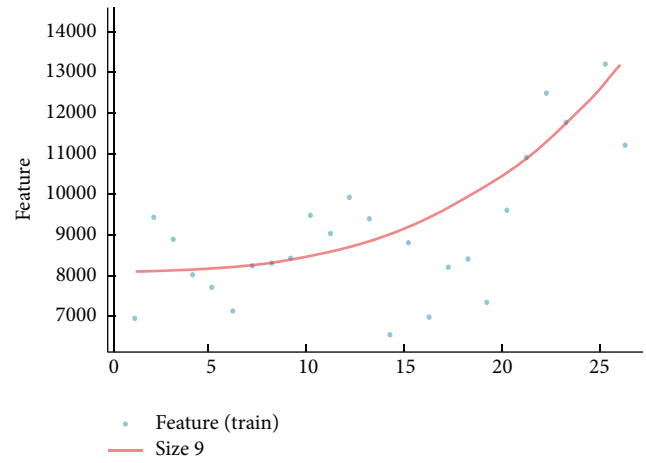
where $x_1(t)$ is a state representing the fault dimension under analysis, $x_2(t)$ is a state associated with an unknown model parameter, U are external inputs to the system (load profile, etc.), $F(x(t), t, U)$ is a general time-varying nonlinear function, and $w_1(t)$ and $w_2(t)$ are white noises (not necessarily Gaussian) [5]. The nonlinear function may represent a model based on first principles, a neural network, or even a fuzzy system. Long term predictions are generated on the basis of the current state pdf estimate, and using kernel functions to reconstruct the state pdf estimate for future time instants,

$$\begin{aligned} &\widehat{p}(x_{t+k} \mid \widehat{x}_{t+k-1}) \\ &\quad \approx \sum_{i=1}^{N} w_{t+k-1}^{(i)} K\left(x_{t+k} - E\left[x_{t+k}^{(i)} \mid \widehat{x}_{t+k-1}^{(i)}\right]\right), \end{aligned} \tag{12}$$

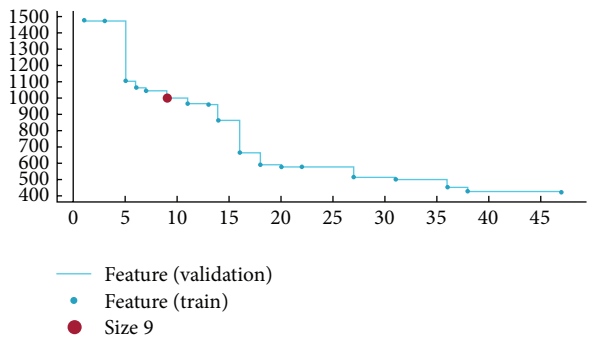
where $K(\cdot)$ is a kernel density function, which may correspond to the process noise pdf, a Gaussian kernel, or a rescaled version of the Epanechnikov kernel.

Size	Fit	Solution
1	1.000	Feature = $8.89e3$
5	0.748	Feature = $6.54e3 + 198x$
3	0.998	Feature = $8.8e3 + x$
6	0.721	Feature = $\frac{4.46e5}{62.8 - x}$
7	0.707	Feature = $7.81e3 + 6.4x^2$
9	0.676	Feature = $8.11e3 + 0.262x^3$
11	0.653	Feature = $8.2e3 + 0.0108x^4$
13	0.649	Feature = $8.27e3 + 0.000415x^5$
14	0.584	Feature = $7.01e3 + 159x - 1.15e3\sin(175x)$
16	0.449	Feature = $7.13e3 + 139x + 84.4x\cos(144x)$
18	0.399	Feature = $7.15e3 + 131x + 5.54x^2\cos(144x)$
20	0.390	Feature = $7.98e3 + 4.37x^2 + 5.25x^2\cos(144x)$
22	0.390	Feature = $7.97e3 + x + 4.35x^2 + 5.27x^2\cos(144x)$

(a)



(b)



(c)

Figure 10: Symbolic regression using Eureka.

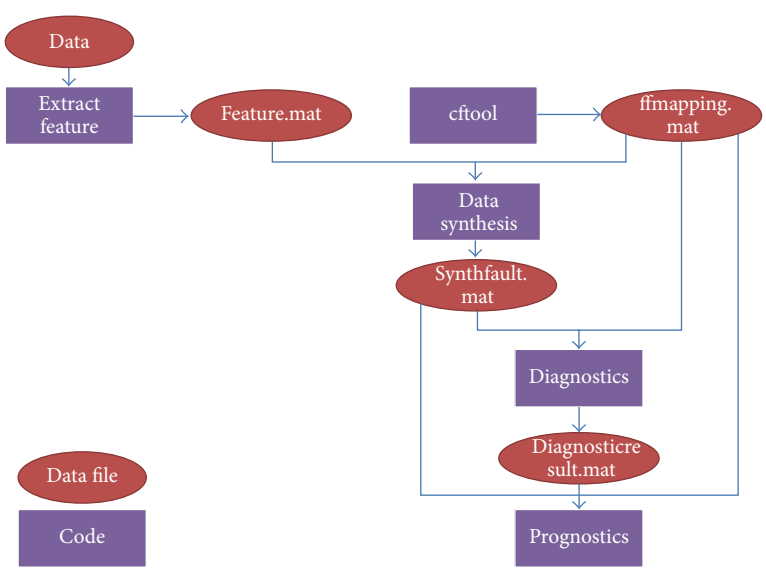


Figure 11: The health management architecture.

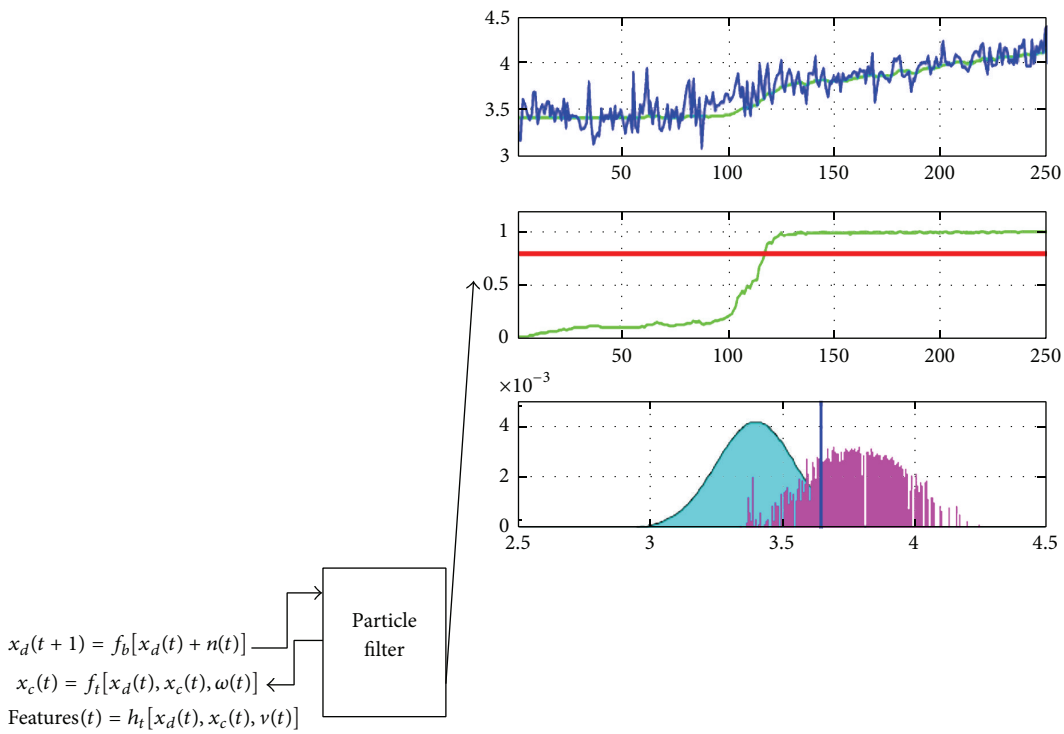


Figure 12: Typical results of the diagnostic algorithm.

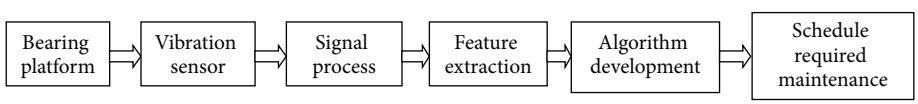


Figure 13: An integrated approach for fault diagnosis and failure prognosis.

7. Results

Using the machinery data, we determine its health status and hence its current fault/failure condition. The main modules of an integrated approach to fault diagnosis and failure prognosis are shown in Figure 13. It includes experiment design, fault classification, data collection, signal process, feature extraction, algorithm development, and schedules required maintenance.

The resulting predicted state pdf contains critical information about the evolution of the fault dimension over time. This information is condensed through the computation of statistics (expectations, 90% confidence intervals), the Time-of-Failure (ToF), or the Remaining Useful Life (RUL) of the faulty system. Figure 16 depicts a typical prognostic configuration.

Data from simulation as well as the experiments were used to implement and evaluate the fault diagnosis and failure prognosis routines. Performance metrics were applied at all levels of the architecture (data analysis/feature extraction, diagnosis, and prognosis).

7.1. Fault Predict. Using Eureka software and the model, we fitted a curve shown in Figure 14. The points represent 27 features extracted in 27 cycles. The red line is the model fitted by the feature. According to the deviation of the data itself, the model was enriched and data were generated with noise injected into the model. The data is shown in Figure 15. The red line is the model and the blue curve is the generated data.

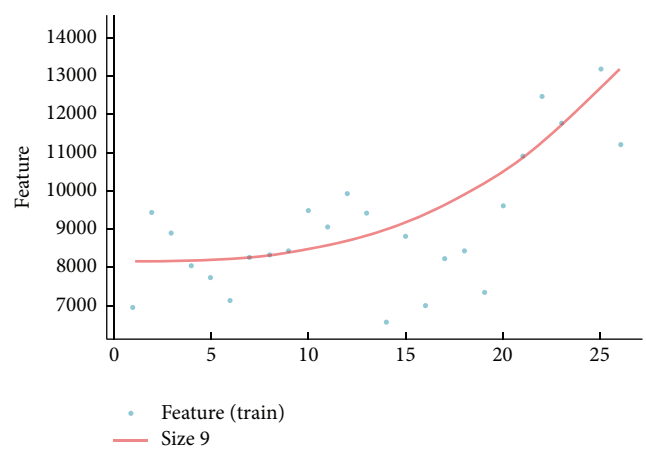


Figure 14: Feature extraction curve fitting.

Figures 16 and 17 show the results obtained when the proposed approach is applied to the problem of fault detection/prediction due to degradation of the bearing lubrication.

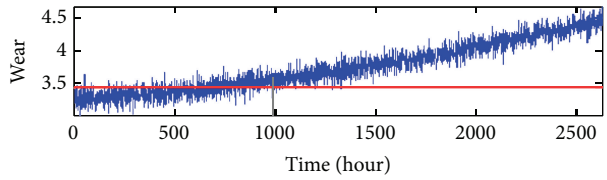


Figure 15: Enrich feature extraction curve fitting.

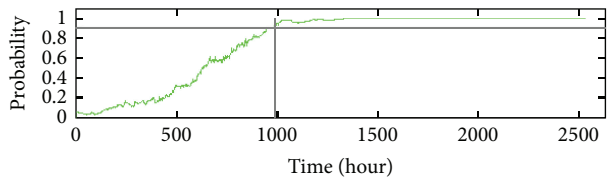


Figure 16: Trending of the selected feature.

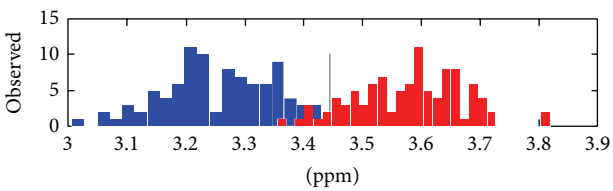


Figure 17: Baseline versus current distributions.

In Figure 19, the feature value is depicted as a function of time. It exhibits the fault growth (green line) and the set threshold, that is, the fault value at a predetermined level (red line). One hundred points in each time window are selected. The blue histogram represents the normal pdf. The blue line is selected as the threshold. 5% of the normal pdf is on the right of the threshold implying that type I error is 5%. The red histogram represents the current pdf. 6% of the current pdf is on the left of the threshold. So, type II error is 6%. Thus, the corresponding confidence in the declaration of the fault condition is 90% meaning that after 998 cycles, there is 94% possibility that the test platform is experiencing a fault.

7.2. Failure Prognosis. For prediction purposes the model is tuned first, as shown in Figure 18. The distance of the bearing run in miles corresponds to time in this case. 5000 particles are used for resampling purposes. The red line corresponds to the measurement data. The blue dot line represents the estimated data.

In Figure 19, the red "line" is the threshold value. The threshold or hazard zone is actually a probability density function (pdf) set by the user and representing the latter's conception from past experiences of the fault dimension a failure event is imminent and must be corrected. Multiple hazard zones may be specified. The histogram along the time/mileage axis (approximate pdf) represents the time or miles traveled to failure. Statistics, such as the mean and standard deviation, are computed to provide the user with useful information, that is, when action should be taken to repair the failing component (earlier or later than the mean time depending on the risk the user is willing to take under the prevailing circumstances). Appropriate

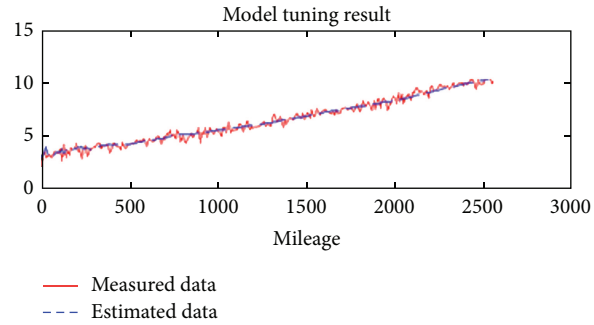


Figure 18: Model tuning.

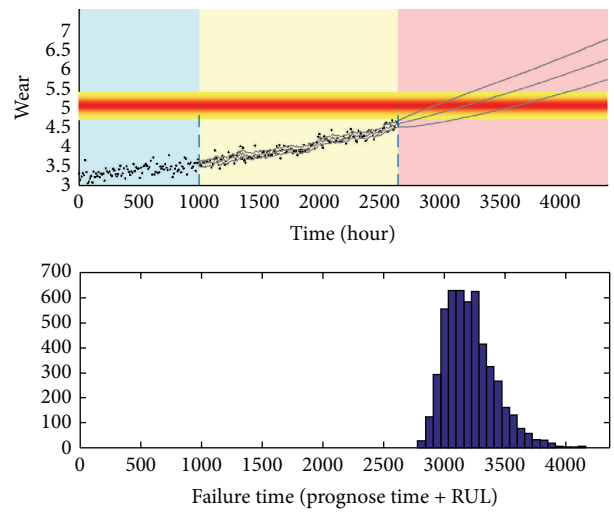


Figure 19: Failure prognosis.

performance metrics are defined and employed relating to the prediction accuracy [1].

8. Conclusions

Prognosis and health management is an important subject worthy of study for automated machining processes. This paper introduces a framework for the process, fault type, and assessment and testing of a critical component from an automated machining process based on novel particle filtering algorithms. Fault diagnosis has been shown to result in detection and isolation of a faulty machine component with user specified accuracy (confidence) and given false alarm rate. The prognostic approach, based on streaming data, an appropriate feature vector, a data-driven symbolic regression model, and a Bayesian estimation process-particle filtering, has been demonstrate to provide accurate results for the test case under consideration. As a result, the maintainer is provided with useful information as to when corrective action is required considering risk issues. The application example employs real fault/failure data derived from a seeded fault test in a bearing of the automated machining process. It provides excellent insight about the effect of model inaccuracies and customer specifications in the algorithm performance.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] G. Vachtsevanos, F. L. Lewis, M. Roemer, A. Hess, and B. Q. Wu, *Intelligent Fault Diagnosis and Prognosis for Engineering System*, Wiley, New York, NY, USA, 2006.

[2] M. E. Orchard and G. J. Vachtsevanos, "A particle-filtering approach for on-line fault diagnosis and failure prognosis," *Transactions of the Institute of Measurement and Control*, vol. 31, no. 3-4, pp. 221–246, 2009.

[3] Narasimhan, "A conceptual model of the alignment-performance link in PSM," 2005.

[4] X. Qi, J. Qi, D. Theilliol et al., "A review on fault diagnosis and fault tolerant control methods for single-rotor aerial vehicles," *Journal of Intelligent & Robotic Systems*, vol. 73, no. 1–4, pp. 535–555, 2014.

[5] M. Orchard, G. Kacprzynski, K. Goebel, B. Saha, and G. Vachtsevanos, "Advances in uncertainty representation and management for particle filtering applied to prognostics," in *Proceedings of the International Conference on Prognostics and Health Management (PHM '08)*, pp. 1–6, IEEE, Denver, Colo, USA, October 2008.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 548050, 13 pages
http://dx.doi.org/10.1155/2015/548050

Research Article

A Unified Approach to Nonlinear Dynamic Inversion Control with Parameter Determination by Eigenvalue Assignment

Yu-Chi Wang, Donglong Sheu, and Chin-E Lin

Department of Aeronautics and Astronautics, National Cheng Kung University, Tainan 701, Taiwan

Correspondence should be addressed to Chin-E Lin; chinelin@mail.ncku.edu.tw

Received 8 September 2015; Revised 13 November 2015; Accepted 16 November 2015

Academic Editor: Rongwei Guo

This paper presents a unified approach to nonlinear dynamic inversion control algorithm with the parameters for desired dynamics determined by using an eigenvalue assignment method, which may be applied in a very straightforward and convenient way. By using this method, it is not necessary to transform the nonlinear equations into linear equations by feedback linearization before beginning control designs. The applications of this method are not limited to affine nonlinear control systems or limited to minimum phase problems if the eigenvalues of error dynamics are carefully assigned so that the desired dynamics is stable. The control design by using this method is shown to be robust to modeling uncertainties. To validate the theory, the design of a UAV control system is presented as an example. Numerical simulations show the performance of the design to be quite remarkable.

1. Introduction

In the development of high-performance aircraft, control difficulties may be encountered over some parts of flight envelope. These difficulties arise from highly nonlinear aerodynamic properties [1] in some flight conditions. In order to solve these control difficulties, nonlinear controllers are required for high-performance aircraft.

Among many control methods, nonlinear dynamic inversion (NDI) is very popular and has been widely studied for flight control designs (e.g., [2, 3]). NDI-based control systems are usually divided into fast- and slow-loop control subsystems according to the multiple time-scale method [4]. In each subsystem, Lie derivatives [5] are used to transform the nonlinear equations into linear equations. Then, linear control design methods can be employed and the control inputs are obtained by converting the linear system control variables into the original coordinates. However, the control systems obtained by feedback linearization [6] may have nonminimum phase problems for affine or nonaffine nonlinear system [7] and robust issues in case of model mismatch. A typical nonminimum phase problem may be found in flight dynamics where the altitude-elevator transfer function usually has a right-half zero. The internal state control [8] is often used to overcome these nonminimum phase problems. In addition, the fuzzy logic control [9] was also applied for solving these kinds of problems. Furthermore, to overcome the robust problems, μ-analysis [10] and H_∞ method [11] were applied. Specially, incremental NDI (INDI) [12] was used to increase the robustness to aerodynamic uncertainties by calculating the control surface deflection changes instead of giving inputs directly.

To circumvent some aforementioned robust problems, an adaptive nonlinear model inversion control [13] was introduced, in which the design concept is similar to the conventional NDI yet without linearizing the nonlinear system. The model inversion method replaces the motion rates with a P-form or PI-form desired dynamics to negate the original dynamics. The choice of parameters in the desired dynamics is based on the bandwidth of response and time scales. The effects of different types of desired dynamics on the resulted control system were discussed [14].

Although the aforementioned NDI approaches are successful in many flight control system designs [14–16] over a large part of the flight envelope, the systems of dynamics in general have to be separated into several subsystems according to the rates of response. There are many cases in which the fast rate and the slow rate might not be

distinguished so clearly however. Also, although the pole assignment method had been introduced to determine the parameters in the desired dynamics, very often the control system must first be transformed into a standard feedback control form from which the standard eigenvalue assignment method can be applied.

In consideration of the aforementioned problems existing in the current literature on control designs, a unified approach to nonlinear dynamic inversion control is proposed in this paper. The equations of motion will not be necessary to be separated into fast rate and slow rate groups, nor will they be limited to an affine system. Feedback linearizations will not be required to transform the nonlinear equations into linear equations. Nonminimum phase problems are solved by eigenvalue assignments for error dynamics. An iterative method for determining the parameters of the desired dynamics from the assigned eigenvalues of error dynamics is proposed. Analysis of robustness to model uncertainties or disturbances is conducted. This method will be ready for design without simplifying the system of equations based on physical insights once the governing dynamic equations are established and the state variables to be tracked are selected. The theory is to be developed in detail in the following sections. A UAV is introduced and its control system is designed with the developed method. Numerical simulations are conducted to validate this method.

2. A Unified Approach to Nonlinear Dynamic Inversion Control

2.1. Nonlinear Dynamic Inversion Control. In general, dynamic equations of motion with control inputs can be expressed by

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}), \tag{1}$$

where $\mathbf{x} \in R^n$ is the state vector, $\mathbf{u} \in R^m$ $(m < n)$ the control vector, and $\mathbf{f}$ the nonlinear function representing the model of dynamics with controls. By extending the concept of dynamic inversion (DI) [4], the control vector $\mathbf{u}$ can be assumed to be computed from

$$\dot{\mathbf{x}}_d = \mathbf{f}(\mathbf{x}_d, \mathbf{u}), \tag{2}$$

where $\mathbf{x}_d \in R^n$ is a desired state vector with its rate of change being designated.

In this paper, the desired dynamics is designated as a set of stable first-order differential equations:

$$\dot{\mathbf{x}}_d = \mathbf{\Omega}(\mathbf{x}_d - \mathbf{x}), \tag{3}$$

where $\mathbf{\Omega} \in R^{n\times n}$ represents a constant matrix with n independent parameters which can be chosen. Substituting (3) into (2) yields

$$\mathbf{\Omega}(\mathbf{x}_d - \mathbf{x}) = \mathbf{f}(\mathbf{x}_d, \mathbf{u}) \tag{4}$$

which constitutes a set of n algebraic equations. Since $m < n$, obviously, $\mathbf{u}$ cannot satisfy (4) if all elements of $\mathbf{x}_d$ are to be designated. It means that only part of $\mathbf{x}_d$ can be designated. So let $\mathbf{x}_d$ be divided into two groups, say $\mathbf{x}_c \in R^m$ and $\mathbf{x}_r \in R^{n-m}$, where $\mathbf{x}_c$ contains the state variables which are to be controlled or designated and $\mathbf{x}_r$ the residual ones. Both $\mathbf{u}$ and $\mathbf{x}_r$ constitute n unknown variables which are to be determined from (4). To solve a set of nonlinear algebraic equations, the Newton-Raphson iteration method can be employed as follows:

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{x}_r \end{bmatrix}_{k+1} = \begin{bmatrix} \mathbf{u} \\ \mathbf{x}_r \end{bmatrix}_k - \begin{bmatrix} \dfrac{\partial \mathbf{f}_{\text{err}}}{\partial \mathbf{u}} & \dfrac{\partial \mathbf{f}_{\text{err}}}{\partial \mathbf{x}_r} \end{bmatrix}^{-1} \mathbf{f}_{\text{err}}, \tag{5}$$

where $\mathbf{f}_{\text{err}} \triangleq \mathbf{f}(\mathbf{x}_d, \mathbf{u}) - \mathbf{\Omega}(\mathbf{x}_d - \mathbf{x})$ and $k = 0, 1, 2, \ldots$ represents the iteration number.

2.2. Parameter Determination. Now, a question arises whether the state vector $\mathbf{x}$ will asymptotically follow the desired vector $\mathbf{x}_d$ if $\mathbf{u}$ is determined from (4) and substituted in (1). In order to answer it, let (1) and (4) be examined more carefully as follows.

Subtracting (4) from (1) yields

$$\dot{\mathbf{x}} - \mathbf{\Omega}(\mathbf{x}_d - \mathbf{x}) = \mathbf{f}(\mathbf{x}, \mathbf{u}) - \mathbf{f}(\mathbf{x}_d, \mathbf{u}). \tag{6}$$

If $\mathbf{x}$ is very close to $\mathbf{x}_d$, then the above equation can be linearized as follows:

$$\dot{\mathbf{x}} - \mathbf{\Omega}(\mathbf{x}_d - \mathbf{x}) \approx \mathbf{f}_{\mathbf{x}}(\mathbf{x}_d, \mathbf{u})(\mathbf{x} - \mathbf{x}_d). \tag{7}$$

With the concept that the desired variables $\mathbf{x}_d$ are near constant, say $\dot{\mathbf{x}}_d \approx \mathbf{0}$, (7) can be rewritten as

$$\dot{\mathbf{x}} - \dot{\mathbf{x}}_d \approx [-\mathbf{\Omega} + \mathbf{f}_{\mathbf{x}}(\mathbf{x}_d, \mathbf{u})](\mathbf{x} - \mathbf{x}_d). \tag{8}$$

Defining an error vector $\mathbf{e} = \mathbf{x} - \mathbf{x}_d$ and replacing the approximate sign with the equal sign lead (8) to

$$\dot{\mathbf{e}} = [-\mathbf{\Omega} + \mathbf{f}_{\mathbf{x}}(\mathbf{x}_d, \mathbf{u})]\,\mathbf{e}. \tag{9}$$

Equation (9) is a set of error dynamics in which the error vector $\mathbf{e}$ will vanish eventually if all the real parts of the eigenvalues of $[-\mathbf{\Omega} + \mathbf{f}_x(\mathbf{x}_d, \mathbf{u})]$ are negative. This is possible if $\mathbf{\Omega}$ is chosen appropriately. It means that the state vector $\mathbf{x}$ can approach the desired vector $\mathbf{x}_d$ once $\mathbf{e}$ approaches $\mathbf{0}$. Recall that $\mathbf{x}_d$ contains $\mathbf{x}_c$, a state vector to be tracked.

Notice that the eigenvalue λ in (9) can be determined by

$$\left|\lambda\mathbf{I} + \mathbf{\Omega} - \mathbf{f}_{\mathbf{x}}(\mathbf{x}_d, \mathbf{u})\right| = 0, \tag{10}$$

where $\mathbf{I}$ is an identity matrix. For simplicity, if $\mathbf{\Omega}$ is diagonal and all elements are the same, say, $\mathbf{\Omega} = \sigma\mathbf{I}$, then (10) can be rewritten as

$$\left|(\lambda + \sigma)\,\mathbf{I} - \mathbf{f}_{\mathbf{x}}(\mathbf{x}_d, \mathbf{u})\right| = 0. \tag{11}$$

In order to make $\mathbf{e}$ bounded, σ can be so chosen that

$$\text{real}(\lambda_i + \sigma) < 0 \tag{12}$$

for all λ_i $(i = 1, 2, \ldots, n)$. Although this method is simple, the resulting σ may be unnecessarily large.

For more general cases, $\mathbf{\Omega}$ contains a set of parameters, σ_k, $(k = 1, 2, 3, \ldots, n)$, which may be chosen to determine the eigenvalues and eigenvectors of $-\mathbf{\Omega} + \mathbf{f}_{\mathbf{x}}(\mathbf{x}_d, \mathbf{u})$. Recall that (3) must be a stable model and, therefore, the simplest way of constructing $\mathbf{\Omega}$ is to let all its elements vanish except those at the diagonal, which are assumed to be $\Omega_{kk} = \sigma_k > 0$. In fact, some off-diagonal elements can also be allowed to exist. For example, let $\Omega_{kk} = \Omega_{ll} = \sigma_k > 0$, and $\Omega_{kl}^2 = \Omega_{lk}^2 = \sigma_l^2 < \sigma_k^2$. It is trivial to prove that the latter case is also a stable model.

Now, assume that the ith $(i = 1, 2, 3, \ldots, n)$ eigenvalue and eigenvector are $\overline{\lambda}_i$ and $\overline{\mathbf{e}}_i$, respectively. Accordingly,

$$\left[-\overline{\mathbf{\Omega}} + \mathbf{f}_{\mathbf{x}}\left(\mathbf{x}_d, \mathbf{u}\right)\right] \overline{\mathbf{e}}_i = \overline{\lambda}_i \overline{\mathbf{e}}_i. \tag{13}$$

In general, the eigenvector $\overline{\mathbf{e}}_i$ can be normalized so that

$$\overline{\mathbf{e}}_i^T \overline{\mathbf{e}}_i = 1. \tag{14}$$

However, if the eigenvalue λ_i is desired rather than $\overline{\lambda}_i$, then it is necessary to adjust the parameters in $\mathbf{\Omega}$. Assume that an increment $(\partial\mathbf{\Omega}/\partial\sigma_k)\Delta\sigma_k$ to $\overline{\mathbf{\Omega}}$ is required. Then

$$\left[-\overline{\mathbf{\Omega}} - \frac{\partial\mathbf{\Omega}}{\partial\sigma_k}\Delta\sigma_k + \mathbf{f}_{\mathbf{x}}\left(\mathbf{x}_d, \mathbf{u}\right)\right] \mathbf{e}_i = \lambda_i \mathbf{e}_i, \tag{15}$$

where λ_i and $\mathbf{e}_i$ are assumed to be approximated to $\overline{\lambda}_i + (\partial\lambda_i/\partial\sigma_k)\Delta\sigma_k$ and $\overline{\mathbf{e}}_i + (\partial\mathbf{e}_i/\partial\sigma_k)\Delta\sigma_k$, respectively, and $\mathbf{e}_i$ is also normalized so that

$$\left(\overline{\mathbf{e}}_i + \frac{\partial\mathbf{e}_i}{\partial\sigma_k}\Delta\sigma_k\right)^T \left(\overline{\mathbf{e}}_i + \frac{\partial\mathbf{e}_i}{\partial\sigma_k}\Delta\sigma_k\right) = 1. \tag{16}$$

Accordingly, the derivative of (13) with respect to σ_k results in

$$\left[-\frac{\partial\mathbf{\Omega}}{\partial\sigma_k}\right] \overline{\mathbf{e}}_i + \left[-\overline{\mathbf{\Omega}} + \mathbf{f}_{\mathbf{x}}\left(\mathbf{x}_d, \mathbf{u}\right)\right] \frac{\partial\mathbf{e}_i}{\partial\sigma_k} = \frac{\partial\lambda_i}{\partial\sigma_k}\overline{\mathbf{e}}_i + \overline{\lambda}_i \frac{\partial\mathbf{e}_i}{\partial\sigma_k} \tag{17}$$

and the derivative of (14) becomes

$$\overline{\mathbf{e}}_i^T \frac{\partial\mathbf{e}_i}{\partial\sigma_k} = 0. \tag{18}$$

Equations (17) and (18) can be rearranged to

$$\begin{bmatrix} \overline{\mathbf{e}}_i & \overline{\lambda}_i\mathbf{I} + \overline{\mathbf{\Omega}} - \mathbf{f}_{\mathbf{x}}\left(\mathbf{x}_d, \mathbf{u}\right) \\ 0 & \overline{\mathbf{e}}_i^T \end{bmatrix} \begin{Bmatrix} \dfrac{\partial\lambda_i}{\partial\sigma_k} \\ \dfrac{\partial\mathbf{e}_i}{\partial\sigma_k} \end{Bmatrix} = \begin{Bmatrix} -\dfrac{\partial\mathbf{\Omega}}{\partial\sigma_k}\overline{\mathbf{e}}_i \\ 0 \end{Bmatrix} \tag{19}$$

from which $\partial\lambda_i/\partial\sigma_k$ can be determined. Note that $\partial\lambda_i/\partial\sigma_k$ represents the variation of λ_i with respect to σ_k. With all $\partial\lambda_i/\partial\sigma_k$, $(k = 1, 2, 3, \ldots, n)$, the ith revised eigenvalue should be

$$\lambda_i = \overline{\lambda}_i + \sum \frac{\partial\lambda_i}{\partial\sigma_k}\Delta\sigma_k, \quad (i = 1, 2, 3, \ldots, n). \tag{20}$$

Recall that, in this equation, $\overline{\lambda}_i$ represents the ith eigenvalue of $[-\overline{\mathbf{\Omega}} + \mathbf{f}_{\mathbf{x}}(\mathbf{x}_d, \mathbf{u})]$, where $\overline{\mathbf{\Omega}}$ contains the parameters $\overline{\sigma}_k$, $(k = 1, 2, 3, \ldots, n)$. Now since λ_i is the ith desired eigenvalue, (20) in fact becomes a set of n simultaneous equations from which $\Delta\sigma_k$ can be solved. With $\overline{\sigma}_k$ being the initial guess, revisions for parameters can be made by

$$\sigma_k = \overline{\sigma}_k + \Delta\sigma_k. \tag{21}$$

Since λ_i, $(i = 1, 2, 3, \ldots, n)$ are nonlinear function of σ_k, $(k = 1, 2, 3, \ldots, n)$, iterative computations of (9) and (19)–(21) are required in order to obtain a set of convergent σ_k. To make it clear, the iteration procedures are summarized as follows.

(1) Let the desired eigenvalues be λ_i $(i = 1, 2, 3, \ldots, n)$ which are all distinct.

(2) Guess a set of parameters $\overline{\sigma}_k$ $(k = 1, 2, 3, \ldots, n)$.

(3) Determine the eigenvalues and eigenvectors of (9). Denote the ith eigenvalue and eigenvector as $\overline{\lambda}_i$ and $\overline{\mathbf{e}}_i$, respectively.

(4) With each $\overline{\sigma}_k$ $(k = 1, 2, 3, \ldots, n)$, determine $\partial\lambda_i/\partial\sigma_k$ $(i = 1, 2, 3, \ldots, n)$ from (19).

(5) Determine $\Delta\sigma_k$ $(k = 1, 2, 3, \ldots, n)$ from (20). If $|\Delta\sigma_k|$ is less than a preset small value, then stop; else continue with the next step.

(6) Determine σ_k from (21).

(7) Replace $\overline{\sigma}_k$ with σ_k and go to step 3.

At this point, it must be mentioned that the iterations will converge only if the initial guesses are very close to the true answers. It is also emphasized that the desired eigenvalues λ_i must be chosen to lie on the left-half s-plane and the resulting parameters in $\mathbf{\Omega}$ must satisfy the stable model requirements in (3).

2.3. Robust Analysis. The nonlinear dynamic inversion control design developed so far is based on a nominal dynamical model. In the real world, however, there always exist some modeling uncertainties which cannot be determined in advance. In order to check if the control design is robust, let the actual dynamical model be represented by

$$\dot{\mathbf{x}} = \mathbf{f}'\left(\mathbf{x}, \mathbf{u}\right). \tag{22}$$

With this assumption, now (6) can be modified to

$$\dot{\mathbf{x}} - \mathbf{\Omega}\left(\mathbf{x}_d - \mathbf{x}\right) = \mathbf{f}'\left(\mathbf{x}, \mathbf{u}\right) - \mathbf{f}\left(\mathbf{x}_d, \mathbf{u}\right) \tag{23}$$

which can be rewritten as follows:

$$\begin{aligned} \dot{\mathbf{x}} + \mathbf{\Omega}\mathbf{x} = \mathbf{\Omega}\mathbf{x}_d + \Big[&\mathbf{f}_{\mathbf{x}}\left(\mathbf{x}_d, \mathbf{u}\right)\left(\mathbf{x} - \mathbf{x}_d\right) + \Delta\mathbf{f}_n\left(\mathbf{x}, \mathbf{x}_d, \mathbf{u}\right) \\ &+ \mathbf{f}'\left(\mathbf{x}, \mathbf{u}\right) - \mathbf{f}\left(\mathbf{x}, \mathbf{u}\right)\Big], \end{aligned} \tag{24}$$

where $\Delta\mathbf{f}_n(\mathbf{x},\mathbf{x}_d,\mathbf{u})$ denotes the nonlinear part of $\mathbf{f}(\mathbf{x},\mathbf{u}) - \mathbf{f}(\mathbf{x}_d,\mathbf{u})$. The solution of the above equation can be represented as follows:

$$\begin{aligned}\mathbf{x}(t) &= e^{-\mathbf{\Omega}t}\mathbf{x}_0 + \left(\mathbf{I} - e^{-\mathbf{\Omega}t}\right)\mathbf{x}_d \\ &+ \int_0^t e^{-\mathbf{\Omega}(t-\tau)}\left\{\mathbf{f}_{\mathbf{x}}\left(\mathbf{x}_d,\mathbf{u}(\tau)\right)\left(\mathbf{x}(\tau)-\mathbf{x}_d\right)\right. \\ &+ \Delta\mathbf{f}_n\left(\mathbf{x}(\tau),\mathbf{x}_d,\mathbf{u}(\tau)\right) + \mathbf{f}'\left(\mathbf{x}(\tau),\mathbf{u}(\tau)\right) \\ &\left.- \mathbf{f}\left(\mathbf{x}(\tau),\mathbf{u}(\tau)\right)\right\} d\tau.\end{aligned} \tag{25}$$

Note that from (9), the eigenvalues of $-\mathbf{\Omega} + \mathbf{f}_x(\mathbf{x}_d,\mathbf{u})$ are all in the left-half s-plane since $\mathbf{\Omega}$ has been determined with that assumption. Also, at this point, it is not unreasonable to assume that both $\Delta\mathbf{f}_n(\mathbf{x}(\tau),\mathbf{x}_d,\mathbf{u}(\tau))$ and the modeling differences $|\mathbf{f}'(\mathbf{x},\mathbf{u}) - \mathbf{f}(\mathbf{x},\mathbf{u})|$ are bounded. Therefore, if the model in (3) is stable enough, the integral part in (25) will vanish along with $e^{-\mathbf{\Omega}t}$. Accordingly, as the time t gets large enough, the state variables $\mathbf{x}$ will approach the desired values $\mathbf{x}_d$ as can be found from (25) even if there are some modeling uncertainties or disturbances.

3. Nonlinear Dynamic Inversion Flight Control System Design for a UAV: An Example

3.1. Flight Dynamics Equations of Motion. To illustrate the theory, a design of flight control system with the method developed is presented. Before the flight control design proceeds, a set of flight dynamics equations of motion must be formulated. Note that all aerodynamic forces and moments result from the relative motions between aircraft and the air. The aircraft is assumed to have a ground velocity:

$$\mathbf{V} = u\mathbf{i} + v\mathbf{j} + w\mathbf{k} = \dot{X}_E\mathbf{I} + \dot{Y}_E\mathbf{J} - \dot{h}\mathbf{K}, \tag{26}$$

where $(\mathbf{i},\mathbf{j},\mathbf{k})$ are unit vectors in the aircraft body moving frame and $(\mathbf{I},\mathbf{J},\mathbf{K})$ the unit vectors in a fixed flat earth frame. The air is assumed to have a velocity:

$$\mathbf{V}_w = u_g\mathbf{i} + v_g\mathbf{j} + w_g\mathbf{k} = V_{w_x}\mathbf{I} + V_{w_y}\mathbf{J} - V_h\mathbf{K} \tag{27}$$

which is also known as the wind velocity. Then, the velocity of the aircraft relative to the air can be represented by

$$\mathbf{V}_a = \mathbf{V} - \mathbf{V}_w = \left(u-u_g\right)\mathbf{i} + \left(v-v_g\right)\mathbf{j} + \left(w-w_g\right)\mathbf{k} \tag{28}$$

from which, the aircraft total speed relative to the air, the angle of attack, and the sideslip angle can be determined, respectively, by the following equations:

$$\begin{aligned}V_a &= \sqrt{\left(u-u_g\right)^2 + \left(v-v_g\right)^2 + \left(w-w_g\right)^2} \\ \alpha &= \tan^{-1}\frac{w-w_g}{u-u_g} \\ \beta &= \tan^{-1}\frac{v-v_g}{\sqrt{\left(u-u_g\right)^2 + \left(w-w_g\right)^2}}.\end{aligned} \tag{29}$$

With the assumptions of fixed flat earth and winds being present, the motions of aircraft with six degrees of freedom can be represented by a set of nonlinear first-order differential equations as follows:

$$\begin{aligned}\dot{X}_E &\triangleq f_X \\ &= V_a\left(C_{11}\cos\alpha\cos\beta + C_{21}\sin\beta + C_{31}\sin\alpha\cos\beta\right) \\ &\quad + V_{w_x}\end{aligned} \tag{30}$$

$$\begin{aligned}\dot{Y}_E &\triangleq f_Y \\ &= V_a\left(C_{12}\cos\alpha\cos\beta + C_{22}\sin\beta + C_{32}\sin\alpha\cos\beta\right) \\ &\quad + V_{w_y}\end{aligned} \tag{31}$$

$$\begin{aligned}\dot{h} &\triangleq f_h \\ &= -V_a\left(C_{13}\cos\alpha\cos\beta + C_{23}\sin\beta + C_{33}\sin\alpha\cos\beta\right) \\ &\quad + V_h\end{aligned} \tag{32}$$

$$\dot{\psi} \triangleq f_\psi = q\sec\theta\sin\phi + r\sec\theta\cos\phi \tag{33}$$

$$\dot{\theta} \triangleq f_\theta = q\cos\phi - r\sin\phi \tag{34}$$

$$\dot{\phi} \triangleq f_\phi = p + q\tan\theta\sin\phi + r\tan\theta\cos\phi \tag{35}$$

$$\begin{aligned}\dot{V}_a &\triangleq f_V \\ &= \frac{F_x}{m}\cos\alpha\cos\beta + \frac{F_y}{m}\sin\beta + \frac{F_z}{m}\sin\alpha\cos\beta + f_{V_w}\end{aligned} \tag{36}$$

$$\begin{aligned}\dot{\alpha} &\triangleq f_\alpha \\ &= -p\cos\alpha\tan\beta + q - r\sin\alpha\tan\beta \\ &\quad - \frac{F_x}{mV_a}\sin\alpha\sec\beta + \frac{F_z}{mV_a}\cos\alpha\sec\beta + f_{\alpha_w}\end{aligned} \tag{37}$$

$$\begin{aligned}\dot{\beta} &\triangleq f_\beta \\ &= p\sin\alpha - r\cos\alpha - \frac{F_x}{mV_a}\cos\alpha\sin\beta + \frac{F_y}{mV_a}\cos\beta \\ &\quad - \frac{F_z}{mV_a}\sin\alpha\sin\beta + f_{\beta_w}\end{aligned} \tag{38}$$

$$\begin{aligned}\dot{p} &\triangleq f_p \\ &= \frac{I_{xz}\left(-I_{xx}+I_{yy}-I_{zz}\right)}{I_{xx}I_{zz}-I_{xz}I_{xz}}pq \\ &\quad + \frac{I_{yy}I_{zz}-I_{zz}I_{zz}-I_{xz}I_{xz}}{I_{xx}I_{zz}-I_{xz}I_{xz}}qr + \frac{I_{zz}}{I_{xx}I_{zz}-I_{xz}I_{xz}}M_x \\ &\quad - \frac{I_{xz}}{I_{xx}I_{zz}-I_{xz}I_{xz}}M_z\end{aligned} \tag{39}$$

$$\dot{q} \triangleq f_q = \frac{I_{xz}}{I_{yy}}\left(p^2 - r^2\right) + \frac{I_{zz}-I_{xx}}{I_{yy}}pr + \frac{1}{I_{yy}}M_y \tag{40}$$

$$\begin{aligned}\dot{r} &\triangleq f_r \\ &= \frac{I_{xx}I_{xx} - I_{xx}I_{yy} + I_{xz}I_{xz}}{I_{xx}I_{zz} - I_{xz}I_{xz}} pq \\ &\quad + \frac{I_{xz}\left(I_{xx} - I_{yy} + I_{zz}\right)}{I_{xx}I_{zz} - I_{xz}I_{xz}} qr - \frac{I_{xz}}{I_{xx}I_{zz} - I_{xz}I_{xz}} M_x \\ &\quad + \frac{I_{xx}}{I_{xx}I_{zz} - I_{xz}I_{xz}} M_z,\end{aligned} \tag{41}$$

where (X_E, Y_E) represents the position in a fixed flat earth frame, h the altitude, C_{ij} the elements of direction cosine matrix for transferring a fixed flat earth frame to the aircraft body moving frame, α the angle of attack, β the sideslip angle, ψ the heading angle, θ the pitch angle, ϕ the bank angle, p the roll rate, q the pitch rate, and r the yaw rate. Also, F_x, F_y, and F_z represent, respectively, three components of the total force in an aircraft body moving frame. The three force components are composed of the thrust T, the lift L, the drag D, the side force Y, and the weight mg (m is the aircraft mass and g the gravity acceleration) by the following equations:

$$\begin{aligned}F_x &= T\cos\theta_T + L\sin\alpha - D\cos\alpha - mg\sin\theta \\ F_y &= Y + mg\sin\phi\cos\theta \\ F_z &= -T\sin\theta_T - L\cos\alpha - D\sin\alpha + mg\cos\phi\cos\theta,\end{aligned} \tag{42}$$

where θ_T is the angle between the thrust and the longitudinal axis. Moreover, M_x, M_y, and M_z represent the roll moment, the pitch moment, and the yaw moment, respectively, about the center of gravity, and I_{xx}, I_{yy}, I_{zz}, and I_{xz} are the components of the moment-of-inertia tensor. Furthermore, the wind disturbances on $\dot{V}_a$, $\dot{\alpha}$, and $\dot{\beta}$ are, respectively, represented by f_{V_w}, f_{α_w}, and f_{β_w} which are expressed as follows:

$$\begin{aligned}f_{V_w} &= \frac{F_{x_w}}{m}\cos\alpha\cos\beta + \frac{F_{y_w}}{m}\sin\beta + \frac{F_{z_w}}{m}\sin\alpha\cos\beta \\ f_{\alpha_w} &= -\frac{F_{x_w}}{mV_a}\sin\alpha\sec\beta + \frac{F_{z_w}}{mV_a}\cos\alpha\sec\beta \\ f_{\beta_w} &= -\frac{F_{x_w}}{mV_a}\cos\alpha\sin\beta + \frac{F_{y_w}}{mV_a}\cos\beta \\ &\quad - \frac{F_{z_w}}{mV_a}\sin\alpha\sin\beta,\end{aligned} \tag{43}$$

where

$$\begin{aligned}F_{x_w} &= m\left(rv_g - qw_g - \dot{u}_g\right) \\ F_{y_w} &= m\left(-ru_g + pw_g - \dot{v}_g\right) \\ F_{z_w} &= m\left(qu_g - pv_g - \dot{w}_g\right).\end{aligned} \tag{44}$$

In (44), F_{x_w}, F_{y_w}, and F_{z_w} are the three components of the force exerted by winds. At this point, it is worthy to mention that although there is no explicit term relating wind disturbances to $\dot{p}$, $\dot{q}$, and $\dot{r}$ in (39)–(41), winds do have effects on V_a, α, and β through which M_x, M_y, and M_z are affected. Also, winds not only have explicit effects on $\dot{V}_a$, $\dot{\alpha}$, and $\dot{\beta}$ in (36)–(38), but also have implicit effects on them through L, Y, and D which obviously depend on V_a, α, and β.

Table 1: The geometric data, weight, and moment of inertia of the UAV.

	Parameter	Value	Units
Reference wing area	S	75	ft^2
Wing span	b	15	ft
Mean aerodynamic chord	$\bar{c}$	5.66	ft
Weight	W	2562.5	lb
x-axis inertia	I_{xx}	296.75	slug/ft^2
y-axis inertia	I_{yy}	1744.1875	slug/ft^2
z-axis inertia	I_{zz}	1971.875	slug/ft^2
x-z product of inertia	I_{xz}	30.6875	slug/ft^2

To validate the method developed in this paper, a UAV as shown in Figure 1 is introduced. The parameters used for analysis are listed in Table 1.

The aerodynamic forces and moments of the UAV are computed by the following equations:

$$\begin{aligned}L &= \bar{q}S\left[C_L\left(M,\alpha,\delta_e\right) + C_{L_q}\left(M,\alpha\right)\frac{\bar{c}q}{2V_a}\right] \\ D &= \bar{q}SC_D\left(M,\alpha,\delta_e\right) \\ Y &= \bar{q}S\left[C_{Y_\beta}\left(M,\alpha\right)\beta + C_{Y_p}\left(M,\alpha\right)\frac{bp}{2V_a}\right. \\ &\quad + C_{Y_r}\left(M,\alpha\right)\frac{br}{2V_a} + \Delta C_Y\left(M,\alpha,\delta_a\right) \\ &\quad \left. + \Delta C_Y\left(M,\alpha,\delta_r\right)\right] \\ M_x &= \bar{q}Sb\left[C_{\ell_\beta}\left(M,\alpha\right)\beta + C_{\ell_p}\left(M,\alpha\right)\frac{bp}{2V_a}\right. \\ &\quad + C_{\ell_r}\left(M,\alpha\right)\frac{br}{2V_a} + \Delta C_\ell\left(M,\alpha,\delta_a\right) \\ &\quad \left. + \Delta C_\ell\left(M,\alpha,\delta_r\right)\right] \\ M_y &= \bar{q}S\bar{c}\left[C_m\left(M,\alpha,\delta_e\right) + C_{m_q}\left(M,\alpha\right)\frac{\bar{c}q}{2V_a}\right] \\ M_z &= \bar{q}Sb\left[C_{n_\beta}\left(M,\alpha\right)\beta + C_{n_p}\left(M,\alpha\right)\frac{bp}{2V_a}\right. \\ &\quad + C_{n_r}\left(M,\alpha\right)\frac{br}{2V_a} + \Delta C_n\left(M,\alpha,\delta_a\right) \\ &\quad \left. + \Delta C_n\left(M,\alpha,\delta_r\right)\right],\end{aligned} \tag{45}$$

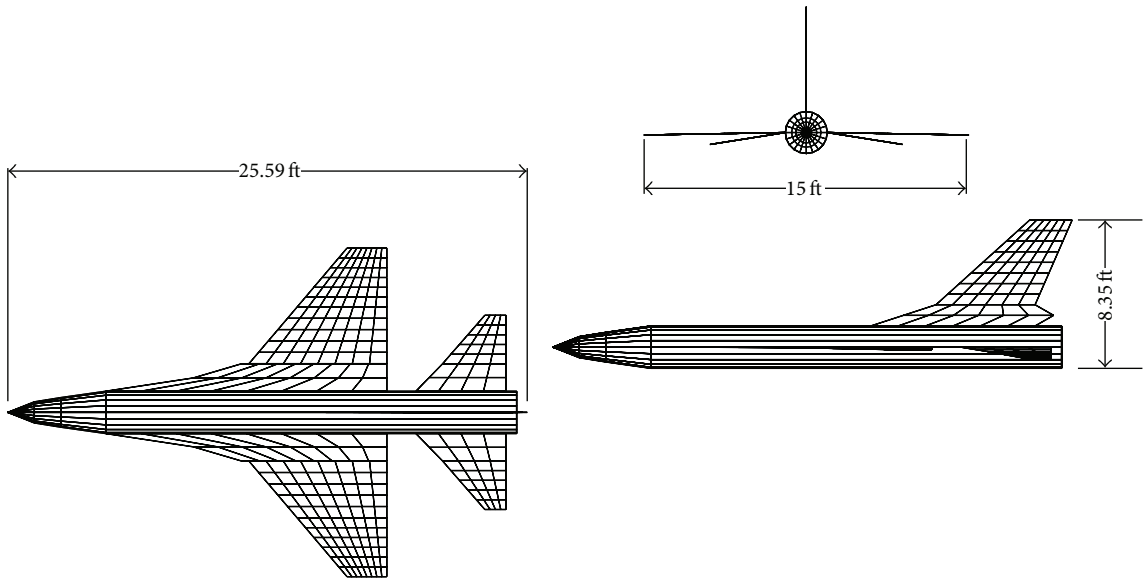


Figure 1: The configuration of the UAV.

where $\overline{q}$ is the dynamic pressure and M the Mach number. The aerodynamic coefficients and stability derivatives are determined by a computer code dubbed "VORSTAB" [17]. The computation results are in the form of discrete data which are then interpolated as continuous functions in the computer program. These functions are represented by $C_L(M,\alpha,\delta_e)$, $C_{L_q}(M,\alpha)$, and so forth, in (45). The nonlinear database includes α ranging from -20° to 40°, δ_e from -24° to 24°, δ_a from -25° to 25°, and δ_r from -30° to 30°. Also, the flight conditions include altitude h ranging from sea level to 40,000 ft and Mach number M from 0.1 to 0.9 and from 1.1 to 1.9. The thrust model of the UAV is assumed as

$$\dot{T} = 0.1\left(T_c - T\right), \tag{46}$$

where T_c is the thrust command. The thrust is assumed to have a limitation, say $0 \le T \le 1.1W$.

3.2. Flight Control Design. In flight control design, only nominal flight dynamics models are considered since modeling uncertainties or wind disturbances, and so forth, cannot be determined in advance. Hence, all terms related to the wind disturbances in the flight dynamics equations of motion are neglected in this stage. For this flight dynamics model, the parameters involved in the control analysis are further elaborated as follows.

(1) In the flight dynamics equations of motion, only h, θ, ϕ, V, α, β, p, q, and r are coupled. Therefore $\mathbf{x} \triangleq \{h,\theta,\phi,V,\alpha,\beta,p,q,r\}^T$, $\mathbf{f} = \{f_h, f_\theta, f_\phi, f_V, f_\alpha, f_\beta, f_p, f_q, f_r\}^T$, and $n = 9$.

(2) The control input vector $\mathbf{u} \triangleq \{\delta_e,\delta_a,\delta_r,T_c\}^T$ and $m = 4$. According to the theory derived in the above, only 4 state variables can be controlled or designated. In this paper, the state vector to be controlled $\mathbf{x}_c \triangleq \{h_c, V_c, \phi_c, \beta_c\}^T$ is chosen and the residual vector is $\mathbf{x}_r \triangleq \{\theta_r, \alpha_r, p_r, q_r, r_r\}^T$.

(3) Determine the state variables and control variables for some trim conditions, on which the control design is based.

(4) The model matrix in (3) is constructed as

$$\mathbf{\Omega} = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & \sigma_4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_4 & 0 & \sigma_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_5 & 0 & 0 & \sigma_8 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_6 & 0 & 0 & \sigma_9 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_7 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_8 & 0 & 0 & \sigma_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_9 & 0 & 0 & \sigma_6 \end{bmatrix}, \tag{47}$$

where the arrangement for the pair (σ_2, σ_4) is enlightened from the phugoid mode in which θ and V are coupled. Similar arrangements are for the pairs (σ_5, σ_8) and (σ_6, σ_9). In order to make the desired dynamics stable, the elements on the diagonal, σ_1, σ_2, σ_3, σ_5, σ_6, and σ_7, must be positive. Also, those off-diagonal elements must satisfy the conditions, $\sigma_4^2 < \sigma_2^2$, $\sigma_8^2 < \sigma_5^2$, and $\sigma_9^2 < \sigma_6^2$. For convenience of later usage, these necessary conditions are dubbed "the stable model requirements."

In the design, the initial cruise flight conditions are $h = 600$ ft and $M = 0.25$. The trim conditions are $\alpha_{\text{trim}} = \theta_{\text{trim}} = 6.045^\circ$, $T_{\text{trim}} = 341.6$ lb, and $\delta_{e_{\text{trim}}} = -0.872^\circ$. Based on these data, the open-loop eigenvalues of $\mathbf{f}_{\mathbf{x}}(\mathbf{x}_{\text{trim}}, \mathbf{u}_{\text{trim}})$ are determined and listed in Table 2.

Table 2: The open-loop eigenvalues of $\mathbf{f}_{\mathbf{x}}(\mathbf{x}_{\text{trim}}, \mathbf{u}_{\text{trim}})$.

$\lambda_{1,2}$	$-1.124786 \pm 1.566875i$	λ_6	-3.129774
$\lambda_{3,4}$	$-0.008857 \pm 0.158058i$	$\lambda_{7,8}$	$-0.297956 \pm 0.410071i$
λ_5	-0.000133	λ_9	-0.023006

Longitudinal dynamics analysis reveals that $\lambda_{1,2}$ and $\lambda_{3,4}$ are closely associated with the short period mode and the phugoid mode, respectively, and lateral dynamics analysis shows that λ_6, $\lambda_{7,8}$, and λ_9 are closely associated with the pure roll mode, the Dutch roll mode, and the spiral mode, respectively. Finally, the rest one, λ_5, can be inferred to be associated with the altitude dynamics.

Intuitively, σ_i $(i = 1, 2, 3, \ldots, 9) \in \boldsymbol{\sigma}$ may just be arbitrarily assigned as long as they satisfy the stable model requirements. However, the resulted error dynamics in (9) may not just be stable. As will be shown in simulations, the desired eigenvalues must not only be able to make the system follow the desired dynamics, but also be able to generate a good set of $\boldsymbol{\sigma}$ which makes $e^{-\boldsymbol{\Omega} t}$ decrease so quickly that the system is also robust if modeling uncertainties exist or wind disturbances are encountered. The choice of these eigenvalues is not only just based on the control analysis alone but must also be simultaneously based on simulations.

Note that, in this case, longitudinal dynamics analysis also reveals a right-half zero $z = 10.85$ in the h-δ_e transfer function. It can therefore be identified as a nonminimum phase problem. For this kind of problem, the desired eigenvalues must be very carefully assigned lest the resulting $\boldsymbol{\sigma}$ does not satisfy the stable model requirements. To determine $\boldsymbol{\sigma}$, a two-way approach ($\boldsymbol{\sigma} \rightleftharpoons \boldsymbol{\lambda}$) is proposed as follows.

(1) If the desired eigenvalues are equal to the open-loop eigenvalues as listed in Table 2, obviously, $\boldsymbol{\sigma} = \mathbf{0}$. Choose small $\boldsymbol{\sigma}(\neq \mathbf{0})$ which satisfies the stable model requirements.

(2) Determine the eigenvalues λ_i $(i = 1, 2, 3, \ldots, 9) \in \boldsymbol{\lambda}$ of the error dynamics in (9).

(3) If the error dynamics is not stable, modify $\boldsymbol{\lambda}$ so that it makes the error dynamics stable. Use the modified values as the desired eigenvalues.

(4) Determine $\boldsymbol{\sigma}$ by following the 7 steps of iteration procedure described in Section 2.2.

(5) If $\boldsymbol{\sigma}$ does not satisfy the stable model requirements, modify $\boldsymbol{\sigma}$ to make it satisfy. Go to step 2.

(6) In each step of simulations, determine $\mathbf{x}_r$ and $\mathbf{u}$ from (4) by using the iteration method described in (5). In this step, only the nominal dynamics model is used.

(7) Use $\mathbf{u}$ in (1) for simulations. The equations of motion may include modeling uncertainties or wind disturbances.

(8) If the simulation results are not satisfactory, modify $\boldsymbol{\sigma}$ and then go to step 2, or modify $\boldsymbol{\lambda}$ and then go to step 4.

Table 3: The designated eigenvalues.

$\lambda_{1,2}$	$-4.22180 \pm 1.57684i$	λ_7	-1.75509
$\lambda_{3,4}$	$-0.112728 \pm 0.0902199i$	λ_8	-0.920091
λ_5	-0.0204641	λ_9	-0.146833
λ_6	-4.73168		

Table 4: The resulting parameters for desired dynamics.

σ_1	0.0201234	σ_6	1.04305
σ_2	0.101328	σ_7	1.60654
σ_3	0.112353	σ_8	0.00193758
σ_4	-0.0000325585	σ_9	0.295443
σ_5	3.09966		

By using these 8 steps of procedure, the designated eigenvalues and the resulting parameters for desired dynamics are obtained and listed in Tables 3 and 4, respectively.

These parameters indeed satisfy the stable model requirements. Note that initially σ_5 is small and $\lambda_{1,2}$ and $\lambda_{3,4}$ are not so deviated from their counterparts of open-loop eigenvalues. However, the simulation results show that the performance in tracking h_c is not good enough. Enlarging σ_5 does improve the performance but makes $\lambda_{1,2}$ and $\lambda_{3,4}$ deviate their counterparts of open-loop eigenvalues a lot. Also note that the large deviations of λ_6, $\lambda_{7,8}$, and λ_9 from their counterparts of open-loop eigenvalues are due to the iteration procedure in determining a set of σ_6, σ_7, σ_8, and σ_9 for satisfying the stable model requirements.

3.3. Flight Simulations. In simulations, the initial conditions are $(X_E, Y_E) = (0, 0)$, $h = 600$ ft, $\psi = 0^\circ$, $\theta = \theta_{\text{trim}} = 6.04^\circ$, $\phi = 0^\circ$, $M = 0.25$, $\alpha = \alpha_{\text{trim}} = 6.04^\circ$, $\beta = 0^\circ$, $p = 0$, $q = 0$, and $r = 0$. The states to be tracked are $h_c = 2,000$ ft, $\psi_c = 90^\circ$, $M_c = 0.40$, and $\beta_c = 0^\circ$. Although the heading angle is to be tracked, for pilots, it would make more sense to regulate the bank angle rather than the heading angle by assuming

$$\phi_c = k_\phi (\psi_c - \psi), \tag{48}$$

where k_ϕ is a constant parameter. In this study, $k_\phi = 0.12$.

At each instant, h_c, V_c, ϕ_c, and β_c are given, the controls δ_e, δ_a, δ_r, and T_c along with the residual state variables α_r, θ_r, p_r, q_r, and r_r can be determined from (4) by using the iteration method described in (5). Note that, in the computation of the control, only nominal flight dynamics equations of motion are used since modeling uncertainties or wind disturbances are not known. Also note that, in using the Newton-Raphson method, the convergence can be guaranteed if the initial guess is sufficiently close to the solution. Since the solution changes only very little in each step of integration, numerical practices show that it takes only 4 iterations to converge to within 0.0001% of the correct value if the solution in the previous step of integration is taken as an initial guess. In the first step of integration, it may be necessary to take a few more iterations, say, 10 iterations, since the solution is different

from that in the trim conditions, which is usually taken as an initial guess.

For illustration, simulation block diagrams including $\mathbf{x}$, $\mathbf{x}_d$, $\mathbf{x}_c$, $\mathbf{x}_r$, and $\mathbf{u}$ are shown in Figure 2.

After the controls δ_e, δ_a, δ_r, and T_c are determined, (30)–(41) are used for simulations. In order to ascertain if the control system is robust or not, a fictitious wind model is put enroute as follows:

$$V_{w_x} = \begin{cases} -\dfrac{30 \times 6080}{3600} \sin\left(\dfrac{2\pi\left(X_E - \left(X_{\text{wo}} - X_{\text{scale}}\right)\right)}{2X_{\text{scale}}}\right) & \text{if } \left|X_E - X_{\text{wo}}\right| < X_{\text{scale}}, \\ 0 & \text{else.} \end{cases} \tag{49}$$

$$V_{w_y} = \begin{cases} \dfrac{30 \times 6080}{3600} \sin\left(\dfrac{2\pi\left(Y_E - \left(Y_{\text{wo}} - Y_{\text{scale}}\right)\right)}{2Y_{\text{scale}}}\right) & \text{if } \left|Y_E - Y_{\text{wo}}\right| < Y_{\text{scale}}, \\ 0 & \text{else.} \end{cases} \tag{50}$$

$$V_h = \begin{cases} \dfrac{10 \times 6080}{3600} \sin\left(\dfrac{2\pi\left(h - \left(H_{\text{wo}} - H_{\text{scale}}\right)\right)}{2H_{\text{scale}}}\right) & \text{if } \left|h - H_{\text{wo}}\right| < H_{\text{scale}}, \\ 0 & \text{else.} \end{cases} \tag{51}$$

$$\dot{V}_{w_x} = \begin{cases} -\dfrac{30 \times 6080}{3600} \times \dfrac{2\pi\dot{X}_E}{2X_{\text{scale}}} \cos\left(\dfrac{2\pi\left(X_E - \left(X_{\text{wo}} - X_{\text{scale}}\right)\right)}{2X_{\text{scale}}}\right) & \text{if } \left|X_E - X_{\text{wo}}\right| < X_{\text{scale}}, \\ 0 & \text{else.} \end{cases} \tag{52}$$

$$\dot{V}_{w_y} = \begin{cases} \dfrac{30 \times 6080}{3600} \times \dfrac{2\pi\dot{Y}_E}{2Y_{\text{scale}}} \cos\left(\dfrac{2\pi\left(Y_E - \left(Y_{\text{wo}} - Y_{\text{scale}}\right)\right)}{2Y_{\text{scale}}}\right) & \text{if } \left|Y_E - Y_{\text{wo}}\right| < Y_{\text{scale}}, \\ 0 & \text{else.} \end{cases} \tag{53}$$

$$\dot{V}_h = \begin{cases} \dfrac{10 \times 6080}{3600} \times \dfrac{2\pi\dot{h}}{2H_{\text{scale}}} \cos\left(\dfrac{2\pi\left(h - \left(H_{\text{wo}} - H_{\text{scale}}\right)\right)}{2H_{\text{scale}}}\right) & \text{if } \left|h - H_{\text{wo}}\right| < H_{\text{scale}}, \\ 0 & \text{else,} \end{cases} \tag{54}$$

where $X_{\text{wo}} = 29{,}727$ ft, $Y_{\text{wo}} = 10{,}180$ ft, and $H_{\text{wo}} = 1{,}000$ ft represent the center position of wind zone, and $X_{\text{scale}} = 5{,}000$ ft, $Y_{\text{scale}} = 5{,}000$ ft, and $H_{\text{scale}} = 1{,}000$ ft represent the maximum range of wind zone from its center. From (49)–(51), it is also known that the three maximum wind velocity components in the earth fixed frame are 30 knots, 30 knots, and 10 knots, respectively.

In this study, two cases of simulations are conducted; one is without wind disturbances and another with wind disturbances. The results are presented in Figures 3–6.

As shown in Figure 3(a), there is no much difference in ground trajectories whether wind disturbances are present or not. While flying, the UAV first suffers vertical wind disturbances in between $t = 15.10$ sec and 32.04 sec and then horizontal wind disturbances in between $t = 50.60$ sec and 90.00 sec as shown in Figures 3(b) and 3(c).

The state variables to be tracked are all plotted in Figures 4(a)–4(e). From Figure 4(a), it is found that h decreases initially when $h_c = 2{,}000$ ft is commanded. A careful study reveals that the elevator angle computed is $\delta_e = -0.733^\circ$, which is not enough to hold the UAV level as $\delta_{e_{\text{trim}}} = -0.872^\circ$ is required. The lowest altitude reached is 578 ft with $\sigma_5 = 3.09966$ in Table 4 being used. Increasing σ_5 will increase the lowest altitude reached but at the expense of increasing overshoot. Fortunately, after descending to 578 ft, the UAV begins to climb. During its climb, the UAV encounters an ascending wind and then a descending wind. But the influences on the altitude are almost negligible. In contrary, the horizontal wind seems to have more influences as it makes the altitude fluctuate. Whether wind disturbances are present or not, the UAV can always approach the commanded altitude asymptotically.

As shown in Figure 4(b), the Mach number does not seem to be so affected by the vertical wind as does by the horizontal wind. In this case, the horizontal wind causes the Mach number to fluctuate up to ±10%. The sharp angles shown in the figure are due to the discontinuity of wind acceleration at the edges of wind zone as shown in Figure 3(c).

When the UAV is commanded to make a 90° turn, its heading angle gradually increases and asymptotically approaches the command as shown in Figure 4(c) whether winds are present or not. Figure 4(d) shows that the bank

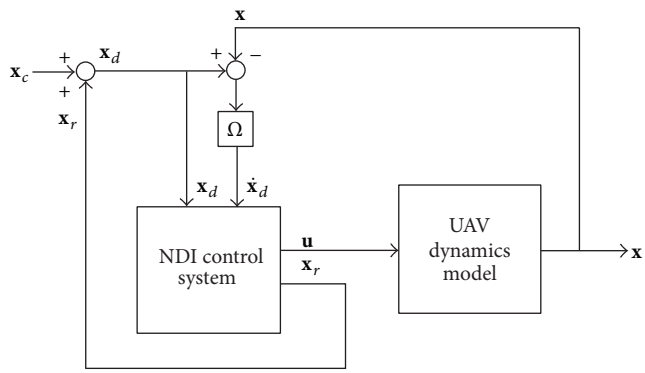


Figure 2: Simulation block diagrams for the NDI flight control system.

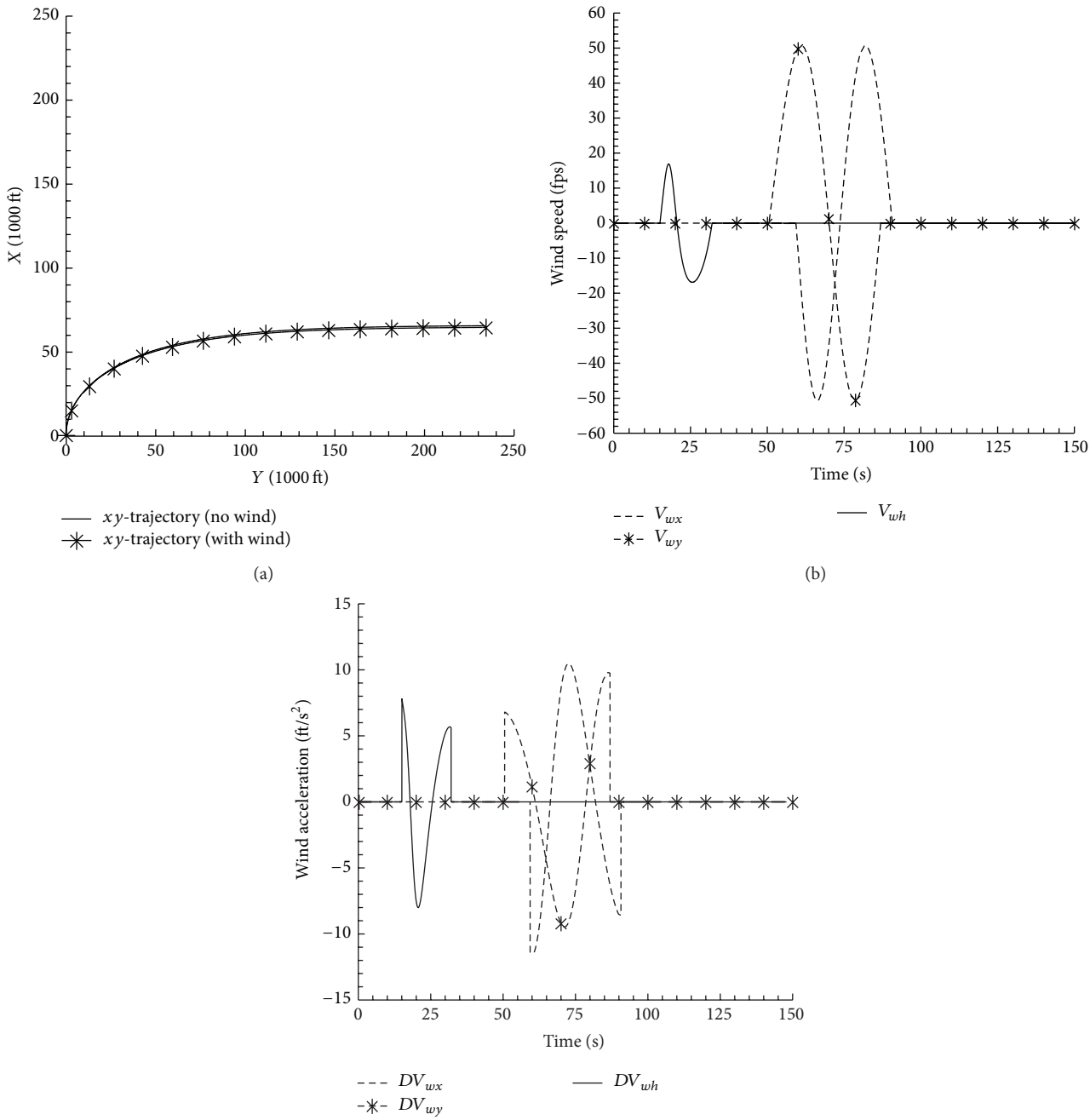


Figure 3: (a) Trajectory in X-Y plane, (b) wind speed, and (c) wind acceleration.

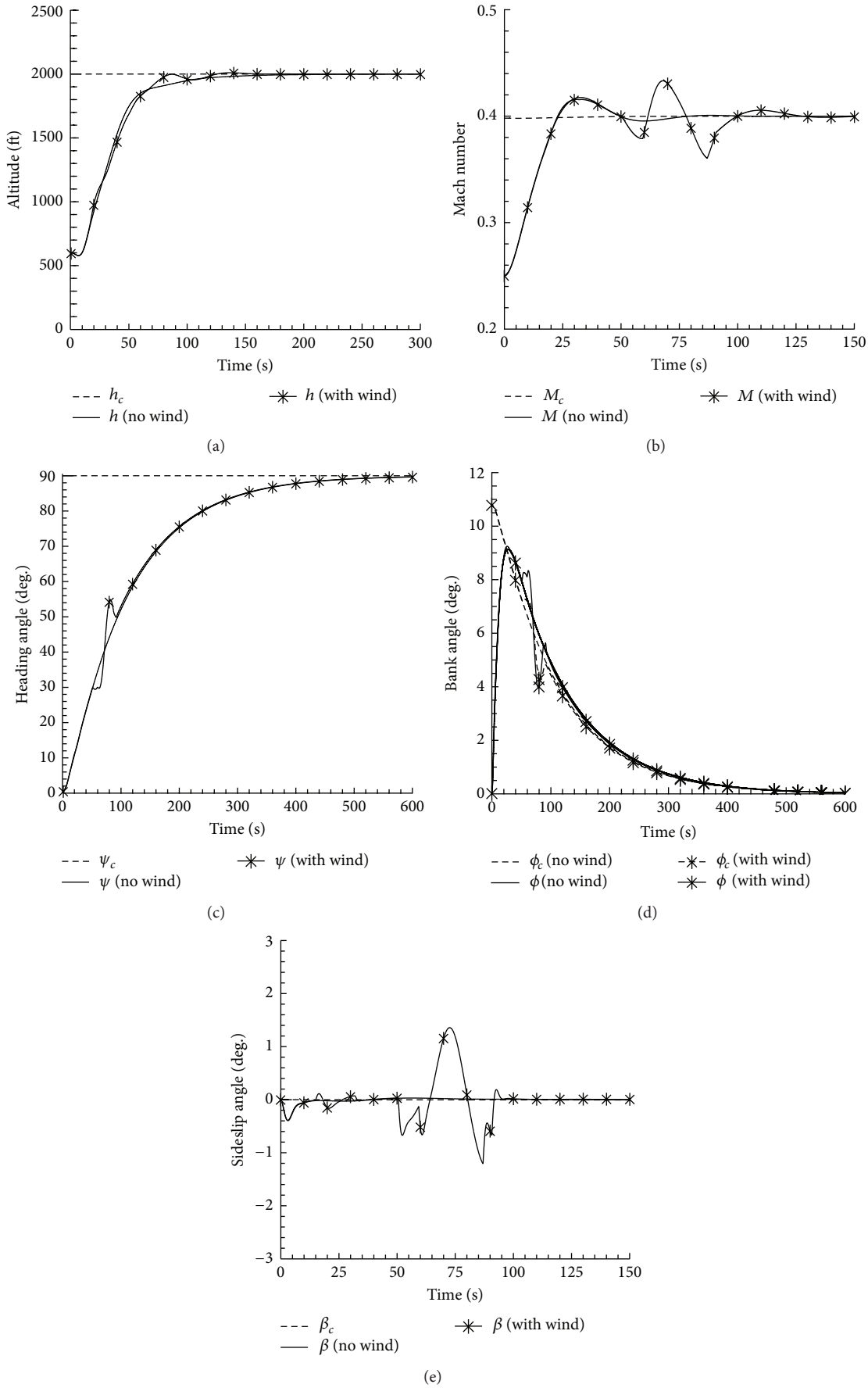


Figure 4: The responses of (a) altitude, (b) Mach number, (c) heading angle, (d) bank angle, and (e) sideslip angle.

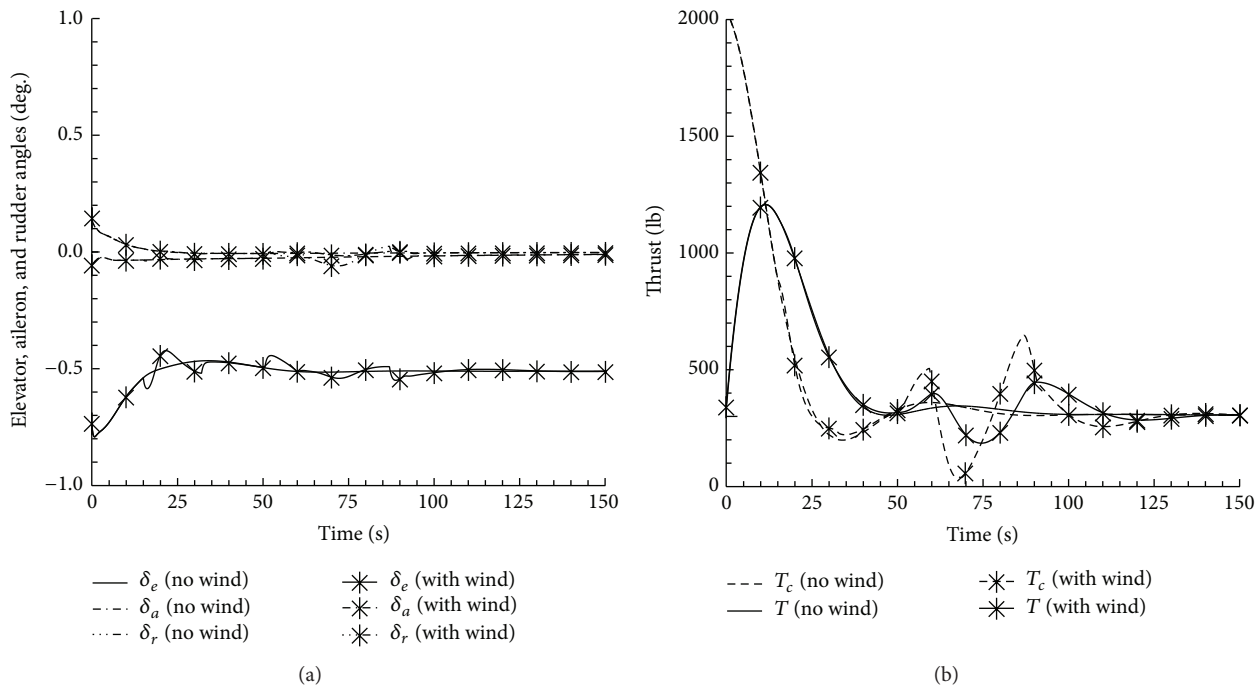


Figure 5: The responses of (a) control surface deflections and (b) thrust.

angle is closely associated with the heading angle as the bank angle is computed based on (48). Although the wind disturbances do not explicitly affect $\dot{\phi}$ in (35), they do implicitly affect it through α and β which in turn affect the aerodynamic moments. Fluctuations in the bank angle and the heading angle are remarkable when the UAV passes through the wind zone but are still tolerable.

Sideslip angle seems to be affected by winds more heavily as shown in Figure 4(e). Fortunately, the control system is robust enough as being able to keep it small.

As shown in Figure 5(a), all control surface deflection angles remain small. The elevator needs only to deflect slightly to rotate the UAV to climb. As mentioned early, $\delta_e = -0.733^\circ$ at the instant when h_c and M_c are given. When the UAV reaches the commanded state conditions, the elevator trim angle approaches -0.510°. Small deflection in aileron angle is enough to make the UAV bank turn and the rudder just keeps very small as does the sideslip angle. All control surface deflection angles are not remarkably affected as the UAV passes through the wind zone.

As shown in Figure 5(b), the thrust command increases very sharply as a demand to increase both the altitude and the Mach number simultaneously is given. Also, when the UAV passes the wind zone, the thrust command fluctuates very violently. Obviously, the thrust is very closely interacted with the Mach number. A low pass filter with time constant 10 sec alleviates the sharp and violent responses a lot for the actual thrust at the expense of delaying its reaction time.

In Figures 6(a)–6(e), responses of all residual state variables along with the computed commands for them are presented. In fact, these commands are generated in the inner system, not input from outer designations. It is interesting to observe how the actual state variables track the commands.

In Figure 6(a), it is observed that, in comparison to $\alpha_{\text{trim}} = 6.04^\circ$, the computed command $\alpha_r = 4.85^\circ$ at the instant when h_c and M_c are given. Then the angle of attack α follows α_r closely without apparent short-period mode oscillations when the designated short-period mode eigenvalues are as high as $\lambda_{1,2} = -4.22180 \pm 1.57684i$. Winds do have remark effect on both α_r and α. Finally, both approach a new value, $\alpha_{\text{trim}} = 2.88^\circ$, for flight conditions at higher altitude and faster speed.

In Figures 6(b) and 6(c), it is observed that although both θ_r and q_r increase in response to the requirement for increasing the cruise altitude, θ and q decrease remarkably. As revealed in explaining h response, the elevator angle computed is not enough to hold $\dot{q} = 0$ and therefore $\dot{q} < 0$ which in turn makes $\dot{\theta} < 0$. The minimum pitch rate is $q = -0.717$ deg./s at $t = 0.78$ sec. In this case, the vertical wind does have more remarkable effects on θ and q than the horizontal wind. Finally, both approach $\theta_{\text{trim}} = \alpha_{\text{trim}} = 2.88^\circ$ and $q_{\text{trim}} = 0$, respectively, for the new flight conditions.

It is very interesting to find that, in Figures 6(d) and 6(e), the shapes of p and r responses resemble, respectively, those of ϕ and β responses as shown in Figures 4(d) and 4(e). The effects of horizontal wind on both responses are more remarkable than those of vertical wind in this case. The decay of r seems to be very slow.

4. Conclusions

In this paper, a unified nonlinear dynamic inversion control system is successfully developed. With this method, the

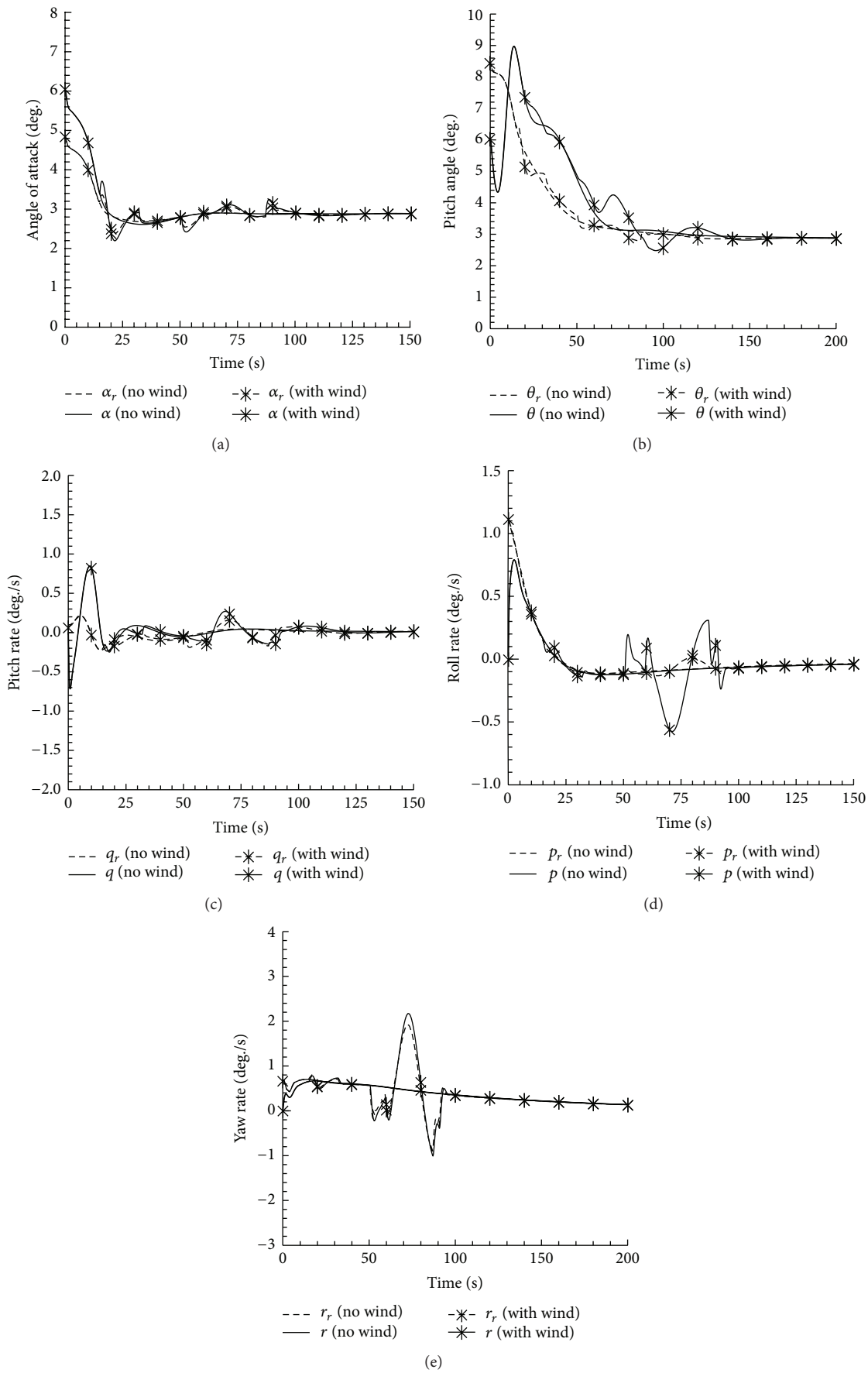


Figure 6: The responses of (a) angle of attack, (b) pitch angle, (c) pitch rate, (d) roll rate, and (e) yaw rate.

parameters for desired dynamics can be determined with a set of assigned eigenvalues for error dynamics. The control system has been proved to be robust to modeling uncertainties or wind disturbances. A nonaffine nonlinear flight dynamics system with right-half zero has been used as an example for the control design. In the design process, it is not necessary to use the feedback linearization to transform the nonlinear equations into linear equations. Numerical simulations of the control system show that the desired state variables can be successfully tracked whether winds are present or not.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

The authors would like to express their gratitude to Chuan-Tau Edward Lan for his helps in computing the aerodynamic data of the UAV used as an example for the control design demonstrated in this paper.

References

[1] M. R. Mendenhall, S. C. Perkins, M. Tomac, A. Rizzi, and R. K. Nangia, "Comparing and benchmarking engineering methods for the prediction of X-31 aerodynamics," *Aerospace Science and Technology*, vol. 20, no. 1, pp. 12–20, 2012.

[2] M. L. Ireland, A. Vargas, and D. Anderson, "A comparison of closed-loop performance of multirotor configurations using non-linear dynamic inversion control," *Aerospace*, vol. 2, no. 2, pp. 325–352, 2015.

[3] J. O. Pedro, A. Panday, and L. Dala, "A nonlinear dynamic inversion-based neurocontroller for unmanned combat aerial vehicles during aerial refuelling," *International Journal of Applied Mathematics and Computer Science*, vol. 23, no. 1, pp. 75–90, 2013.

[4] N. Hovakimyan, E. Lavretsky, and A. Sasane, "Dynamic inversion for nonaffine-in-control systems via time-scale separation. Part I," *Journal of Dynamical and Control Systems*, vol. 13, no. 4, pp. 451–465, 2007.

[5] Z. Xie, Y. Xia, and M. Fu, "Robust trajectory-tracking method for UAV using nonlinear dynamic inversion," in *Proceedings of the IEEE 5th International Conference on Cybernetics and Intelligent Systems (CIS '11)*, pp. 93–98, Qingdao, China, September 2011.

[6] G. Gao, J. Wang, and X. Wang, "Adaptive fault-tolerant control for feedback linearizable systems with an aircraft application," *International Journal of Robust and Nonlinear Control*, vol. 25, no. 9, pp. 1301–1326, 2015.

[7] H. N. Foghahaayee, M. B. Menhaj, and H. A. Talebi, "Weakly and strongly non-minimum phase systems: properties and limitations," *International Journal of Control*, pp. 1–16, 2015.

[8] L. Fiorentini and A. Serrani, "Adaptive restricted trajectory tracking for a non-minimum phase hypersonic vehicle model," *Automatica*, vol. 48, no. 7, pp. 1248–1261, 2012.

[9] A. R. Babaei, M. Mortazavi, and M. H. Moradi, "Fuzzy sliding mode autopilot design for nonminimum phase and nonlinear UAV," *Journal of Intelligent and Fuzzy Systems*, vol. 24, no. 3, pp. 499–509, 2013.

[10] N. Dadkhah and B. Mettler, "Control system design and evaluation for robust autonomous rotorcraft guidance," *Control Engineering Practice*, vol. 21, no. 11, pp. 1488–1506, 2013.

[11] Z. Lang and A. Wu, "Study on dual-loop controller of helicopter based on the robust hinfinite loop shaping method," *Applied Mechanics and Materials*, vol. 130–134, pp. 1182–1185, 2012.

[12] S. Sieberling, Q. P. Chu, and J. A. Mulder, "Robust flight control using incremental nonlinear dynamic inversion and angular acceleration prediction," *Journal of Guidance, Control, and Dynamics*, vol. 33, no. 6, pp. 1732–1742, 2010.

[13] A. Rahideh, A. H. Bajodah, and M. H. Shaheed, "Real time adaptive nonlinear model inversion control of a twin rotor MIMO system using neural networks," *Engineering Applications of Artificial Intelligence*, vol. 25, no. 6, pp. 1289–1297, 2012.

[14] J. Georgie and J. Valasek, "Evaluation of longitudinal desired dynamics for dynamic-inversion controlled generic reentry vehicles," *Journal of Guidance, Control, and Dynamics*, vol. 26, no. 5, pp. 811–819, 2003.

[15] G. A. Smith and G. Meyer, "Aircraft automatic flight control system with model inversion," *Journal of Guidance, Control, and Dynamics*, vol. 10, no. 3, pp. 269–275, 1987.

[16] I. Yang, D. Lee, and D. S. Han, "Designing a robust nonlinear dynamic inversion controller for spacecraft formation flying," *Mathematical Problems in Engineering*, vol. 2014, Article ID 471352, 12 pages, 2014.

[17] C. T. E. Lan, "VORSTAB, a computer program for calculating lateral directional stability derivatives with vortex flow effect," NASA CR-172501, NASA, 1985.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 313154, 13 pages
http://dx.doi.org/10.1155/2015/313154

Research Article

Computing and Controlling Basins of Attraction in Multistability Scenarios

John Alexander Taborda[1] and Fabiola Angulo[2]

[1] *Facultad de Ingeniería, Programa de Ingeniería Electrónica, Magma Ingeniería, Universidad del Magdalena, Apartado Postal 2121630, Santa Marta, Colombia*
[2] *Departamento de Ingeniería Eléctrica, Electrónica y Computación, Percepción y Control Inteligente, Facultad de Ingeniería y Arquitectura, Universidad Nacional de Colombia (Sede Manizales), Campus La Nubia, Bloque Q, Manizales 170003, Colombia*

Correspondence should be addressed to Fabiola Angulo; fangulog@unal.edu.co

Received 17 October 2015; Accepted 16 November 2015

Academic Editor: Rongwei Guo

The aim of this paper is to describe and prove a new method to compute and control the basins of attraction in multistability scenarios and guarantee monostability condition. In particular, the basins of attraction are computed only using a submap, and the coexistence of periodic solutions is controlled through fixed-point inducting control technique, which has been successfully used until now to stabilize unstable periodic orbits. In this paper, however, fixed-point inducting control is used to modify the domains of attraction when there is coexistence of attractors. In order to apply the technique, the periodic orbit whose basin of attraction will be controlled must be computed. Therefore, the fixed-point inducting control is used to stabilize one of the periodic orbits and enhance its basin of attraction. Then, using information provided by the unstable periodic orbits and basins of attractions, the minimum control effort to stabilize the target periodic orbit in all desired ranges is computed. The applicability of the proposed tools is illustrated through two different coupled logistic maps.

1. Introduction

Complex bifurcation scenarios have been observed in nonlinear dynamic systems from virtually all areas of science, including a broad range of natural sciences, mechanical and electrical engineering, and economics and other areas of the social sciences [1–3]. Theoretical and applied researches have explained various bifurcation scenarios [4–6], and analytical, numerical, and experimental works have contributed to unraveling the complexity inherent to chaotic motion [3].

Coupled chaotic maps are a set of special discrete-time dynamical systems that can describe chemical, epidemiological, physiological, biological, or engineering systems [7]. Interesting nonlinear phenomena have been reported in them. For example, in [8, 9], coupled logistic maps were analyzed and new scenarios for transition to chaos were found via the creation and destruction of multilayered tori. In those papers, novel routes to chaos were described, and authors found that, depending on the coupling constant value, the system approaches different periodic attractors.

Control methods of coexisting attractors in multistability scenarios have been widely studied in the last decades. In [10], periodic signals were replaced by chaotic ones in order to eliminate multiple domains of attraction. In [11, 12], the influence of noise on preference and dominance of attractors in multistable systems was studied. In [13, 14], multistability was controlled using small periodic modulation of a system parameter. In [15], the basins of attraction (BA) were controlled using harmonic and stochastic modulation. In [16], an impulsive force was used in order to perturb one attractor of the system and to change its response such that the system response evolved to another attractor. In [17], control of multistability scenarios based on the selection of a particular attractor by periodic external modulation was presented. A complete report of control of multistability can be found in [18].

In this paper, we prove a different technique to control domains of attraction in multistability scenarios which is called Fixed-Point Inducting Controller (FPIC). The FPIC is a feed-forward controller that forces the system solution to evolve to an existing desired attractor which cannot be always reached for the uncontrolled case because of initial conditions. This technique was initially proposed in [19] and successfully applied in [20, 21]. In particular, in [21], experiments showed the good performance of the controlled system, where the FPIC was used as a second control loop. However, in all previous works published until now, FPIC has been only used to stabilize unstable periodic orbits.

The main contributions of this paper lie in numerical analysis and control design areas. The development of a novel methodology to compute and analyze the basins of attraction reinforced the numerical analysis. The methodology consists in decomposing the map of the system in p submaps, where p corresponds to the order of the periodic orbit to be controlled. To have the BA, the long-time response terms of this submap are depicted according to key colours. Apart from help to the control design, this way to compute the basin of attraction can be seen as the basins of attraction of period-1 orbits which are computed every sixth or fifth iteration for linear and nonlinear coupling, respectively, and the control goal can be thought of as the control of period-1 orbit. On the other hand, the methodology to design the controller based on FPIC technique belongs to control theory. When the proposed controller is used, the system is forced to evolve to a known periodic orbit that exists in the uncontrolled map. Moreover, by using the information obtained from bifurcation diagrams and basins of attraction, it is possible to compute the minimum control effort required to stabilize the target orbit in the defined region. In particular, the coexistence of periodic solutions in coupled logistic maps [8, 9, 22] is controlled by widening the basin of attraction of a period-p orbit that coexists with another one, and the minimum control effort is computed aided by the unstable period-2 orbit for the linear coupling and by unstable period-1 orbits for the nonlinear coupling case.

The paper is organized as follows. In Section 2, the coupled logistic maps are presented. The coexistence of period-6 and period-5 orbits is identified. In Section 3, the methodology to compute the basins of attraction is explained, and the procedure is applied to the linear and nonlinear coupling maps. In Section 4, the methodology to compute the controller is presented and applied to the considered systems and the minimum control effort required to guarantee monostabilization of the periodic orbit is computed. In Section 5, a brief discussion of the results is presented, and finally, in Section 6, the conclusions are given.

2. Coupled Logistic Maps

Coupled chaotic maps provide a source of bifurcation scenarios with nonlocal phenomena and coexistence of attractors. To design the proposed controller, two maps with different coupling mechanisms are chosen. For linear coupling map, there are two period-6 orbits that coexist in a range of the parameter set. Similarly, for nonlinear coupling case, two period-5 orbits coexist.

2.1. Coupled Logistic Maps: Linear Coupling. This system is described by

$$\begin{aligned} x(k+1) &= f_x(k) + \epsilon\left(y(k) - x(k)\right), \\ y(k+1) &= f_y(k) + \epsilon\left(x(k) - y(k)\right), \end{aligned} \tag{1}$$

where $\epsilon \in (-2, 2)$ is the coupling constant and $f_z(k) = rz(1-z)$, $r \in (1, 4)$, is the parameter associated with the nonlinear part of (1). Since the system is symmetrical, there is an invariant line Δ and the restriction of the 2D map to Δ reduces it to a 1D map (the logistic map). The dynamics on this invariant set help us to study the dynamics and bifurcations of the 2D system. Moreover, symmetrical trajectories are generated by symmetrical initial conditions leading to symmetrical basins of attraction [22]. Novel routes to chaos through torus-breakdown mechanism of this system were reported in [9] and different dynamics were characterized in the parameter region given by (ϵ, r).

Figure 1 shows the dynamic behavior of (1) when r is varied and two symmetrical initial conditions are considered. Note that a pair of coexisting period-6 orbits are identified in these diagrams for $r \in [3.045, 3.065]$, which makes the main difference between Figures 1(a) and 1(b). The dynamical behavior in the rest of the interval is very similar and other differences cannot be easily seen. Table 1 shows the specific values of the two coexisting period-6 orbits. It can be observed that the solutions are mutually symmetrical in the sense that $(x, y) = (y, x)$.

2.2. Coupled Logistic Maps: Nonlinear Coupling. This system is defined by

$$\begin{pmatrix} x(k+1) \\ y(k+1) \end{pmatrix} = \begin{pmatrix} f_x(k) + \epsilon\left(f_y(k) - f_x(k)\right) \\ f_y(k) + \epsilon\left(f_x(k) - f_y(k)\right) \end{pmatrix}, \tag{2}$$

where $f_x(k)$ and $f_y(k)$ were defined before, $\epsilon \in (-2, 2)$, and $r \in (1, 4)$. Figure 2 shows an interesting coexistence scenario for this system. Two mutually symmetrical period-5 cycles coexist in it when $\epsilon = 1.5$ and $r \in [2.775, 2.815]$.

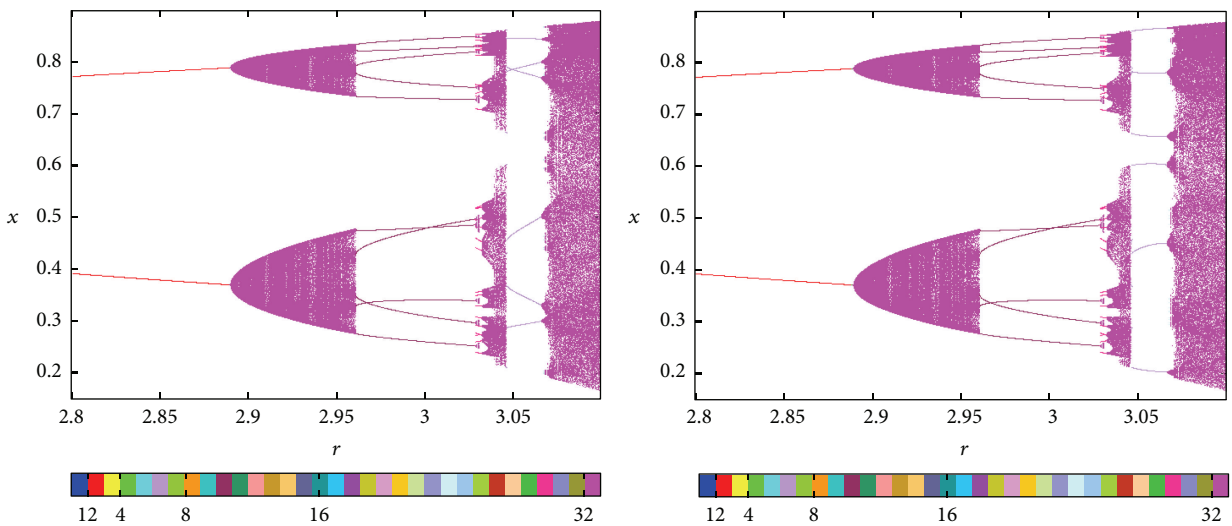
3. Methodology to Compute the Basins of Attraction

In this section, a methodology to compute the basins of attraction (BA) is developed. By using this methodology, it is possible to find the minimum control effort required to stabilize a desired periodic orbit.

3.1. Computation of the Basins of Attractions in Linear Coupling Maps. The BA corresponds to the set of initial conditions whose long-time response (LTR) approaches the attractor. The first and second period-6 orbits were defined in Table 1. All numerical results associated with the second period-6 orbit are differentiated from the first one with an

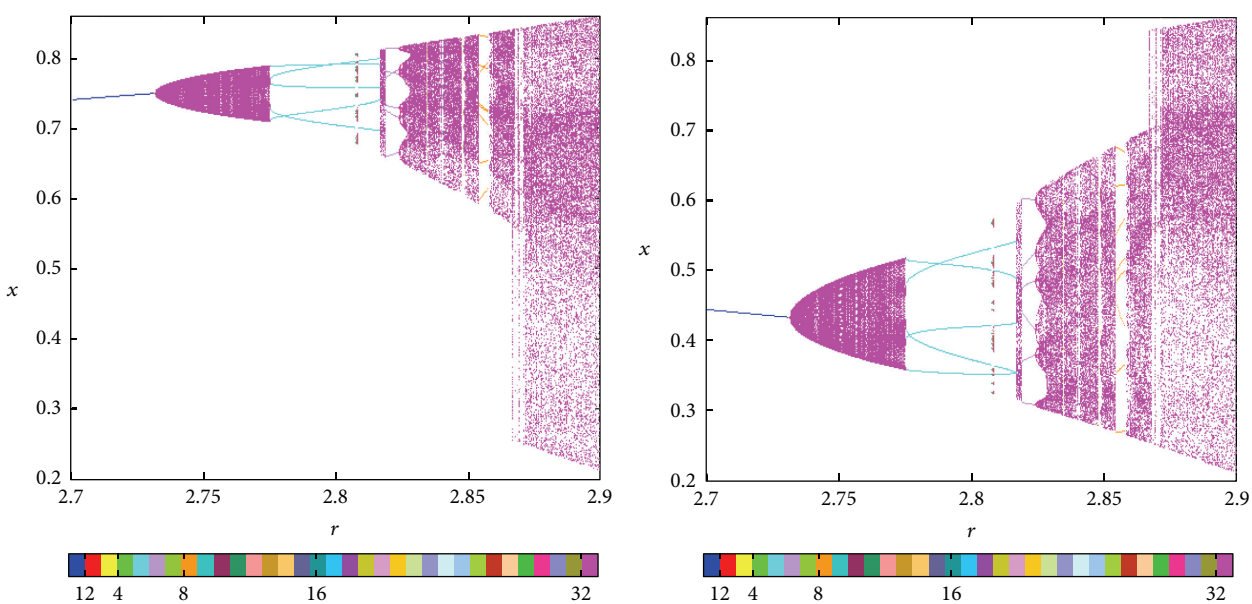
Table 1: Two coexisting period-6 orbits for the linear coupling maps when $\epsilon = 0.2725$ and $r = 3.065$. For the sake of simplicity, the periodic orbit values are presented with 8 decimal digits.

First period-6 orbit	Second period-6 orbit
$(x^*, y^*)_1 = (0.79744636, 0.20236605)$	$(\overline{x}^*, \overline{y}^*)_1 = (0.20236605, 0.79744636)$
$(x^*, y^*)_2 = (0.33291675, 0.65689339)$	$(\overline{x}^*, \overline{y}^*)_1 = (0.65689339, 0.33291675)$
$(x^*, y^*)_3 = (0.76896860, 0.60251974)$	$(\overline{x}^*, \overline{y}^*)_1 = (0.60251974, 0.76896860)$
$(x^*, y^*)_4 = (0.49915797, 0.77939325)$	$(\overline{x}^*, \overline{y}^*)_1 = (0.77939325, 0.49915797)$
$(x^*, y^*)_5 = (0.84261193, 0.45063017)$	$(\overline{x}^*, \overline{y}^*)_1 = (0.45063017, 0.84261193)$
$(x^*, y^*)_6 = (0.29965625, 0.86559446)$	$(\overline{x}^*, \overline{y}^*)_1 = (0.86559446, 0.29965625)$



(a) $(x_0, y_0) = (0.1, 0.5)$ (b) $(x_0, y_0) = (0.5, 0.1)$

Figure 1: Bifurcation diagrams of (1) with $\epsilon = 0.2725$ and two symmetrical initial conditions.



(a) (b)

Figure 2: Bifurcation diagrams of (2) with $\epsilon = 1.5$ and two symmetrical initial conditions.

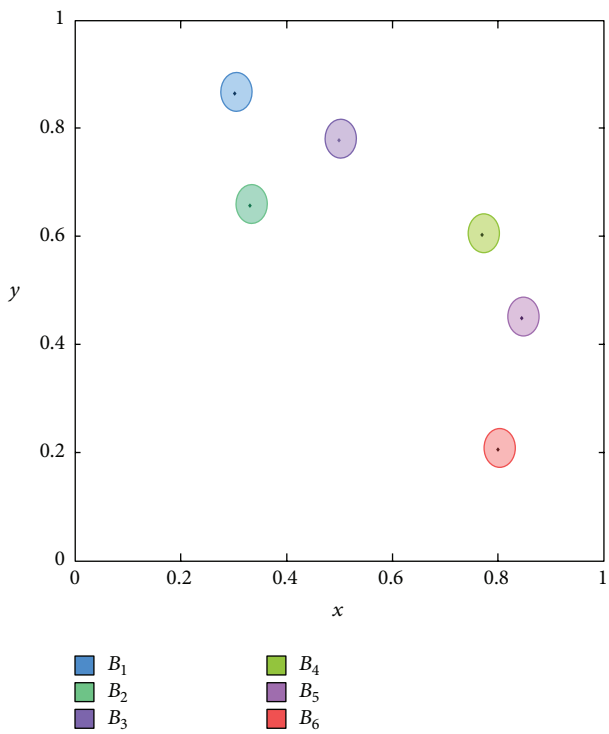


Figure 3: Colour key to compute the BA.

upper bar. In order to compute the BA and after applying FPIC control, we associate with each point k of the first period-6 orbit a section B_k that is given by $|x_{\mathrm{LTR}} - x^*| < \delta$, such as what is illustrated in Figure 3. A similar procedure is developed to compute the BA of the second period-6 orbit, whose sections are defined as $\overline{B}_k$.

To compute the LTR terms, we start from any initial condition inside the considered region ($0 < x_1 < 1$ and $0 < x_2 < 1$) and compute its successive iterations until the solution approaches the attractor. The general map to describe the evolution of the system is P, and it is given by

$$\begin{aligned} P := \mathbf{x}(0) &\longmapsto \mathbf{x}(1) \longmapsto \mathbf{x}(2) \longmapsto \mathbf{x}(3) \longmapsto \mathbf{x}(4) \\ &\longmapsto \mathbf{x}(5) \longmapsto \mathbf{x}(6) \longmapsto \mathbf{x}(7) \longmapsto \mathbf{x}(8) \longmapsto \mathbf{x}(9) \\ &\longmapsto \cdots \mathbf{x}(m), \end{aligned} \tag{3}$$

where m is high enough. The map P is decomposed in six submaps P_i. Each P_i is formed by the set of points given by $\mathbf{x}(i-1+6j)$, where $i = 1 \cdots 6$ and $j \in \mathbb{Z}^+ + \{0\}$. If the LTR terms of P_i lie inside some B_k (or $\overline{B}_k$ when the BA for the second period-6 orbit is computed), it is coloured according to the colour key displayed in Figure 3. If the LTR terms of P_i lie outside these regions, then it is gray coloured. This procedure is repeated for a lot of initial conditions.

The BA of the first period-6 orbit (second period-6 orbit) can be computed and depicted only using P_i ($\overline{P}_i$) submaps for a fixed i, and its diagram may have up to seven colors. The gray coloured points correspond to the set of initial conditions whose LTR terms diverge from the first period-6 orbit (second period-6 orbit), and the other points (depicted with six different colors according to Figure 3) correspond to the initial conditions whose LTR terms lie inside some B_k ($\overline{B}_k$) region. The BA computed for the first period-6 orbit (second period-6 orbit) in this way is noted as $\mathrm{BA}(P_i)$ ($\mathrm{BA}(\overline{P}_i)$). Figure 4(a) shows the diagram associated with $\mathrm{BA}(P_1)$. All $\mathrm{BA}(P_i)$ have the same information but are coloured in different way. Figure 4(b) shows $\mathrm{BA}(\overline{P}_1)$. These BA let us compute the minimum control effort required to stabilize the periodic orbits.

3.2. Computation of the Basins of Attraction for Nonlinear Coupling Maps. The procedure explained before is applied to the nonlinear coupling maps. $\mathrm{BA}(P_i)$ and $\mathrm{BA}(\overline{P}_i)$ of the two period-5 orbits are also symmetrical. Figures 5(a) and 5(b) show $\mathrm{BA}(P_1)$ and $\mathrm{BA}(\overline{P}_1)$, respectively. The remaining diagrams $\mathrm{BA}(P_i)$ and $\mathrm{BA}(\overline{P}_i)$ for $i = 2 \cdots 5$ have similar behavior. As we said before, all $\mathrm{BA}(P_i)$ carry the same information but are coloured in different way.

4. Methodology to Control the Basins of Attraction

In this section, domains of attraction in coexistence scenarios are controlled using FPIC technique. This technique has been proven to stabilize unstable and chaotic systems [19–21]. In those papers, analytical or numerical values of the steady-state control inputs were needed to guarantee the stabilization of the equilibrium point because FPIC was used as a second control loop. On the contrary, in this work, FPIC technique is used to control the BA in coupled logistic maps to guarantee stabilization of period-6 (for linear coupling) and period-5 (for nonlinear coupling) orbits, using the values of the periodic orbit to be controlled.

4.1. Fixed-Point Inducting Control

Theorem 1. *Consider a discrete-time system*

$$\mathbf{z}(k+1) = \mathbf{g}(\mathbf{z}(k)). \tag{4}$$

It is assumed that period-p solutions coexist in the system and two or more of them are stable. A stable period-p solution is noted in a compact form as $\mathbf{z}^*$, and it is composed of p points such that it can be expressed as

$$\begin{aligned} \mathbf{z}^* = \left\{z_1^*, z_2^*, \ldots, z_p^*\right\} = \left\{\left(x_1^*, x_2^*, \ldots, x_n^*\right)_1, \right. \\ \left. \left(x_1^*, x_2^*, \ldots, x_n^*\right)_2, \ldots, \left(x_1^*, x_2^*, \ldots, x_n^*\right)_p\right\}. \end{aligned} \tag{5}$$

The objective of the controller is to expand the domain of attraction of a given period-p orbit, by forcing the system to evolve to this specific solution. Equation (6) shows the control law based on FPIC, with E as the control parameter. This

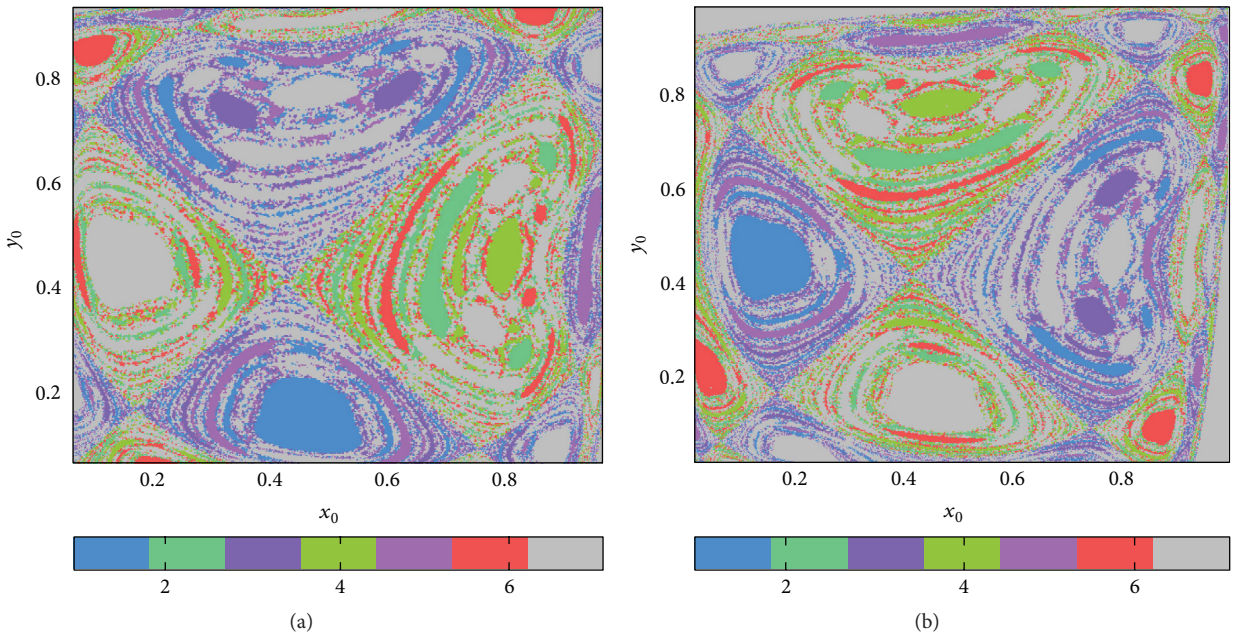


FIGURE 4: (a) BA(P_1). (b) BA($\overline{P}_1$). $r = 3.065$ and $\epsilon = 0.2725$.

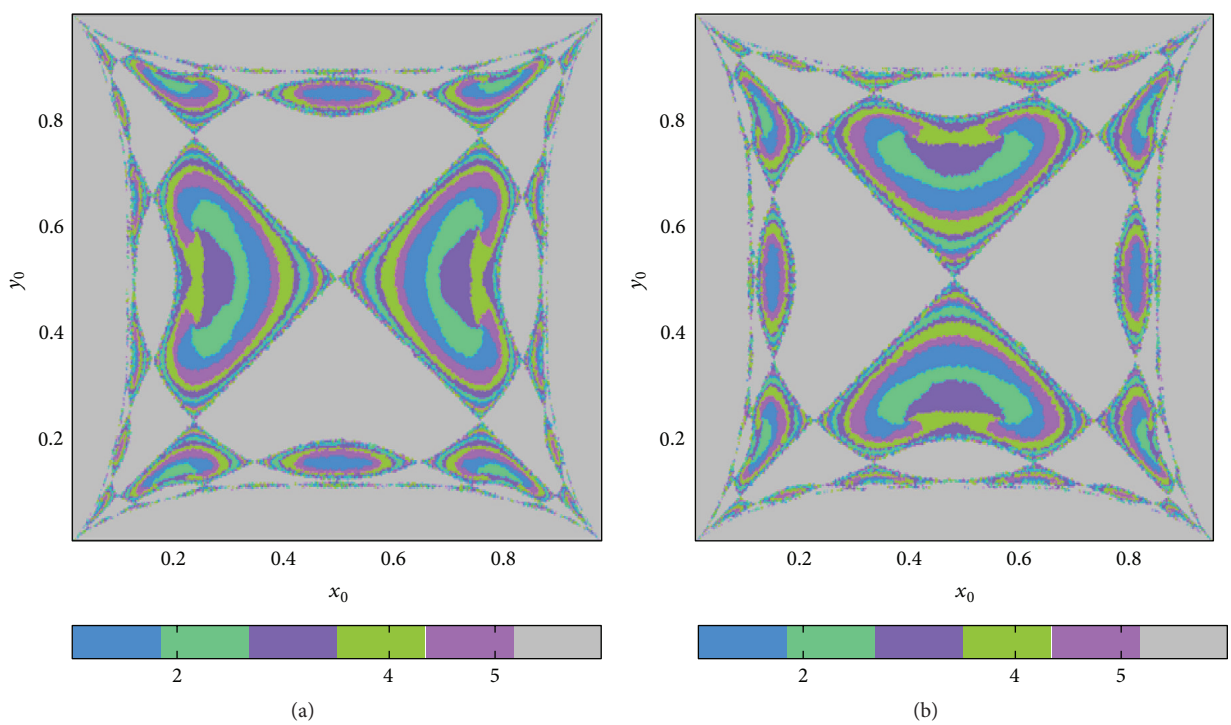


FIGURE 5: (a) BA(P_1). (b) BA($\overline{P}_1$). $r = 2.8$ and $\epsilon = 1.5$.

controller stabilizes the target periodic orbit and expands its attraction domain as E increases. Hence,

$$\mathbf{z}(k+1) = \begin{cases} \dfrac{1}{E+1}\mathbf{g}(\mathbf{z}(k)) + \dfrac{E}{E+1}\mathbf{z}_j^* & \text{if } (k \bmod p) = 0 \\ \mathbf{g}(\mathbf{z}(k)) & \text{if } (k \bmod p) \neq 0, \end{cases} \quad (6)$$

where j is defined depending on the analysis of the controlled system.

Proof. The periodic orbits of the controlled system are the same as the uncontrolled one. This is because after p iterations $\mathbf{g}(z_j^*) = z_j^*$ (by definition of periodic orbit) and thus considering only the upper part of (6), it can be seen that the periodic

orbits remain unchanged. Hence, when E is high enough, the controller transfers the state instantly to some B_k when $k \bmod p = 0$, irrespective of the state value. In this way, the new state (that is near to the periodic solution) is the initial condition for the next iteration obtaining convergence to the periodic solution because the periodic orbit is stable. □

Remarks. (i) From a practical point of view, the idea is not to fix E to a high value because the control effort increases as E does. To control systems whose initial conditions are away from the periodic orbit can lead to oversizing of the control effort. For example, for the systems analyzed in this work, the solution quickly diverges when the initial conditions are outside the considered region and the controller cannot be used.

(ii) The minimum value of E that should be put in (6) can be found based on numerical analysis, where the unstable period orbits that coexist in the dynamical system play an important role.

4.2. Control of Period-6 Orbits in Linear Coupling Maps. Let us consider the map defined by (1) with $r = 3.065$ and $\epsilon = 0.2725$. We will stabilize the second period-6 orbit using FPIC control. In order to achieve the goal, we use the second point $(\overline{x}^*, \overline{y}^*)_2$ as the target control; that is, $\overline{\mathbf{x}}_2^* := (0.6569, 0.3329)$. Taking into account (6), the linear coupling map is perturbed every sixth iteration according to the control law defined by

$$\begin{pmatrix} x(k+1) \\ y(k+1) \end{pmatrix} = \begin{cases} \dfrac{1}{E+1}\begin{pmatrix} f_x(k) + \varepsilon\left(y(k) - x(k)\right) \\ f_y(k) + \varepsilon\left(x(k) - y(k)\right) \end{pmatrix} + \dfrac{E}{E+1}\begin{pmatrix} \overline{x}^* \\ \overline{y}^* \end{pmatrix} & \text{if } (k \bmod 6) = 0 \\ \begin{pmatrix} f_x(k) + \varepsilon\left(y(k) - x(k)\right) \\ f_y(k) + \varepsilon\left(x(k) - y(k)\right) \end{pmatrix} & \text{if } (k \bmod 6) \neq 0. \end{cases} \tag{7}$$

Figure 6 shows $\mathrm{BA}(\overline{P}_1)$ diagram for different E values. Colors were assigned as defined above. The elements of $\mathrm{BA}(\overline{P}_1)$ that tends to $\overline{B}_2$ increase as parameter E increases. Comparing Figures 4(b) and 6, it can be seen that the green coloured region enhances forming clear narrow contours. Now, the BA is only formed by two sets of points. The first one includes all points whose LTR terms end in $\overline{B}_2$ which are green coloured. The second one includes all points whose LTR terms are not inside any $\overline{B}_i$ section for $i = 1, 2, 3, 4, 5, 6$, which are gray coloured. This implies that no long-time term of the submap $\overline{P}_1$ lies in $\overline{B}_i$ section, for $i = 1, 3, 4, 5, 6$. The green contours expand as E increases, widening the BA and diminishing the gray coloured parts.

Figure 7 shows $\mathrm{BA}(P_1)$ when $\mathbf{x}_2^* := (0.3329, 0.6569)$ is used as the reference in the FPIC controller. The same values of E have been used.

4.2.1. Selection of Minimum Control Effort. The minimum control effort E_{cr} is defined as the minimum value of E that forces the system to evolve to the controlled orbit. The selection of E_{cr} can be computed by studying the behavior of unstable periodic orbits (UPOs) around the periodic orbits.

The metamorphosis of controlled $\mathrm{BA}(\overline{P}_1)$ or $\mathrm{BA}(P_1)$ is influenced by two unstable period-2 solutions of the system that are schematized in Figure 8. The fixed points of these period-2 orbits are named V_1, V_2, V_3, and V_4.

For the controlled system, the points V_i are hard to cover by the expansion of $\mathrm{BA}(P_1)$. These points are always on the frontiers between the green and gray regions as it is shown in Figure 9. To find E_{cr}, we consider Δ_i region around each point V_i. E_{cr} is the minimum value such that all points in Δ_i^0 regions converge to V_i points. Figure 9 shows examples of Δ_i regions when $E = 1.5$. The neighborhood near V_i requires more control effort to converge to B_2.

Δ_i can be expressed as $(\Delta_x, \Delta_y) := (v_x + \delta\sin(\theta), v_y + \delta\cos(\theta))$, where δ is fixed to a small value and θ is varied in the range $[0, 2\pi]$. Figure 10 illustrates the behavior of $\mathrm{BA}(P_1)$ around Δ_i when E is varied. The convergence of all points to Δ_i is reached when $E_{\mathrm{cr}} \approx 1.78$.

4.2.2. Selection of the Reference. Similar results can be obtained if we choose other equilibrium points to perturb the map P_i. Table 3 summarizes the control effort requirements. Note that the critical control effort E_{cr} varies depending on the used reference. Coexistence phenomenon can be controlled with less effort if $(x^*, y^*) = (0.8426, 0.4506)$ or $(\overline{x}^*, \overline{y}^*) = (0.4506, 0.8426)$ are used to control the map.

4.3. Control of Period-5 Orbits in Nonlinear Coupling Maps. Now, we control two coexisting period-5 cycles of (2) when $r = 2.8$ and $\epsilon = 1.5$. The reference is introduced every fifth iteration as is presented in

$$\begin{pmatrix} x(k+1) \\ y(k+1) \end{pmatrix} = \begin{cases} \dfrac{1}{E+1}\begin{pmatrix} f_x(k) + \varepsilon\left(f_y(k) - f_x(k)\right) \\ f_y(k) + \varepsilon\left(f_x(k) - f_y(k)\right) \end{pmatrix} + \dfrac{E}{E+1}\begin{pmatrix} \overline{x}^* \\ \overline{y}^* \end{pmatrix} & \text{if } (k \bmod 5) = 0 \\ \begin{pmatrix} f_x(k) + \varepsilon\left(f_y(k) - f_x(k)\right) \\ f_y(k) + \varepsilon\left(f_x(k) - f_y(k)\right) \end{pmatrix} & \text{if } (k \bmod 5) \neq 0. \end{cases} \tag{8}$$

Table 2: Two coexisting period-5 orbits for nonlinear coupling maps when $\epsilon = 1.5$ and $r = 2.8$. For the sake of simplicity, the periodic orbit values are presented with 8 decimal digits.

First period-5 orbit	Second period-5 orbit
$(x^*, y^*)_1 = (0.79175845, 0.35258819)$	$(\overline{x}^*, \overline{y}^*)_1 = (0.35258819, 0.79175845)$
$(x^*, y^*)_2 = (0.72790518, 0.37290574)$	$(\overline{x}^*, \overline{y}^*)_2 = (0.37290574, 0.72790518)$
$(x^*, y^*)_3 = (0.70487469, 0.50446287)$	$(\overline{x}^*, \overline{y}^*)_3 = (0.50446287, 0.70487469)$
$(x^*, y^*)_4 = (0.75867944, 0.52373859)$	$(\overline{x}^*, \overline{y}^*)_4 = (0.52373859, 0.75867944)$
$(x^*, y^*)_5 = (0.79131428, 0.41974570)$	$(\overline{x}^*, \overline{y}^*)_5 = (0.41974570, 0.79131428)$

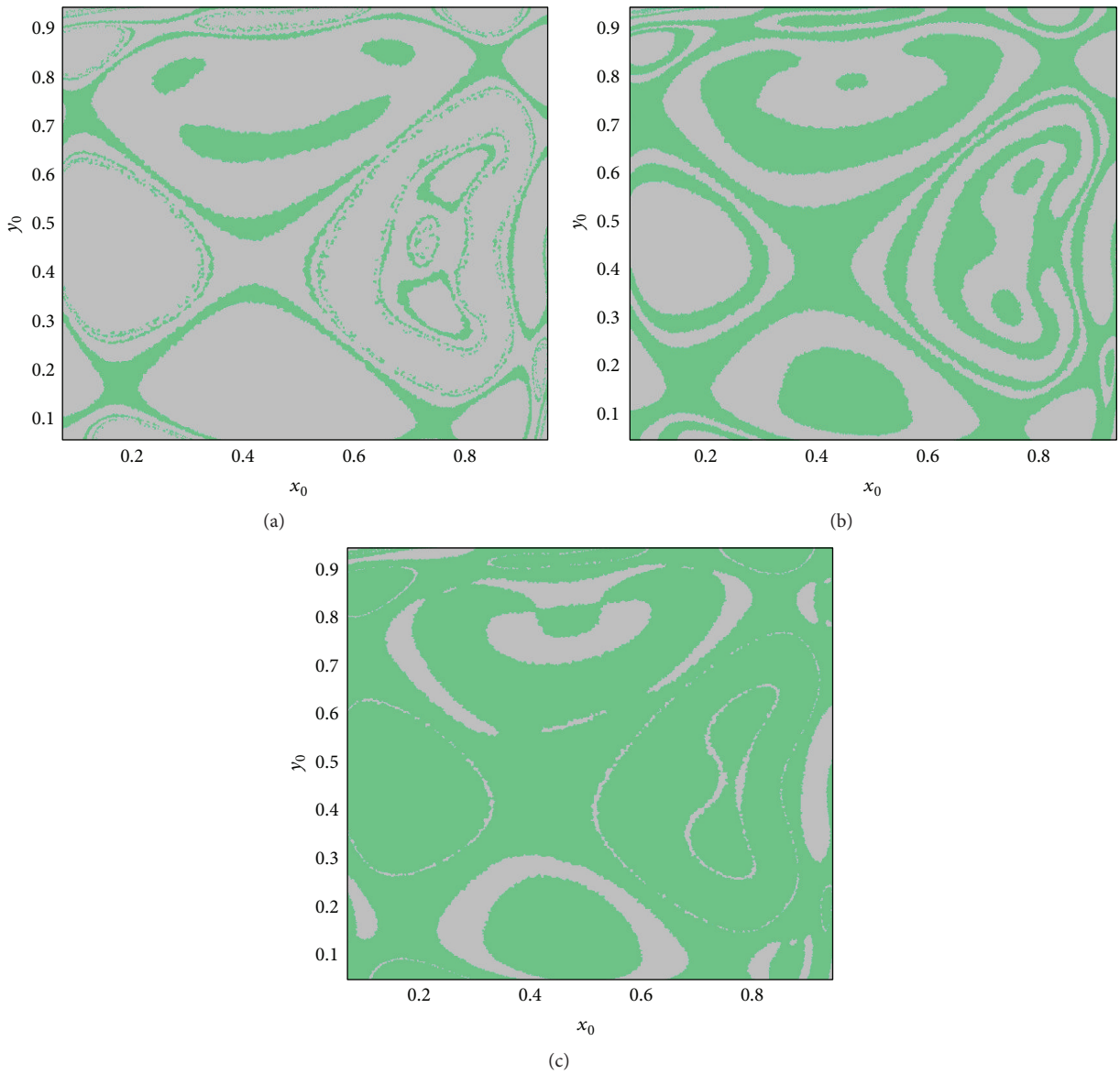


Figure 6: (a) $\mathrm{BA}(\overline{P}_1)$ for $E = 0.85$. (b) $\mathrm{BA}(\overline{P}_1)$ for $E = 1.5$. (c) $\mathrm{BA}(\overline{P}_1)$ for $E = 1.7$. $r = 3.065$ and $\epsilon = 0.2725$ in all cases.

According to Table 2, we use $(\overline{x}^*, \overline{y}^*)_1 := (0.3526, 0.7918)$ to control the system. Figure 11 presents $\mathrm{BA}(\overline{P}_1)$ of the controlled system when three different values of E are considered.

Selection of Minimum Control Effort. As in the previous case, the minimum control effort E_{cr} can be computed by studying the behavior of UPOs around the periodic orbits. Figure 12 shows an illustrative sketch of bifurcation scenario of the uncontrolled system. The unstable period-1 orbits are noted as N_1 and N_2. The metamorphosis of the controlled maps $\mathrm{BA}(\overline{P}_i)$ is influenced by unstable period-1 solutions of the system. We use points N_1 and N_2 to compute E_{cr} in each case.

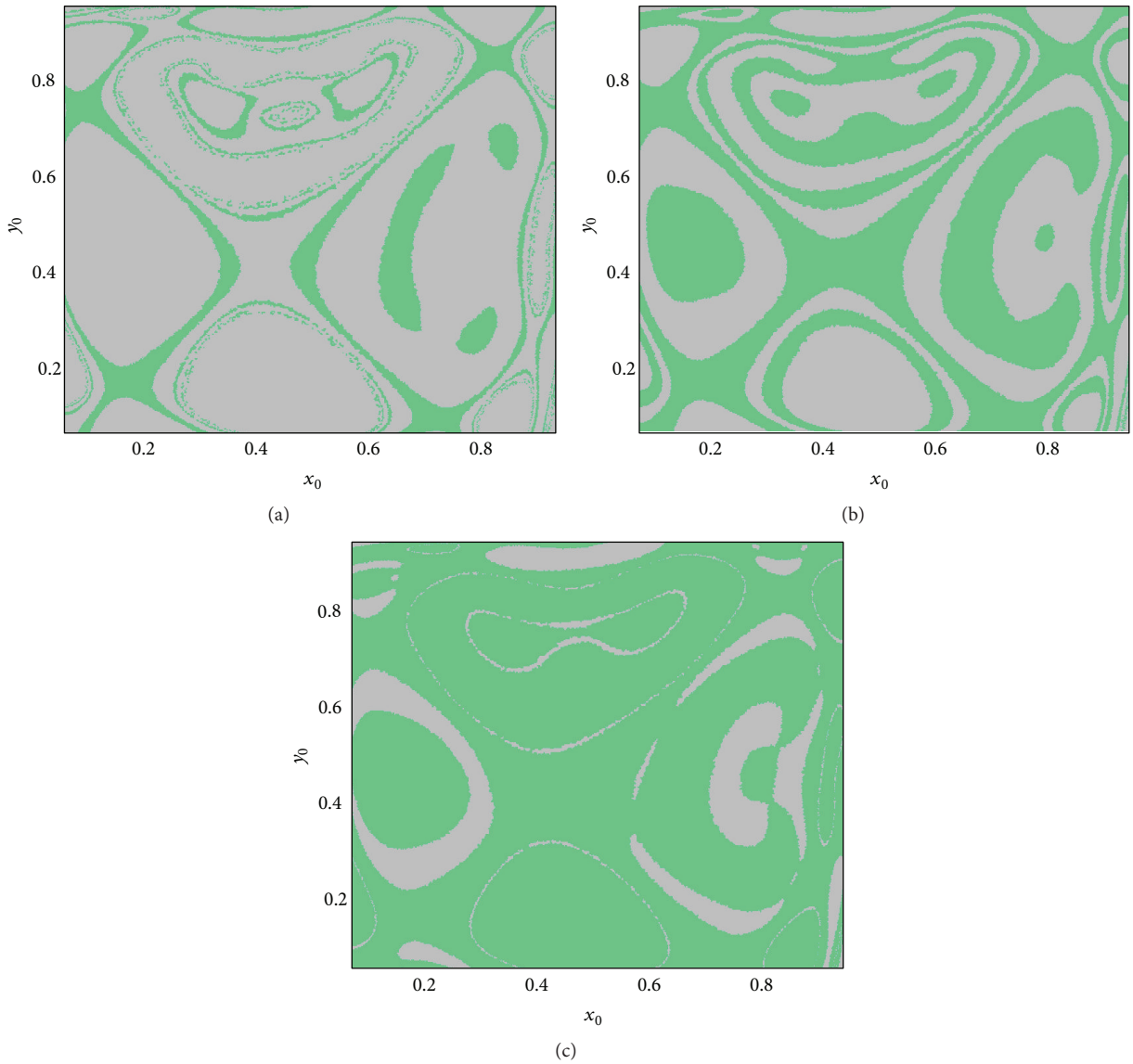


FIGURE 7: (a) BA(P_1) for $E = 0.85$. (b) BA(P_1) for $E = 1.5$. (c) BA(P_1) for $E = 1.7$. $r = 3.065$ and $\epsilon = 0.2725$ for all cases.

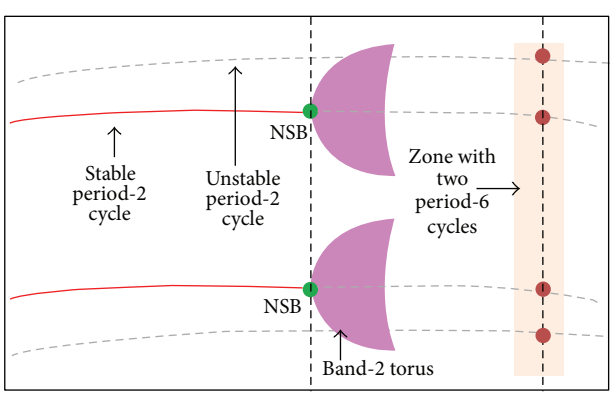


FIGURE 8: Sketch of unstable period-2 solutions V_i in coexistence zone.

Table 4 presents the minimum value of parameter E to control coexistence phenomenon. All initial conditions belonging to the considered region converge to $\overline{\mathbf{x}}^*$ when $E > E_{cr}$. The lowest control effort that avoids coexistence phenomenon is reached when the map is controlled with the first point of the periodic orbit. If we induce the control using other points of the periodic orbit, then we need a greater control effort. Figure 13 shows the symmetrical behavior between BA(P_1) and BA($\overline{P}_1$) (compare with Figure 11). Critical control effort required in both cases has the same value.

5. Discussion

The results presented above give rise to a new approach for coexistence control in multistable systems. The proposed

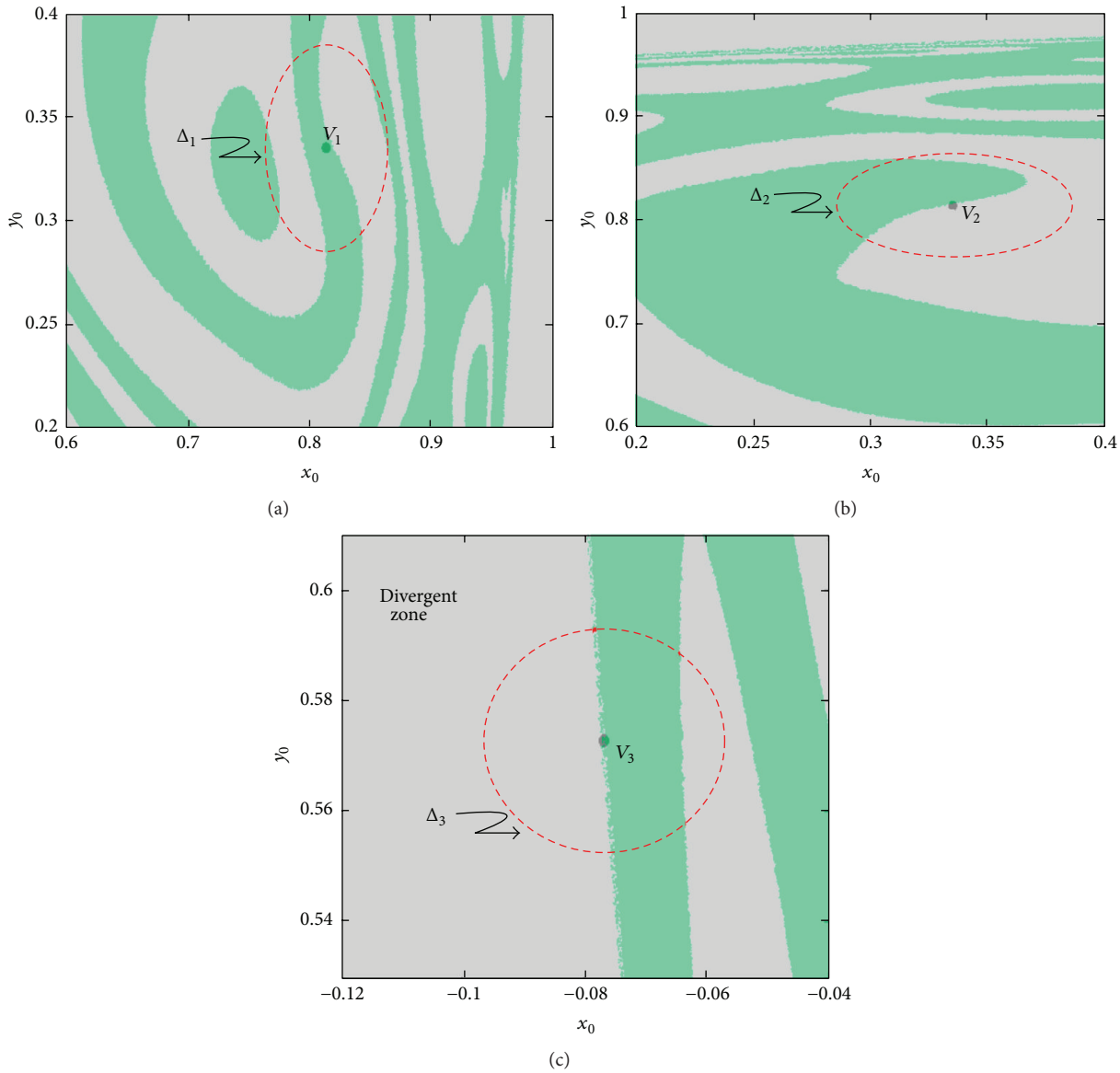


Figure 9: V_i points and Δ_i sections in controlled linear map with $E = 1.5$. (a) $V_1 = (0.81326798, 0.33518226)$. (b) $V_2 = (0.33518226, 0.81326798)$. (c) $V_3 = (-0.07696298, 0.57288468)$.

Table 3: Summary of FPIC requirements in coexistence control of two period-6 cycles.

Maps	$\mathbf{x}^*$	$\overline{\mathbf{x}}^*$	E_{cr}
$P_1, \overline{P}_1$	(0.7974, 0.2023)	(0.2023, 0.7974)	3.371
$P_2, \overline{P}_2$	(0.3329, 0.6568)	(0.6568, 0.3329)	1.774
$P_3, \overline{P}_3$	(0.7689, 0.6025)	(0.6025, 0.7689)	2.507
$P_4, \overline{P}_4$	(0.4991, 0.7793)	(0.7793, 0.4991)	1.492
$P_5, \overline{P}_5$	(0.8426, 0.4506)	(0.4506, 0.8426)	1.258
$P_6, \overline{P}_6$	(0.2996, 0.8655)	(0.8655, 0.2996)	4.112

Table 4: Summary of FPIC requirements in coexistence control of two period-5 cycles.

Maps	$\mathbf{x}^*$	$\overline{\mathbf{x}}^*$	E_{cr}
$P_1, \overline{P}_1$	(0.7918, 0.3526)	(0.3526, 0.7918)	0.6099
$P_2, \overline{P}_2$	(0.7279, 0.3729)	(0.3729, 0.7279)	0.6873
$P_3, \overline{P}_3$	(0.7049, 0.5045)	(0.5045, 0.7048)	1.2312
$P_4, \overline{P}_4$	(0.7588, 0.5237)	(0.5237, 0.7588)	1.1892
$P_5, \overline{P}_5$	(0.7913, 0.4197)	(0.4197, 0.7913)	0.7504

methodology has significant differences with other methods reported in the literature, some of which are commented on below.

(i) The Applicability of the Method. Our process is sequential, systematic, and applicable to any multistable system, even though we use low-dimensional discrete maps as illustrative

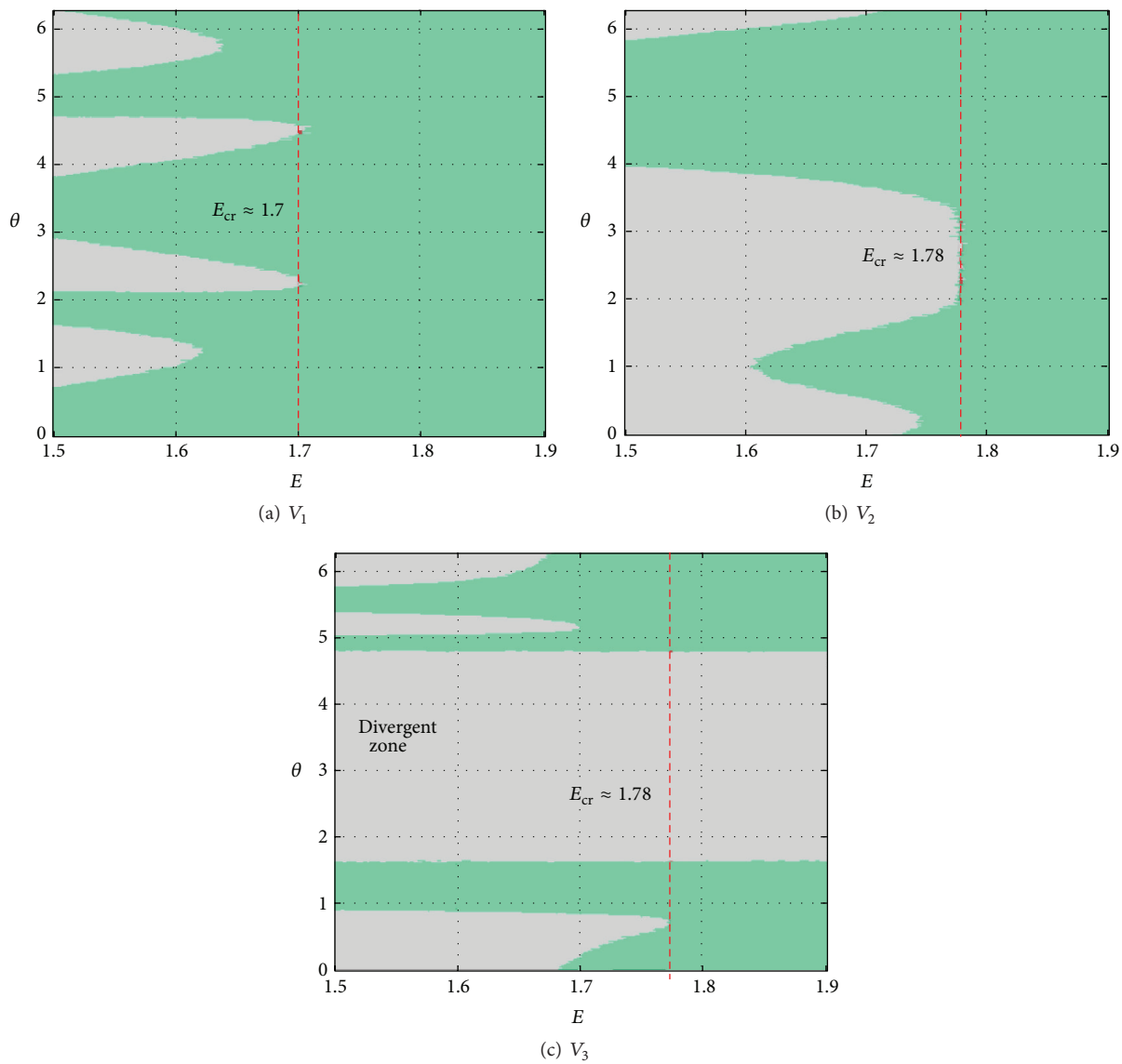


Figure 10: Selection of critical value E_{cr}. Diagram E versus θ to select minimum value of E that guarantees total convergence of P_1 to $x^* = 0.3329$ and $y^* = 0.6568$.

examples. One of the most common methods to control multistability, which is based on slow harmonic modulation, has restrictions on its application because the attractor to be destroyed should be a focus type [23].

(ii) The Requirement of Dynamic Behavior Knowledge. Fixed-point inducting control needs to know an approximate value of the periodic orbit to be stabilized, while all nonfeedback methods do not require a priori information of the system to be controlled [18]. However, this is not a major problem due to the simplicity of the required information and the multiple ways to obtain it, by either theoretical, empirical, or experimental sources.

(iii) The Control Scheme, the Choice of the Attractor, and the Number of Parameters to Be Tuned. The most representative methods for multistability control can be classified into three categories: *nonfeedback*, *feedback*, and *stochastic* schemes [18]. We propose a *feed-forward* control configuration which has not been widely exploited. Particularly, the control law based on FPIC criterion has advantages with respect to the selection of the target orbit and the parameters setting. For example, when the coexistence control is realized by parameter modulation, the general form given by $p = p_0 + p_c \sin(2\pi f_c t + \psi_0)$ has four parameters to be tuned [24]; then, the selection of a particular attractor depends on the right choice of these parameters. In contrast, to successfully apply

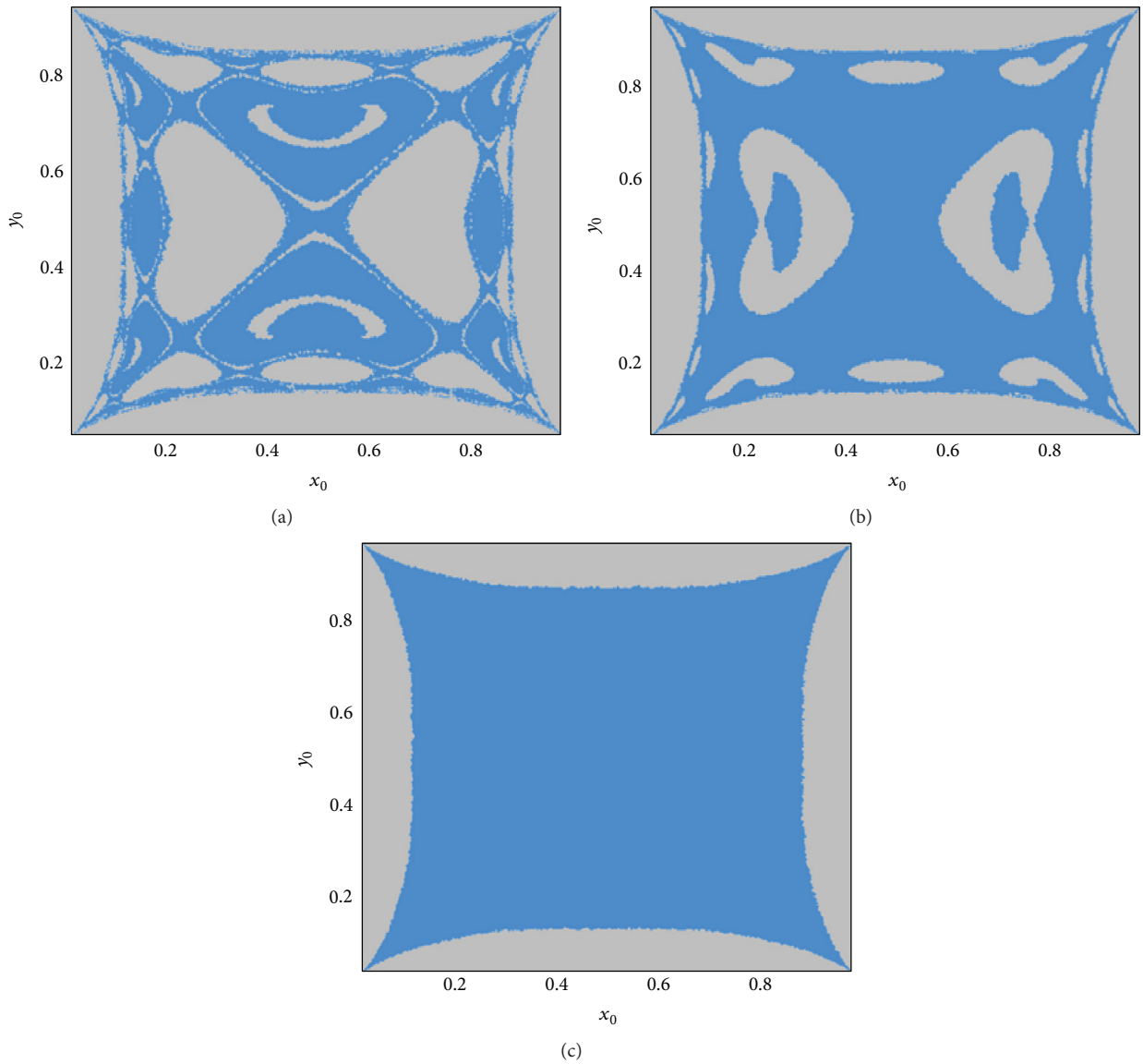


FIGURE 11: BA($\overline{P}_1$) for $E = 0.25$. (b) BA($\overline{P}_1$) for $E = 0.45$. (c) BA($\overline{P}_1$) for $E = 0.65$. $r = 2.8$ and $\epsilon = 1.5$ for all cases.

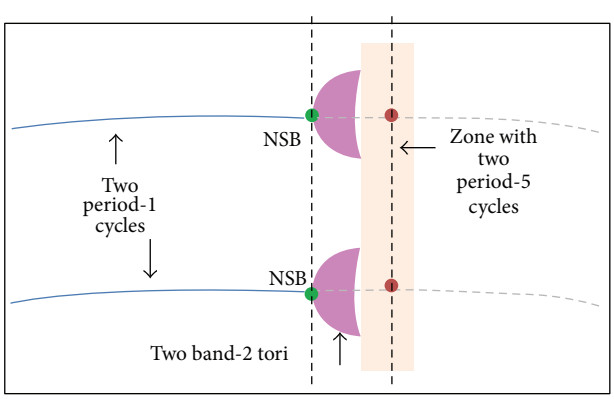


FIGURE 12: Sketch of unstable period-1 solutions N_i in coexistence zone.

FPIC control only it is needed to tune one parameter and to know the target orbit.

6. Conclusions

The coexistence of periodic solutions has been controlled using FPIC technique and the required control effort to obtain stabilization of the periodic orbits was computed.

It has been proven that FPIC controller can be used to control BA. FPIC requires virtually no information about the system; just approximated knowledge of a periodic orbit is needed. Thus, this is a good control technique if the system is only known approximately, which is a fundamental requirement for other control methods.

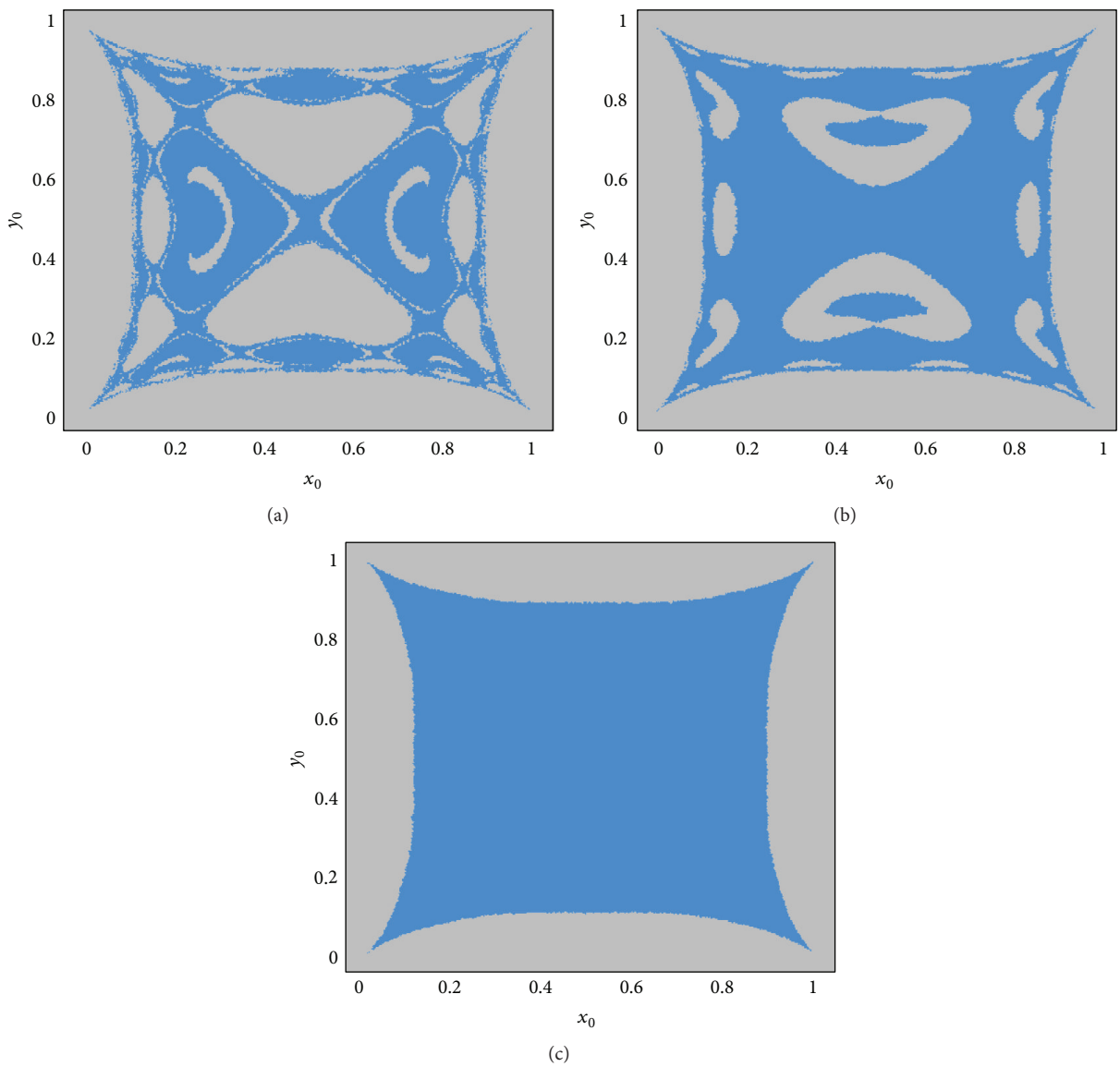


Figure 13: BA(P_1) for $E = 0.25$. (b) BA(P_1) for $E = 0.45$. (c) BA(P_1) for $E = 0.65$. $r = 2.8$ and $\epsilon = 1.5$ for all cases.

The applicability of the proposed tools has been illustrated in linear and nonlinear coupling logistic maps. However, the developed tools are generic and can be applied to any system with complex scenarios.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was partially supported by Universidad Nacional de Colombia, Manizales, Projects 28528 and 22546, Vicerrectoría de Investigación, DIMA.

References

[1] E. Ott, *Chaos in Dynamical Systems*, Cambridge University Press, Cambridge, UK, 2002.

[2] R. L. Devaney, *An Introduction to Chaotic Dynamical Systems*, Addison-Wesley, Reading, Mass, USA, 2nd edition, 1989.

[3] E. Mosekilde, *Topics in Nonlinear Dynamics—Applications to Physics, Biology and Economic Systems*, World Scientific, Singapore, 1996.

[4] S. Wiggins, *Global Bifurcations and Chaos: Analytical Methods*, Springer, New York, NY, USA, 1988.

[5] Z. T. Zhusubaliyev and E. Mosekilde, *Bifurcations and Chaos in Piecewise-Smooth Dynamical Systems*, World Scientific, 2003.

[6] S. Banerjee and G. Verghese, *Nonlinear Phenomena in Power Electronics Attractors, Bifurcations, Chaos, and Nonlinear Control*, IEEE Press, New York, NY, USA, 2001.

[7] A. Polynikis, M. di Bernardo, and S. J. Hogan, "Synchronizability of coupled PWL maps," *Chaos, Solitons and Fractals*, vol. 41, no. 3, pp. 1353–1367, 2009.

[8] Z. T. Zhusubaliyev and E. Mosekilde, "Multilayered tori in a system of two coupled logistic maps," *Physics Letters A*, vol. 373, no. 10, pp. 946–951, 2009.

[9] Z. T. Zhusubaliyev and E. Mosekilde, "Novel routes to chaos through torus breakdown in non-invertible maps," *Physica D: Nonlinear Phenomena*, vol. 238, no. 5, pp. 589–602, 2009.

[10] L. M. Pecora and T. L. Carroll, "Pseudoperiodic driving: eliminating multiple domains of attraction using chaos," *Physical Review Letters*, vol. 67, no. 8, article 945, 1991.

[11] K. Kaneko, "Dominance of milnor attractors and noise-induced selection in a multiattractor system," *Physical Review Letters*, vol. 78, no. 14, pp. 2736–2739, 1997.

[12] S. Kraut, U. Feudel, and C. Grebogi, "Preference of attractors in noisy multistable systems," *Physical Review E*, vol. 59, no. 5, pp. 5253–5260, 1999.

[13] A. N. Pisarchik and B. K. Goswami, "Annihilation of one of the coexisting attractors in a bistable system," *Physical Review Letters*, vol. 84, no. 7, pp. 1423–1426, 2000.

[14] B. K. Goswami and A. N. Pisarchik, "Controlling multistability by small periodic perturbation," *International Journal of Bifurcation and Chaos*, vol. 18, no. 6, pp. 1645–1673, 2008.

[15] A. N. Pisarchik and R. Jaimes-Reategui, "Control of basins of attraction in a multistable fiber laser," *Physics Letters A*, vol. 374, no. 2, pp. 228–234, 2009.

[16] Y. Liu, M. Wiercigroch, J. Ing, and E. Pavlovskaia, "Intermittent control of coexisting attractors," *Philosophical Transactions of the Royal Society A*, vol. 371, no. 1993, 15 pages, 2013.

[17] R. Sevilla-Escoboza, A. N. Pisarchik, R. Jaimes-Reátegui, and G. Huerta-Cuellar, "Selective monostability in multi-stable systems," *Proceedings of the Royal Society A*, vol. 471, no. 2180, 2015.

[18] A. N. Pisarchik and U. Feudel, "Control of multistability," *Physics Reports*, vol. 540, no. 4, pp. 167–218, 2014.

[19] F. Angulo, *Dynamical analisys of PWM-controlled power electronic converters based on the zero average dynamics (ZAD) strategy [Ph.D. thesis]*, Polytechnic University of Catalonia, Barcelona, Spain, 2004 (Spanish), http://www.tdx.cesca.es/TDX-0727104-095928/.

[20] A. El Aroudi, F. Angulo, G. Olivar, B. G. M. Robert, and M. Feki, "Stabilizing a two-cell DC-DC buck converter by fixed point induced control," *International Journal of Bifurcation and Chaos*, vol. 19, no. 6, pp. 2043–2057, 2009.

[21] F. E. Hoyos, A. Rincón, J. A. Taborda, N. Toro, and F. Angulo, "Adaptive quasi-sliding mode control for permanent magnet DC motor," *Mathematical Problems in Engineering*, vol. 2013, Article ID 693685, 12 pages, 2013.

[22] L. Gardini, R. H. Abraham, D. Fournier Prunaret, and R. J. Record, "A double logistic map," *International Journal of Bifurcation and Chaos*, vol. 4, no. 1, pp. 145–176, 1994.

[23] A. N. Pisarchik, "Controlling the multistability of nonlinear systems with coexisting attractors," *Physical Review E*, vol. 64, no. 4, Article ID 046203, 2001.

[24] E. N. Egorov and A. A. Koronovskii, "Dynamical control in multistable systems," *Technical Physics Letters*, vol. 30, no. 3, pp. 186–189, 2004.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 198380, 8 pages
http://dx.doi.org/10.1155/2015/198380

Research Article

Robust Nonlinear H^∞ Control Design via Stable Manifold Method

Yoshiki Abe,[1] Gou Nishida,[2] Noboru Sakamoto,[3] and Yutaka Yamamoto[1]

[1] *Department of Applied Analysis and Complex Dynamical Systems, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan*
[2] *Department of Mechanical and Environmental Informatics, Graduate School of Information Science and Engineering, Tokyo Institute of Technology, 2-12-1 W8-1, O-okayama, Meguro-ku, Tokyo 152-8552, Japan*
[3] *Department of Mechatronics, Faculty of Science and Engineering, Nanzan University, 18 Yamazato-cho, Shyowa-ku, Nagoya 466-8673, Japan*

Correspondence should be addressed to Gou Nishida; g.nishida@ieee.org

Received 8 September 2015; Accepted 5 November 2015

Academic Editor: Wenguang Yu

This paper proposes a systematic numerical method for designing robust nonlinear H^∞ controllers without a priori lower-dimensional approximation with respect to solutions of the Hamilton-Jacobi equations. The method ensures the solutions are globally calculated with arbitrary accuracy in terms of the stable manifold method that is a solver of Hamilton-Jacobi equations in nonlinear optimal control problems. In this realization, the existence of stabilizing solutions of the Hamilton-Jacobi equations can be derived from some properties of the linearized system and the equivalent Hamiltonian system that is obtained from a transformation of the Hamilton-Jacobi equation. A numerical example is shown to validate the design method.

1. Introduction

Robust controls have been extensively studied to suppress the effects of disturbances or noises on performances of controllers. In particular, the appearance of robust H^∞ control [1] caused the paradigm shift in control theory. The linear H^∞ control has been extended to deal with nonlinear systems [2–4]. The nonlinear H^∞ control design can be described as a problem of solving Hamilton-Jacobi-Isaac equations. However, it is difficult to directly solve Hamilton-Jacobi-Isaac equations as against Riccati equations in the linear case that many practical solving methods have been elaborated. According to the latest reference book [4], there is no systematic numerical approach for solving the Hamilton-Jacobi-Isaac equations at present. Although a lot of efforts have been made [5–13], all the contributions are still valid in a local region around the equilibrium on which low-dimensional approximations of the solutions are valid. Some possible approaches that may yield exact and global solutions are also reviewed in [4].

On the other hand, an effective numerical solver for Hamilton-Jacobi equations in nonlinear optimal control problems that is called the stable manifold method was recently presented [14]. The method has been applied to various control problems [15]. However, their results are basically on pure optimal controls, and robust control designs have not been sufficiently studied in the framework. Optimal controllers without careful thought on robustness might cause instability in systems with disturbances. Thus, the development of robust controls is quite important in nonlinear control design using the stable manifold method.

This paper clarifies the way of implementing robust nonlinear H^∞ control design to the stable manifold method [14]. We believe that our result is the premier report of realizing the nonlinear H^∞ control without a priori lower-dimensional approximation with respect to solutions of the Hamilton-Jacobi equations. The conventional approximate methods based on the Taylor expansion for solving the equations have the critical problem that the valid range of the approximation is unextendable [15]. In our approach, the solutions of the equation can be systematically calculated in a global domain with arbitrary accuracy in terms of the stable manifold method. In our method, we transform the Hamilton-Jacobi-Isaac equation to an equivalent Hamiltonian system under

the assumption that there are no cross-product terms in cost functions, and there is no need to restrict the weight on the control to an identity matrix, which is relaxed from the typical simplification on weights. The existence of stabilizing solutions of the Hamilton-Jacobi(-Isaac) equations can be checked by the stabilizability of the linearized system. The numerical scheme of the stable manifold method is based on the separation of the linear part of the Hamiltonian system that is equivalent to the Hamilton-Jacobi(-Isaac) equation from the nonlinear part. The separation can be achieved if a given system is stabilizable, and the transformation for the separation can be systematically given. Hence, we can apply this method to a wide range of nonlinear control systems. The robust performance of the controller can be designed by choosing the design parameter γ that means the upper bound of the worst response, that is, the H^∞ norm of the system defined by $\mathscr{L}^2$-gain.

This paper is organized as follows: Section 2 makes a brief summary of basic definitions of robust nonlinear H^∞ control. Section 3 shows that the nonlinear H^∞ control design can be converted with the stable manifold method. In Section 4, we show the validity of the nonlinear H^∞ controller derived from the stable manifold method by showing a robustness improvement of a controlled vehicle model [16] under disturbances modeled as an artificial effect of side winds and a rough road surface. In this numerical experimentation, we can see that the nonlinear H^∞ controller achieved a higher robust performance than a linear H^∞ controller in the case that a nonlinear optimal regulator fails stabilization under the disturbances.

2. Summary of Robust Nonlinear H^∞ Control

This section makes a brief summary of basic definitions of robust nonlinear H^∞ control.

2.1. Nonlinear H^∞ Control Design. In this paper, we consider the following standard form of control systems as an objective.

Definition 1. Let one consider the following control system defined on a smooth n-dimensional manifold $\mathscr{X} \subseteq \mathbb{R}^n$:

$$\Sigma : \begin{cases} \dot{x} = f(x) + g_1(x)w + g_2(x)u, \\ y = x, \\ z = h(x) + k(x)u, \end{cases} \tag{1}$$

where the vectors $x(t) \in \mathscr{X}$, $u(t) \in \mathscr{U} \subseteq \mathbb{R}^p$, $w(t) \in \mathscr{W}$, $y(t) \in \mathbb{R}^n$, and $z(t) \in \mathbb{R}^q$ denote state variables, control inputs, disturbances, outputs that can be directly measured, and outputs that are controlled, respectively. In (1), one has defined $\mathscr{U}$ and $\mathscr{W}$, respectively, as the set of admissible controls and the set of admissible disturbances, where a function is called admissible if the function is defined on some time interval and it is piecewise continuous. Furthermore, one denotes the initial state at the time t_0 by $x(t_0) = 0$, and the functions $f: \mathscr{X} \to \mathscr{V}(\mathscr{X})$, $g_1: \mathscr{X} \to \mathscr{M}^{n\times r}(\mathscr{X})$, $g_2: \mathscr{X} \to \mathscr{M}^{n\times p}(\mathscr{X})$, $h_1: \mathscr{X} \to \mathbb{R}^s$, and $k: \mathscr{X} \to \mathscr{M}^{p\times m}(\mathscr{X})$ are assumed to be real C^∞-functions of x, where $\mathscr{V}$ is the vector space of all smooth vector fields over $\mathscr{X}$ and $\mathscr{M}^{i\times j}(\mathscr{X})$ is the ring of $(i \times j)$ matrices over $\mathscr{X}$.

To the system Σ, we consider the following conditions for simplification.

Assumption 2. (1) $x = 0$ is a unique equilibrium point of the system Σ in (1) when $u = 0$ and $w = 0$.

(2) $f(0) = 0$, $h(0) = 0$, and $k^\top(x)k(x) > 0$ hold.

(3) There exists a unique solution $x(t, t_0, x_0, u)$ on the time interval $[t_0, \infty] \in \mathbb{R}$ that continuously depends on the initial condition x_0.

In robust nonlinear H^∞ control, the effect of the signal w to the reference output z is evaluated by the following inequality that will be related with an $\mathscr{L}^2$-gain in the next definition.

Definition 3. System (1) is said to have an $\mathscr{L}^2$-gain less than or equal to γ from w to z in $\mathscr{X}$ if

$$\|z(t)\|_2^2 \le \gamma^2 \|w(t)\|_2^2 + \beta(x_0), \quad \forall T > t_0 \tag{2}$$

for any $x_0 \in \mathscr{X}$, a fixed u, and some bounded C^0-function $\beta: \mathscr{X} \to \mathbb{R}$ such that $\beta(0) = 0$, where one has defined the $\mathscr{L}^2$-norm

$$\|\nu\|_2 = \left(\int_{t_0}^{T} \|\nu(t)\|^2 \, dt \right)^{1/2} \tag{3}$$

for any $\nu: [t_0, T] \subset \mathbb{R} \to \mathbb{R}^n$, where $\|\cdot\|$ means the Euclidean norm on $\mathbb{R}^n$; that is, $\|\nu(t)\|^2 = \nu^\top(t)\nu(t)$.

According to Definition 3, the usual H^∞ norm in a frequency domain can be interpreted as the following $\mathscr{L}^2$-gain that is the induced norm from $\mathscr{L}^2$ to $\mathscr{L}^2$ in the time domain.

Definition 4. One defines the following H^∞ norm of the system Σ:

$$\|\Sigma\|_{\mathscr{H}^\infty} = \sup_{w \in \mathscr{L}^2 \cap \mathscr{L}_c^\infty \setminus \{0\}} \frac{\|z\|_2}{\|w\|_2}, \quad x(t_0) = 0, \tag{4}$$

where $w \in \mathscr{L}^2 \cap \mathscr{L}_c^\infty \setminus \{0\}$ means that $w \in \mathscr{L}^2$ satisfies $\sup_t |w(t)| \le c$ for some constant c and $w \ne 0$.

Remark 5. In the linear H^∞ control design, the disturbance is defined as a function in $\mathscr{L}^2$. On the other hand, in the nonlinear H^∞ control design, the class of disturbances is limited as $w \in \mathscr{L}^2 \cap \mathscr{L}_c^\infty \setminus \{0\}$, because an asymptotical stability does not always hold in a global domain.

By using the above definitions, we state the main problem that is treated in this paper.

Definition 6 (nonlinear H^∞ control problem). Let $\gamma > 0$ be a constant that is a design parameter with respect to disturbances. Then, find a control input u satisfying $\|\Sigma\|_{\mathscr{H}^\infty} \le \gamma$ for the system Σ in (1).

We will rephrase the above problem as the following minimax optimization problem.

Definition 7 (H^∞ differential game). Consider the cost function

$$J(u,w) = \inf_u \sup_w \int_{t_0}^{T} \left\{ \|z(t)\|_2^2 - \gamma^2 \|w(t)\|_2^2 \right\} dt, \quad T > t_0. \tag{5}$$

Then, find the input u that minimizes $J(u,w)$ while the disturbance w maximizes $J(u,w)$ under the constraint described by the system Σ in (1). Furthermore, such solutions (u^*, w^*) must shape a saddle-point equilibrium such that

$$J(u^*, w) \le J(u^*, w^*) \le J(u, w^*) \tag{6}$$

for any disturbance w and any input u that can stabilize the system Σ with the disturbance w^*.

Remark 8. The problem in Definition 7 is not the same problem in Definition 6 in a precise sense; that is, the set of solutions of the problem in Definition 7 is included in that of Definition 6. If the system Σ has a $\mathscr{L}^2$-gain, then the evaluation function J in (5) takes a nonpositive value in the first problem. However, solutions of the second problem are not always nonpositive. Thus, we must check the nonpositiveness separately from solving the second problem.

Remark 9. Finding the worst disturbance w^* is not included in the first problem in Definition 6.

2.2. Hamilton-Jacobi-Isaac Equation. Such a two-person zero-sum game as in Definition 7 has a solution if the value function

$$V(x,t) = \inf_u \sup_w \int_t^T \left\{ \|z(\tau)\|_2^2 - \gamma^2 \|w(\tau)\|_2^2 \right\} d\tau \tag{7}$$

is C^1, and V satisfies the dynamic-programming equation

$$-\frac{\partial V}{\partial t} = \inf_u \sup_w \left\{ \frac{\partial V}{\partial x}\dot{x} + \|z(t)\|_2^2 - \gamma^2 \|w(t)\|_2^2 \right\}, \quad V(T,x) = 0. \tag{8}$$

Now, we consider the infinite-time horizon problem under the conditions $\lim_{T\to\infty} J(u,w)$ remains bounded and the $\mathscr{L}^2$-gain of the system remains finite; that is, we find a time-independent positive-semidefinite function $V: \mathscr{X} \to \mathbb{R}$ satisfying the relation

$$\begin{aligned} &H(x,p,u,w) \\ &\quad = p^\top \{ f(x) + g_1(x) w + g_2(x) u \} + z^\top(t) z(t) \\ &\qquad - \gamma^2 w^\top(t) w(t), \\ &\inf_u \sup_w H(x,p,u,w) = 0, \quad V(0) = 0 \end{aligned} \tag{9}$$

that is called *the Hamilton-Jacobi-Isaac equation*, where we have defined $p = (\partial V/\partial x)^\top$. From the stationary conditions $\partial H/\partial u = 0$ and $\partial H/\partial w = 0$, we obtain the following explicit forms of optimal solutions:

$$\begin{aligned} u^* &= -\frac{1}{2} K^{-1}(x) \Xi(x,p), \\ w^* &= \frac{1}{2\gamma^2} g_1^\top(x) p, \end{aligned} \tag{10}$$

where we have defined $K(x) = k^\top(x)k(x) > 0$ and $\Xi(x,p) = g_2^\top(x)p + 2k^\top(x)h(x)$. Then, the Hamilton-Jacobi-Isaac equation can be written as

$$\begin{aligned} H(x,p,u,w) = {} & p^\top f(x) + \frac{1}{4\gamma^2} p^\top g_1(x) g_1^\top(x) p \\ & - \frac{1}{4} \Xi^\top(x,p) K^{-1}(x) \Xi(x,p) \\ & + h^\top(x) h(x) = 0. \end{aligned} \tag{11}$$

Indeed, the Hamiltonian H in (11) can be transformed into

$$\begin{aligned} H(x,p,u,w) = {} & H(x,p,u^*,w^*) \\ & - \gamma^2 (w - w^*)^\top (w - w^*) \\ & + (u - u^*)^\top K(x) (u - u^*) \end{aligned} \tag{12}$$

that means the solutions u^* and w^* determine the saddle point of the Hamiltonian.

From the above preliminaries, we can obtain the following fact.

Theorem 10 (see [17]). *If there exists a function $V(x) \in C^1$ such that $H(x,p) = 0$, $p = (\partial V/\partial x)^\top$, $V(x) \ge 0$, and $V(0) = 0$ for the Hamiltonian $H(x,p)$ in (11), then u^* and w^* in (10) are the solution of the system Σ in (1), and the $\mathscr{L}^2$-gain of the system Σ is less than or equal to γ.*

3. Nonlinear H^∞ Control Design Using Stable Manifold Method

This section derives the way of converting the nonlinear H^∞ control design with the stable manifold method from the viewpoint of the Hamiltonian representation of Hamilton-Jacobi-Isaac equations.

3.1. Stabilizing Solution of Hamilton-Jacobi Equations. Before explaining the implementation of the linear and nonlinear H^∞ control designs to the stable manifold method, we make a brief summary of basic results on the solvability of Hamilton-Jacobi equations.

Assumption 11. We assume that $h(x)^\top k(x) = 0$ for all $x \in \mathscr{X}$. For example, in this case, we can write $z = h(x) + k(x)u$ with $h(x) = [h_1(x), 0]^\top$, and $k(x) = [0, k_2(x)]^\top$.

Remark 12. In the typical settings [4, 17], the condition $K(x) = k^{\top}(x)k(x) = I$ that means the unity weighting on the control is introduced to reduce (11) to be a simple quadratic form with respect to g_2 without the weight K^{-1} in addition to the condition in Assumption 11. However, in control designs using the stable manifold method, such a simplification is not necessary.

Proposition 13 (see [17]). *Let one consider the following approximations:*

$$\begin{aligned} f(x) &= Ax + \mathcal{O}\left(|x|^2\right), \\ \overline{R}(x) &= R + \mathcal{O}(|x|), \\ \overline{Q}(x) &= \frac{1}{2}x^{\top}Qx + \mathcal{O}\left(|x|^3\right) \end{aligned} \tag{13}$$

in (11), where $\overline{R}(x) := R_2(x) - (1/\gamma^2)g_1(x)g_1^{\top}(x)$, $\overline{Q}(x) := h^{\top}(x)h(x)$, and A, R, and Q are constant matrixes. If $V(x)$ is assumed to be a quadratic form of symmetric matrix P, the Hamilton-Jacobi-Isaac equation can be reduced to the Riccati equation:

$$PA + A^{\top}P - PRP + Q = 0. \tag{14}$$

Definition 14. A solution of the Riccati equation (14) is called a *stabilizing solution* if $A - RP$ is a stable matrix.

Theorem 15 (see [17]). *Consider the Hamilton-Jacobi equation $H(x,p) = p^{\top}f(x) - (1/2)p^{\top}R_2(x)p + q(x) = 0$ in nonlinear optimal control problems, where $p = (\partial V/\partial x)^{\top}$ and $R_2(x) = g_2(x)K^{-1}(x)g_2^{\top}(x)$. If the Riccati equation derived from the Hamilton-Jacobi equation has a stabilizing solution, then there exists a stabilizing solution $V(x)$ of the Hamilton-Jacobi equation such that $f(x) - R_2(x)p(x)$ is asymptotically stable.*

3.2. Calculation of Stabilizing Solutions via Stable Manifold Method. In this section, we clarify H^{∞} control design procedures in stable manifold method. The objective of the stable manifold method [14] is to calculate a stable manifold of stabilizing solutions of the Hamilton-Jacobi equation by using the following iterative numerical scheme:

(1) Transform the equivalent Hamiltonian system of the Hamilton-Jacobi equation as

$$\begin{bmatrix} \dot{x}' \\ \dot{p}' \end{bmatrix} = \begin{bmatrix} F & 0 \\ 0 & -F^{T} \end{bmatrix} \begin{bmatrix} x' \\ p' \end{bmatrix} + \begin{bmatrix} n_s\left(t, x', p'\right) \\ n_u\left(t, x', p'\right) \end{bmatrix} \tag{15}$$

by the coordinate transformation

$$\begin{bmatrix} x' \\ p' \end{bmatrix} = \begin{bmatrix} I & S \\ P & PS + I \end{bmatrix}^{-1} \begin{bmatrix} x \\ p \end{bmatrix}, \tag{16}$$

where S is the matrix that is a solution of Lyapunov equation $FS + SF^{T} = F$ and $F = A - RP$.

(2) Calculate sequences $\{x_k'(t,\xi)\}$ and $\{p_k'(t,\xi)\}$ determined by

$$\begin{aligned} x_{k+1}'(t,\xi) &= e^{Ft}\xi \\ &\quad + \int_0^t e^{F(t-s)} n_s\left(s, x_k'(s,\xi), p_k'(s,\xi)\right) ds, \\ p_{k+1}'(t,\xi) &= -\int_t^{\infty} e^{-F^{T}(t-s)} n_u\left(s, x_k'(s,\xi), p_k'(s,\xi)\right) ds \end{aligned} \tag{17}$$

for a certain parameter $\xi \in \mathbb{R}^n$, where $x_0'(t,\xi) = e^{Ft}\xi$ and $p_0'(t,\xi) = 0$.

(3) By iteratively applying (17), extend a solution along an initial vector ξ in a plain surface spanned by P under the condition that the Hamiltonian of the right side of (11) is sufficiently close to zero.

(4) If a solution passes through a desired initial state of control systems, then the iteration is finished. If not, back to procedure (2) and try with other ξ.

We can actually transform the Hamilton-Jacobi-Isaac equation (11) into the following Hamiltonian system.

Lemma 16. *Under Assumption 11, (11) can be transformed into the equivalent Hamiltonian system:*

$$\begin{aligned} \dot{x} &= \frac{\partial H^{\top}}{\partial p} = f(x) - \frac{1}{2}\left(R_2(x) - \frac{1}{\gamma^2}g_1(x)g_1^{\top}(x)\right)p, \\ \dot{p} &= -\frac{\partial H^{\top}}{\partial x} \\ &= -\frac{\partial f}{\partial x}^{\top}(x)p - \frac{1}{2\gamma^2}p^{\top}\frac{\partial g_1}{\partial x}(x)g_1^{\top}(x)p \\ &\quad + \frac{1}{4}\left(\frac{\partial}{\partial x}p^{\top}R_2(x)p\right)^{\top} - 2\frac{\partial h^{\top}}{\partial x}(x)h(x), \end{aligned} \tag{18}$$

where we have defined $R_2(x) = g_2(x)K^{-1}(x)g_2^{\top}(x)$.

From the facts discussed in the previous section, we can obtain the condition for the applicability of the stable manifold method.

Theorem 17. *Let us consider a nonlinear H^{∞} control problem for system (1). For the Riccati equation (14) corresponding to the Hamilton-Jacobi-Isaac equation (11) of the problem under the approximation (13), if the Hamiltonian matrix*

$$H = \begin{bmatrix} A & -R \\ -Q & -A^{\top} \end{bmatrix} \tag{19}$$

does not have eigenvalues on the imaginary axis and (A, R) is stabilizable, then we can calculate the stabilizing solution of the Hamilton-Jacobi-Isaac equation by using the stable manifold method.

Proof. A stable manifold can be described by $p = (\partial V/\partial x)^{\top}$, and such a function $V(x)$ exists if the Hamiltonian matrix of

the Riccati equation corresponding to the Hamilton-Jacobi-Isaac equation does not have eigenvalues on the imaginary axis [17]. Indeed, this fact is used in the proof of Theorem 15. If the linearized system (A, R) is stabilizable and $R \geq 0$ or $R \leq 0$, there exists the stabilizing solution of the Riccati equation [17]. Now, we assumed that $K > 0$; then $K^{-1} > 0$; that is, $R \geq 0$ or $R \leq 0$, because R is the linear part of $\overline{R}(x) := R_2(x) - (1/\gamma^2) g_1(x) g_1^\top(x)$, where $R_2(x) = g_2(x) K^{-1}(x) g_2^\top(x)$. Hence, there also exists a stabilizing solution $V(x)$ of the Hamilton-Jacobi-Isaac equation according to Theorem 15. Consequently, in such a case, we can directly find p derived from the stabilizing solution $V(x)$ by the stable manifold method. The Hamiltonian system representation in (15) can be given by the system in Lemma 16 and the linearization in (13). □

4. Numerical Example

We will check the validity of the nonlinear H^∞ control design via the stable manifold method by showing a robustness improvement of a controlled vehicle model [16].

4.1. Control Model. We assume that the left side and right side wheels of a vehicle have the same property, and the vehicle should be stabilized to some direction under a constant speed. Then, the equivalent 2-wheel model with respect to yawing without rolling and pitching motions is given as follows:

$$\begin{aligned} x &= [\beta \;\; r \;\; \theta \;\; \delta \;\; Y]^\top, \\ w &= [w_1 \;\; w_2]^\top, \end{aligned} \tag{20}$$

$$f(x) = \left[-\frac{\sin\beta}{mV_0} F_x + \frac{\cos\beta}{mV_0} F_y - r \;\; \frac{2l_f}{I} C_f \cos\delta - \frac{2l_r}{I} C_r \;\; r \;\; 0 \;\; V_0 \sin(\beta + \theta) \right]^\top, \tag{21}$$

$$\begin{aligned} g_1(x) &= \begin{bmatrix} \dfrac{\cos(\beta+\theta)}{(mV_0)} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}^\top, \\ g_2(x) &= [0 \;\; 0 \;\; 0 \;\; 1 \;\; 0]^\top, \end{aligned} \tag{22}$$

where the control input u is the steering angle speed, the state vector x consists of the slip angle β at center of gravity (COG), the yaw rate r, the direction θ, the steering angle δ, and the lateral position Y of the vehicle, and note that the vertical position is ignored under the assumption of motions around a constant speed. Furthermore, the translational forces F_x and F_y and the cornering force of each wheel Y_i are written as follows:

$$\begin{aligned} F_x &= 2Y_f \sin(\beta_f + \delta) + 2Y_r \sin\beta_r, \\ F_y &= 2Y_f \cos(\beta_f + \delta) + 2Y_r \cos\beta_r, \\ Y_i &= C_i \cos\beta_i \end{aligned} \tag{23}$$

for $i = \{f, r\}$ that means the front and the rear wheels, respectively, where β_i is the slip angle of wheels, C_i is the lateral force of wheels, and C_i and β are related by

$$C_i = \mu N_i \sin\left[a \tan^{-1}\left\{b\beta_i - c\left(b\beta_i - \tan^{-1}(b\beta_i)\right)\right\}\right], \tag{24}$$

where $a = 1.23$, $b = 3.25$, and $c = -6.00$ are experimental parameters, $\mu = 0.2$ is a friction constant between road surface and tire, and $N_f = 5.48$ and $N_r = 4.21$ are vertical loads of each wheel. In (22), the following physical parameters are used: the constant speed $V_0 = 17.7$, the mass $m = 990$, the moment of inertia $I = 683$, the distance from front axle to COG $l_f = 1.0$, and the distance from rear axle to COG $l_r = 1.3$.

4.2. Disturbance Models. We applied the following disturbance to the model during simulations:

$$\begin{aligned} w_1 &= \begin{cases} 0.47 & (0 \leq t \leq 0.5) \\ 0 & (\text{otherwise}), \end{cases} \\ w_2 &= \begin{cases} \dfrac{1}{40} \displaystyle\sum_{k=60}^{100} \sin(kt) & (2 \leq t \leq 5) \\ 0 & (\text{otherwise}) \end{cases} \end{aligned} \tag{25}$$

that mean artificial effects of side winds and rough road surfaces (see Figure 1). However, the particular information of these disturbances defined by the above relations is not used in the design of H^∞ controllers, but we only determine the upper bound of the disturbance, that is, γ as a design parameter.

4.3. Additional Calculation. According to Theorem 15, we must check an obtained function $V(x)$ is nonnegative. Because the stable manifold method gives the pair of the variables (x, p) as a solution, we must calculate V from p obtained from the simulation by

$$V(x(t)) = \int_0^\infty p^\top \dot{x}\, dt. \tag{26}$$

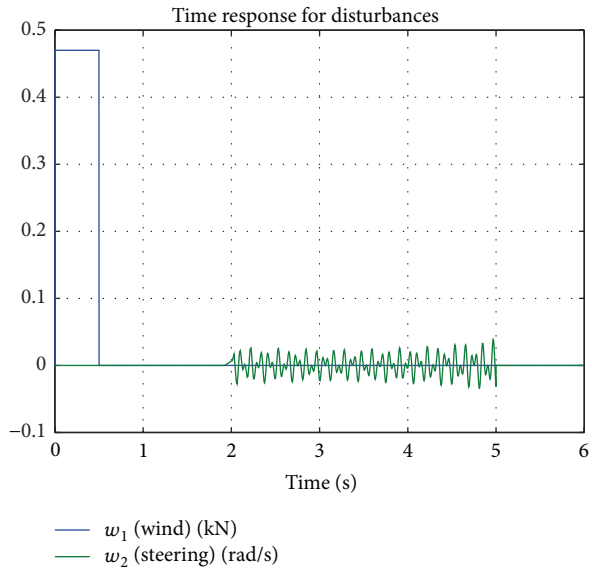


Figure 1: Disturbances.

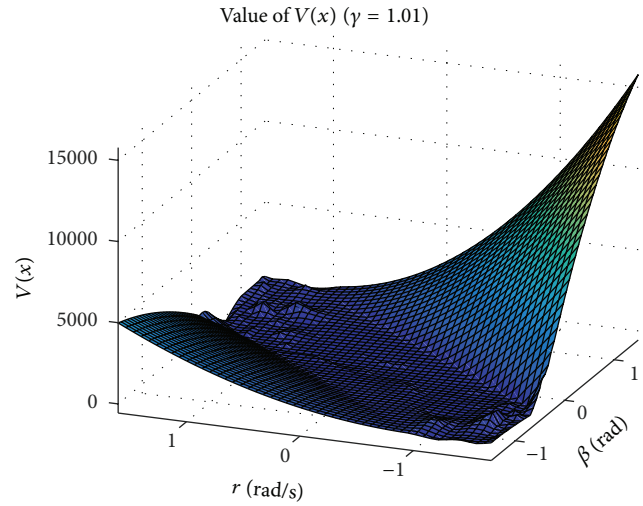


Figure 3: Stabilizing solution (view 2).

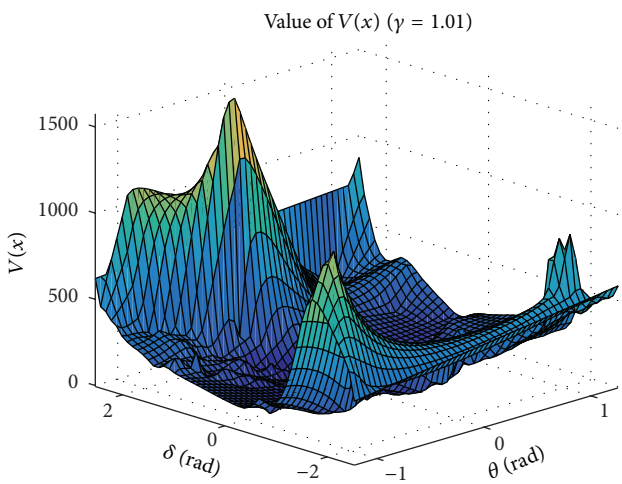


Figure 2: Stabilizing solution (view 1).

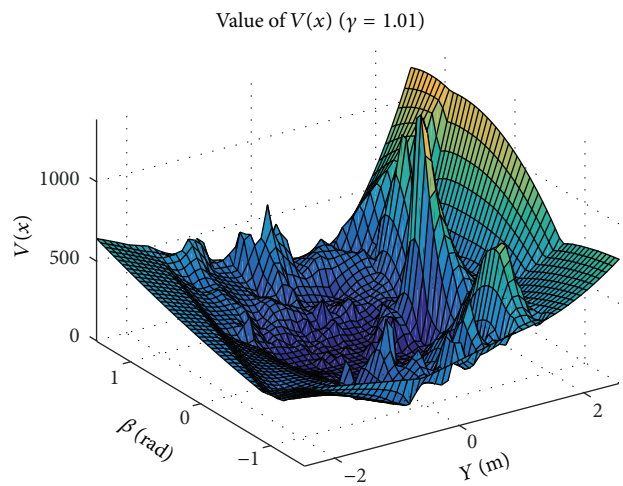


Figure 4: Stabilizing solution (view 3).

4.4. Numerical Results. We carried out the simulation using the stable manifold method for the model with $\gamma = 1.01$. Figures 2–4 show the three projections of the stabilizing solution $V(x)$ calculated by (26), where please note that $V(x)$ is defined on the fifth-dimensional space of x. We can see that $V(x)$ is nonnegative.

Figures 5–10 show the time plots of the state variables controlled by the linear and nonlinear H^∞ controllers. The convergence performance of the time responses was improved by the nonlinear H^∞ controller. Then, the values of the objective functions of the linear and nonlinear controls were $J = -0.0809$ and $J = -0.0818$, respectively. Indeed, from these figures, we can see that the amplitude of the control input generated by the nonlinear H^∞ controller is smaller than that of the linear H^∞ controller.

On the other hand, we did the simulation for the same model controlled by a nonlinear optimal regulator that does

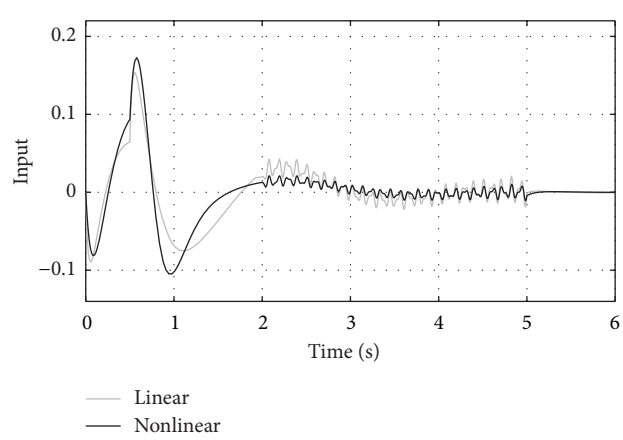


Figure 5: Time response of inputs.

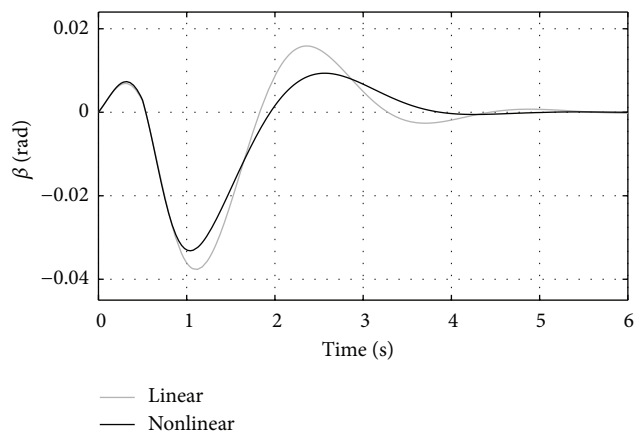


Figure 6: Time response of slip angles.

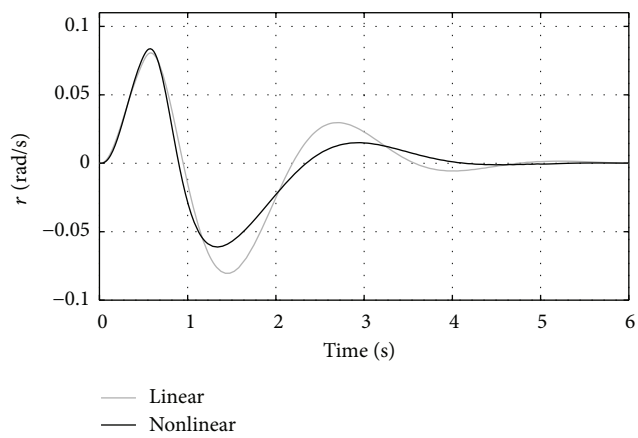


Figure 7: Time response of yaw rates.

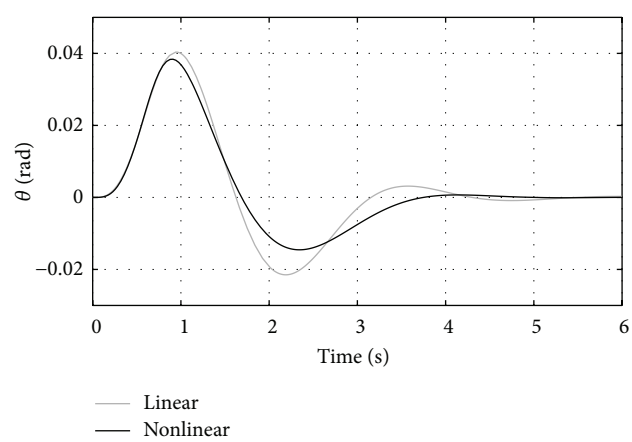


Figure 8: Time response of directions.

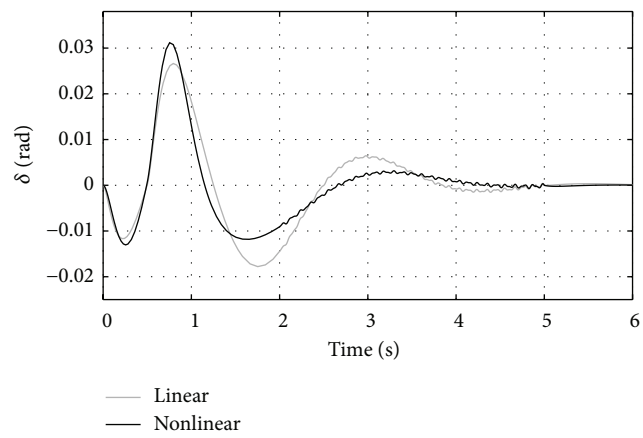


Figure 9: Time response of steering angles.

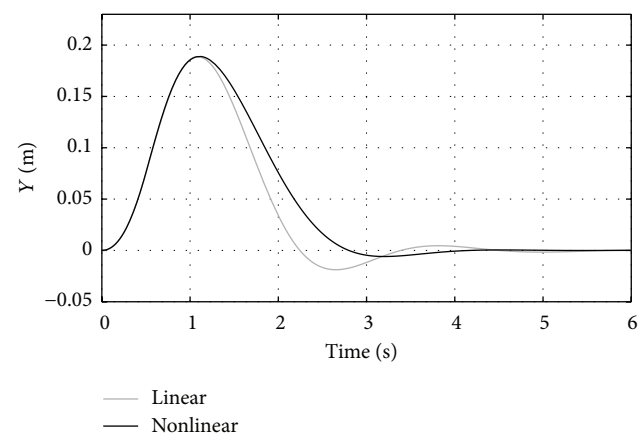


Figure 10: Time response of lateral positions.

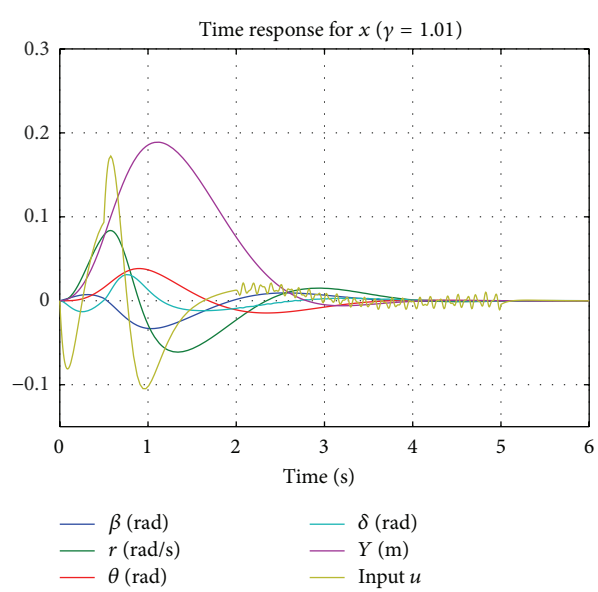


Figure 11: Time responses of nonlinear H^{∞} control.

not have any guarantee with respect to robustness. Figure 12 shows the time plot of the state variables by the nonlinear optimal regulator with the unit weight 1 to control inputs. However, the trajectory diverged; that is, the system became unstable. Note that Figure 11 is the plot in which the time

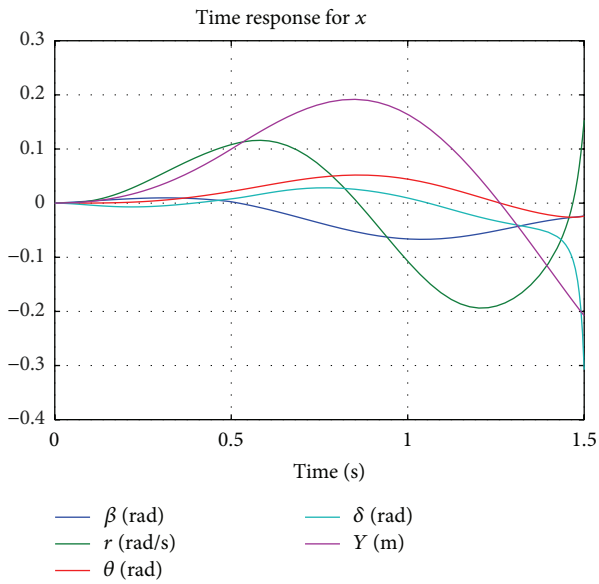


FIGURE 12: Time responses of nonlinear optimal regulator.

responses of the nonlinear H^∞ controller in Figures 5–10 are collected.

Consequently, we can confirm that the nonlinear H^∞ controller achieved a higher robust performance in this case.

5. Conclusion

We proposed the way of integrating robust nonlinear H^∞ control design to the stable manifold method. Furthermore, the numerical experimentation was shown for checking the validity of the robust control for the vehicle model with disturbances. The stable manifold method does not require the information of analytical solutions, and we only have to prepare the description of nonlinear systems. Hence, we expect this method to be applied to a lot of control objects that could not be considered due to theoretical difficulties before.

At present, we realized only the full-state feedback case. The output feedback case can be considered as a challenging future work.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by JSPS Grants-in-Aid for Scientific Research (C) no. 26420415 and JSPS Grants-in-Aid for Challenging Exploratory Research no. 26630197. N. Sakamoto was supported by Nanzan University Pache Research Subsidy I-A-2 for the 2015 academic year.

References

[1] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*, Prentice Hall, 1995.

[2] J. W. Helton and M. R. James, *Extending H-Infinity Control to Nonlinear Systems: Control of Nonlinear Systems to Achieve Performance Objectives*, SIAM Books, 1987.

[3] A. J. van der Schaft, *L_2-Gain and Passivity Techniques in Nonlinear Control*, Springer, London, UK, 2nd edition, 2000.

[4] M. D. S. Aliyu, *Nonlinear H_∞-Control, Hamiltonian Systems and Hamilton-Jacobi Equations*, CRC Press, 2011.

[5] D. L. Lukes, "Optimal regulation of nonlinear dynamical systems," *SIAM Journal on Control and Optimization*, vol. 7, pp. 75–100, 1969.

[6] S. T. Glad, "Robustness of nonlinear state-feedback: a survey," *Automatica*, vol. 23, no. 4, pp. 425–435, 1987.

[7] A. J. van der Schaft, "L_2-gain analysis of nonlinear systems and nonlinear state feedback H_∞-control," *IEEE Transactions on Automatic Control*, vol. 37, no. 6, pp. 770–784, 1992.

[8] A. Isidori and W. Kang, "H_∞ control via measurement feedback for general nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 40, no. 3, pp. 466–472, 1995.

[9] R. W. Beard, G. N. Saridis, and J. T. Wen, "Galerkin approximations of the generalized Hamilton-Jacobi-Bellman equation," *Automatica*, vol. 33, no. 12, pp. 2159–2177, 1997.

[10] B.-S. Chen and Y.-C. Chang, "Nonlinear mixed H_2/H_∞ control for robust tracking design of robotic systems," *International Journal of Control*, vol. 67, no. 6, pp. 837–857, 1997.

[11] H. Guillard, S. Monaco, and D. Normand-Cyrot, "Approximated solutions to nonlinear discrete-time H_∞-control," *IEEE Transactions on Automatic Control*, vol. 40, no. 12, pp. 2143–2148, 1995.

[12] P. Tsiotras, M. Corless, and M. A. Rotea, "An L_2 disturbance-attenuation solution to the nonlinear benchmark problem," *International Journal of Robust and Nonlinear Control*, vol. 8, no. 4-5, pp. 311–330, 1998.

[13] M. J. Yazdanpanah, K. Khorasani, and R. V. Patel, "Uncertainty compensation for a flexible-link manipulator using nonlinear H_∞ control," *International Journal of Control*, vol. 69, no. 6, pp. 753–771, 1998.

[14] N. Sakamoto and A. J. van der Schaft, "Analytical approximation methods for the stabilizing solution of the Hamilton-Jacobi equation," *IEEE Transactions on Automatic Control*, vol. 53, no. 10, pp. 2335–2350, 2008.

[15] N. Sakamoto, "Case studies on the application of the stable manifold approach for nonlinear optimal control design," *Automatica*, vol. 49, no. 2, pp. 568–576, 2013.

[16] R. Rajamani, *Vehicle Dynamics and Control*, Mechanical Engineering Series, Springer, 2nd edition, 2012.

[17] A. J. van der Schaft, "On a state space approach to nonlinear H_∞ control," *Systems & Control Letters*, vol. 16, no. 1, pp. 1–8, 1991.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 645168, 8 pages
http://dx.doi.org/10.1155/2015/645168

Research Article

Adaptive MIMO Supervisory Control Design Using Modeling Error

Zhi-Ren Tsai[1,2] and Yau-Zen Chang[3]

[1]*Department of Computer Science & Information Engineering, Asia University, Wufeng District, Taichung 41354, Taiwan*
[2]*Department of Medical Research, China Medical University Hospital, China Medical University, Taichung 40402, Taiwan*
[3]*Department of Mechanical Engineering, Chang Gung University, Taoyuan 33302, Taiwan*

Correspondence should be addressed to Yau-Zen Chang; zen@mail.cgu.edu.tw

Received 14 July 2015; Revised 14 September 2015; Accepted 15 September 2015

Academic Editor: Rongwei Guo

This paper proposes an adaptive control scheme for nonlinear systems with significant nonminimum phase dynamics. The scheme is composed of an inner-level adaptive fuzzy PD control law and an outer-level supervisory control law. Importantly, the inner-level controller of the two-level scheme is designed based on a fuzzy model, which takes nonminimum phase phenomenon and modeling error explicitly into account. The scheme is both much simpler in design and more applicable to general nonlinear systems when compared with most existing nonlinear controllers. Effectiveness of the proposed control strategy is demonstrated by numerical simulation of the control of a five-degree-of-freedom aircraft system in the face of bursting disturbances.

1. Introduction

Many critical dynamic systems, such as aircraft, are nonminimum phase, MIMO, and highly nonlinear, which undergo significant disturbances and parameter variation during operation. To control these systems, robust control [1], optimal tuning of fuzzy controllers with output sensitivity function [2], adaptive control [3–5], and feedback linearization with discrete sliding-mode control [6] have attracted much attention from both academic and industrial communities due to their robustness to uncertainties. Recently, many interests have been focused on applying these techniques to flight control systems, such as [7, 8]. However, for systems with significant nonminimum phase phenomenon, direct application of these approaches tends to introduce unstable zero dynamics.

For instance, in [9], the nonminimum phase plants are approximated by minimum phase models. The research [10] applied the output regulation theory to solve the output tracking problem, but a set of partial differential equations must be solved. The control scheme of [11] is based on decomposing the aircraft dynamics into a minimum phase part and a nonminimum phase part. Inversion is used on the minimum phase part to obtain asymptotic tracking, while a robust linear control approach is used to stabilize the nonminimum phase part, which is linearized at equilibrium. As this strategy is based on local linearization of the nonminimum phase part, the result can only apply to simplified models.

By estimating parameters online, adaptive control can adapt to a controlled system with varying or unknown parameters. Nevertheless, in spite of the prosperous literature of adaptive control, practical application of these control strategies on MIMO systems has been restricted by the lack of assurance in closed-loop stability. Among them, the adaptive neural controller of [3] is too complex to implement, while the adaptive fuzzy terminal sliding-mode controller of [4, 5] is applicable only to robotic manipulators.

The proposed adaptive control scheme is inspired by [12], which was developed for SISO nonlinear systems based on the feedback linearization technique, with the distinction that the scheme is extended to nonminimum phase MIMO control systems.

The scheme is composed of an inner-level tracking control law and an outer-level supervisory control law.

The design procedure is hence divided into two parts. First, an adaptive fuzzy-model-based PD control scheme is designed at the inner level to achieve robust output tracking. Special care is taken for the nonminimum phase fuzzy subsets in the control law by restricting parameter magnitudes in the singular-value decomposition operation. Next, a supervisory controller is employed at the outer level to minimize both the approximation error between the fuzzy model and the plant and the effects of external disturbance. Effectiveness of the adaptive control scheme is demonstrated by simulation results of the fight control of a complete 5-DOF aircraft model.

2. Problem Formulation

System dynamics of the plant are firstly represented in a general MIMO state-space representation as

$$\begin{aligned} \dot{x} &= F(x) + G(x) \cdot u + w_0, \\ y &= H \cdot x, \end{aligned} \tag{1}$$

where $x \in \mathfrak{R}^{n\times 1}$ is the state vector, $u \in \mathfrak{R}^{m\times 1}$ is the control vector, $w_0 \in \mathfrak{R}^{n\times 1}$ is the disturbance vector, $y \in \mathfrak{R}^{N\times 1}$ is the output vector, and F, G are corresponding nonlinear matrices in state vectors with H being a constant matrix, all of compatible dimensions.

Equation (1) can be further represented in output vector y as

$$\begin{aligned} \dot{y} &= H \cdot F(x) + H \cdot G(x) \cdot u + H \cdot w_0 \\ &= f(x) + g(x) \cdot u + w \\ &= h_f \cdot A_f + \sum_{i=1}^{L} h_i \cdot B_i \cdot u + \left[f(x) - h_f \cdot A_f \right] \\ &\quad + \left[g(x) - \sum_{i=1}^{L} h_i \cdot B_i \right] \cdot u + w \\ &= h_f \cdot A_f + \sum_{i=1}^{L} h_i \cdot B_i \cdot u + \bar{e}_{\text{mod}} \\ &\triangleq h_f \cdot A_f + B_c \cdot u + \bar{e}_{\text{mod}}, \end{aligned} \tag{2}$$

where

$$\begin{aligned} B_c &\triangleq \sum_{i=1}^{L} h_i \cdot B_i, \\ h_f &= \begin{bmatrix} h_{f1} & 0 & 0 & 0 \\ 0 & h_{f2} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & h_{fn} \end{bmatrix}, \\ A_f &= \left[A_{f1} \;\; A_{f2} \;\; \cdots \;\; A_{fn} \right]^T, \\ f_1 &= \sum_{i=1}^{p} h_{1i} \cdot A_{1i} = h_{f1} \cdot A_{f1}, \\ f_2 &= \sum_{i=1}^{p} h_{2i} \cdot A_{2i} = h_{f2} \cdot A_{f2}, \ldots, \\ f_n &= \sum_{i=1}^{p} h_{ni} \cdot A_{ni} = h_{fn} \cdot A_{fn}, \\ \bar{e}_{\text{mod}} &= \left[f(x) - h_f \cdot A_f \right] + \left[g(x) - \sum_{i=1}^{L} h_i \cdot B_i \right] \cdot u \\ &\quad + w, \end{aligned} \tag{3}$$

and the external disturbance $w = H \cdot w_0$.

In the last representation, it is assumed that $g(x)$ is bounded and is away from singularity in a compact set. Furthermore, $f(x)$ and $g(x)$ are identified in fuzzy form as $h_f(y) \cdot A_f(t)$ and $\sum_{j=1}^{L} h_j(y) \cdot B_j$, respectively, where the fuzzy logic systems are universal approximations which can uniformly approximate nonlinear continuous functions to arbitrary accuracy [13–15].

3. Controller Design for the Nonminimum Phase Dynamics

Firstly, the tracking error is defined as

$$e(t) = -\bar{e}, \tag{4}$$

where $\bar{e} = (y_r - y)$ (y_r is reference input); we have that

$$\begin{aligned} \dot{e}(t) &= h_f \cdot A_f + \sum_{i=1}^{L} h_i \cdot B_i \cdot u + \bar{e}_{\text{mod}} - \dot{y}_r \\ &= h_f \cdot A_f + B_c \cdot u + \bar{e}_{\text{mod}} - \dot{y}_r, \end{aligned} \tag{5}$$

where u is a combination of two signals [16]:

$$u = u_F + u_S, \tag{6}$$

with

$$\begin{aligned} u_F &= \left(1 - I^*\right) \cdot B_c^{-1} \\ &\quad \cdot \left\{ -h_f(y) \cdot A_f(t) + \dot{y}_r + K_P \cdot e(t) + K_D \cdot \dot{e}(t) \right\}, \end{aligned} \tag{7}$$

$$\begin{aligned} u_S &= -I^* \cdot \operatorname{sgn}\left(B_c \cdot P \cdot e(t)\right) \\ &\quad \cdot \left\{ \left| B_c^{-1} \cdot \left[h_f(y) \cdot A_f(t) - \dot{y}_r \right] \right| + e_U \right\}. \end{aligned} \tag{8}$$

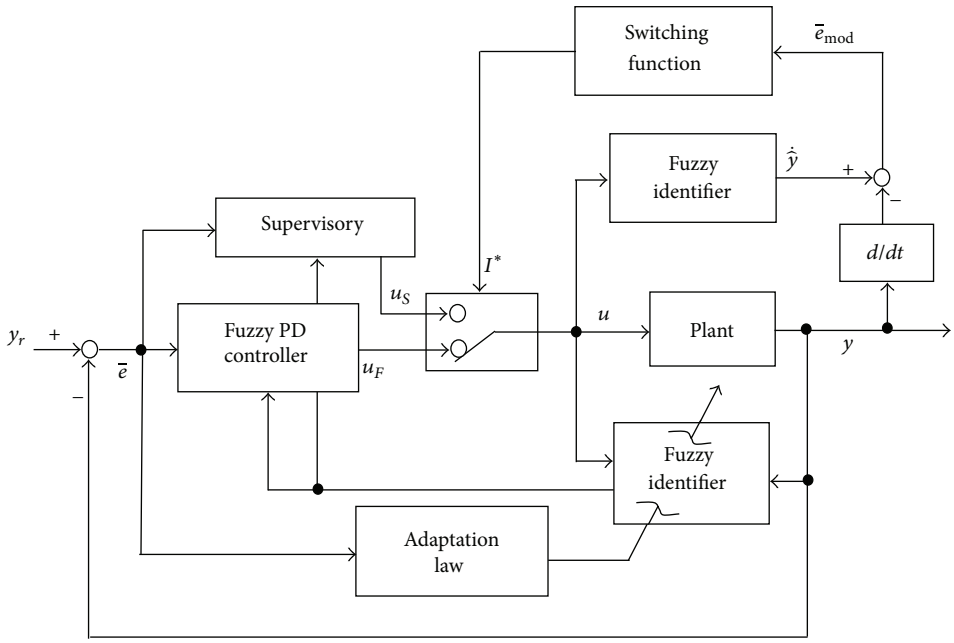


Figure 1: The proposed two-level switching control scheme.

In (7), the proportional gain of the inner fuzzy control law is designed as

$$K_P = \alpha - R^{-1} \cdot P. \tag{9}$$

The switching variable in both (7) and (8) is defined as

$$\begin{aligned} I^* &= 0, \quad \text{if } \left\|\bar{e}_{\text{mod}}\right\| \le \bar{e}_U, \\ I^* &= 1, \quad \text{otherwise}, \end{aligned} \tag{10}$$

with

$$\begin{aligned} &\left|B_c^{-1} \cdot \bar{e}_{\text{mod}}\right| \le e_U, \\ &\alpha = \begin{bmatrix} -\alpha_1 & 0 & 0 & 0 \\ 0 & -\alpha_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & -\alpha_n \end{bmatrix}, \\ &P > 0, \\ &R > 0. \end{aligned} \tag{11}$$

A complete control scheme of the two-level architecture is shown in Figure 1.

To avoid encountering singularity of the control law, the singular-value decomposition of the matrix B_c is introduced as follows:

$$B_c = U \cdot S \cdot V^T, \tag{12}$$

where

$$S = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \sigma_n \end{bmatrix} \tag{13}$$

and σ_i is replaced by ε if $\sigma_i \le \varepsilon$, where ε is a small value.

Substituting the adaptive fuzzy PD controller u_F (7) into (5), we have

$$\dot{e}(t) = K_P \cdot e(t) + h_f(y) \cdot \widetilde{A}_f(t) + e_{\text{mod}}, \tag{14}$$

after algebra manipulations, where

$$\widetilde{A}_f(t) = A_f^* - A_f(t), \tag{15}$$

and the modeling error

$$\begin{aligned} e_{\text{mod}} = {} & \left[f(x) - h_f(y) \cdot A_f^*\right] \\ & + \left[g(x) - \sum_{i=1}^{L} h_i(y) \cdot B_i\right] \cdot u + w + K_D \\ & \cdot \dot{e}(t). \end{aligned} \tag{16}$$

In the following derivation, we need the following condition to be satisfied [17–20]:

$$J \le e(0)^T \cdot P \cdot e(0) + \gamma_f^{-1} \cdot \operatorname{trace}\left(\widetilde{A}_f(0)^T \cdot \widetilde{A}_f(0)\right), \quad (17)$$

where

$$J = \int_0^{t_f} \left[e(t)^T \cdot \left(Q + P^T \cdot R^{-T} \cdot P\right) \cdot e(t) - \rho^2 \cdot e_{\text{mod}}^T \cdot e_{\text{mod}}\right] \cdot dt, \quad (18)$$

the weighting factor $\gamma_f > 0$, $\rho^2 \cdot I \ge R$, the matrix $Q > 0$, and $\|X\|_F = \sqrt{\operatorname{trace}(X^T \cdot X)}$ is the Frobenius norm of matrix X. Derivation of this condition, (17), is given in the Appendix.

Hence, we have that

$$\begin{aligned} &\dot{\widetilde{A}}_f(t) = -\dot{A}_f = -\gamma_f \cdot h_f(y)^T \cdot P \cdot e(t), \\ &\left[\alpha^T \cdot P + P \cdot \alpha + Q - P^T \cdot R^{-T} \cdot P\right] = -\rho^{-2} \cdot P^T \cdot P. \end{aligned} \quad (19)$$

Furthermore, to guarantee boundedness of A_f, the parameter update laws must be modified as follows:

$$\dot{A}_f = \begin{cases} \gamma_f \cdot h_f(y)^T \cdot P \cdot e(t), & \text{if } \|A_f\| < M_f \text{ or } \left(\|A_f\| = M_f,\ \dot{A}_f^T \cdot A_f \le 0\right), \\ F\left(\gamma_f \cdot h_f(y)^T \cdot P \cdot e(t)\right), & \text{otherwise}, \end{cases} \quad (20)$$

where M_f is a positive design parameter and the projection function $F(\cdot)$ is defined as

$$\begin{aligned} &F\left(\gamma_f \cdot h_f(y)^T \cdot P \cdot e(t)\right) \\ &\quad = \gamma_f \cdot h_f(y)^T \cdot P \cdot e(t) - \gamma_f \\ &\qquad \cdot \frac{A_f \cdot A_f^T \cdot h_f(y)^T \cdot P \cdot e(t)}{\|A_f\|^2}. \end{aligned} \quad (21)$$

Next, the supervisor control law of (8) is designed by the following Lyapunov candidate:

$$V = e(t)^T \cdot P \cdot e(t). \quad (22)$$

Its time derivative, $\dot{V}$, can be obtained as

$$\begin{aligned} \dot{V} &= \left[h_f \cdot A_f + B_c \cdot u_S + \bar{e}_{\text{mod}} - \dot{y}_r\right]^T \cdot P \cdot e(t) \\ &\quad + e(t)^T \cdot P \cdot \left[h_f \cdot A_f + B_c \cdot u_S + \bar{e}_{\text{mod}} - \dot{y}_r\right] \\ &= 2e(t)^T \cdot P \cdot B_c \cdot u_S + 2e(t)^T \cdot P \\ &\quad \cdot \left[h_f \cdot A_f + \bar{e}_{\text{mod}} - \dot{y}_r\right]. \end{aligned} \quad (23)$$

Substituting (8) into (23) yields

$$\begin{aligned} \dot{V} &\le 2e(t)^T \cdot P \cdot B_c \cdot u_S + 2\left|e(t)^T \cdot P \cdot B_c\right| \\ &\quad \cdot \left|B_c^{-1} \cdot \left(h_f \cdot A_f + \bar{e}_{\text{mod}} - \dot{y}_r\right)\right| \le 0. \end{aligned} \quad (24)$$

Hence, we can infer that if the supervisory control signal (8) is injected into fuzzy system (2), time derivative of the Lyapunov candidate $\dot{V} \le 0$ and system (2) is UUB stable.

4. Numerical Simulation

In this section, the proposed control strategy is applied on a five-degree-of-freedom aircraft system described in [21] for performance evaluation. We consider the angle of attack α and the roll angle ϕ as outputs to be tracked. Tracking of angle of attack is directly related to tracking of normal acceleration [21], which plays an important role in many practical maneuvers.

Let $b = 3$ be the reference length (m), $\bar{c} = 2$ the mean aerodynamic chord (m), $g = 9.8$ the gravitational acceleration (m/s^2), $I = 50$ the moment of inertia (kg-m^2), p the roll angle rate, q the pitch angle rate, r the yaw angle rate, $Q = 80$ the dynamic pressure (kg/m^2), $S = 5$ the reference wing area (m^2), $V = 100$ the aircraft velocity (m/s), θ the pitch angle, $\delta_a = 0$ the aileron deflection, δ_e the elevator deflection, δ_r the rudder deflection, and $m = 100$ the mass of aircraft (kg); the aircraft dynamics can be written as [21]

$$\begin{aligned} &\begin{bmatrix} \dot{p} \\ \dot{q} \\ \dot{r} \end{bmatrix} = \begin{bmatrix} f_p \\ f_q \\ f_r \end{bmatrix} + \begin{bmatrix} L_{\delta a} & 0 & L_{\delta r} \\ 0 & M_{\delta e} & 0 \\ N_{\delta a} & 0 & N_{\delta r} \end{bmatrix} \cdot \begin{bmatrix} \delta_a \\ \delta_e \\ \delta_r \end{bmatrix}, \\ &\begin{bmatrix} \dot{\alpha} \\ \dot{\beta} \\ \dot{\varphi} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} f_\alpha \\ f_\beta \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -t_\beta \cdot c_\alpha & 1 & -t_\beta \cdot s_\alpha \\ s_\alpha & 0 & -c_\alpha \\ 1 & t_\theta \cdot s_\varphi & t_\theta \cdot c_\varphi \\ 0 & c_\varphi & -s_\varphi \end{bmatrix} \cdot \begin{bmatrix} p \\ q \\ r \end{bmatrix} \\ &\qquad + w, \end{aligned} \quad (25)$$

where

$$
\begin{aligned}
\phi &= y_1, \\
\alpha &= y_2, \\
L_{\delta a} &= I \cdot Q \cdot S \cdot b \cdot C_{l\delta a}, \\
L_{\delta r} &= I \cdot Q \cdot S \cdot b \cdot C_{l\delta r}, \\
M_{\delta e} &= I \cdot Q \cdot S \cdot \bar{c} \cdot C_{m\delta e}, \\
N_{\delta a} &= I \cdot Q \cdot S \cdot b \cdot C_{n\delta a}, \\
N_{\delta r} &= I \cdot Q \cdot S \cdot b \cdot C_{n\delta r}, \\
f_p &= \frac{Q \cdot S \cdot I \cdot b^2 \cdot C_{lp} \cdot p}{2V}, \\
f_q &= \frac{Q \cdot S \cdot I \cdot \bar{c}^2 \cdot C_{mq} \cdot q}{2V}, \\
f_r &= \frac{Q \cdot S \cdot I \cdot b^2 \cdot C_{nr} \cdot r}{2V}, \\
f_\alpha &= \frac{-Q \cdot S \cdot C_{L\alpha} \cdot \alpha + m \cdot g \cdot \left(c_\theta \cdot c_\phi \cdot c_\alpha + s_\theta \cdot s_\alpha\right)}{m \cdot V \cdot c_\beta}, \\
f_\beta &= \frac{Q \cdot S \cdot C_{Y\beta} \cdot \beta + m \cdot g \cdot \left[s_\theta \cdot c_\alpha \cdot s_\beta + c_\theta \cdot s_\phi \cdot c_\beta - c_\theta \cdot c_\phi \cdot s_\alpha \cdot s_\beta\right]}{m \cdot V}, \\
t_\beta &= \tan(\beta), \\
s_\beta &= \sin(\beta), \\
c_\beta &= \cos(\beta), \\
s_\theta &= \sin(\theta), \\
c_\theta &= \cos(\theta), \\
t_\theta &= \tan(\theta), \\
c_\alpha &= \cos(\alpha), \\
s_\alpha &= \sin(\alpha), \\
s_\phi &= \sin(\phi), \\
c_\phi &= \cos(\phi).
\end{aligned}
\tag{26}
$$

In the following simulation, we assume

$$
\begin{aligned}
C_{l\delta a} &= -10^{-4}, \\
C_{l\delta r} &= 10^{-2}, \\
C_{m\delta e} &= -1.6 \times 10^{-4}, \\
C_{n\delta a} &= 10^{-2}, \\
C_{n\delta r} &= -10^{-4}, \\
C_{lp} &= -3.8 \times 10^{-2}, \\
C_{mq} &= -0.9 \times 10^{-2},
\end{aligned}
$$

$$C_{nr} = -4.5 \times 10^{-3},$$
$$C_{L\alpha} = 2.8 \times 10^{-1},$$
$$C_{Y\beta} = -2.8. \tag{27}$$

For the inner fuzzy control law, we select the following membership functions:

$$\begin{aligned}
\mu_{F_1^1}(y_1) &= \exp\left[-\frac{1}{2}\left(\frac{y_1 - c_{11}}{d_{11}}\right)^2\right],\\
\mu_{F_1^2}(y_1) &= \exp\left[-\frac{1}{2}\left(\frac{y_1 - c_{12}}{d_{12}}\right)^2\right],\\
\mu_{F_1^3}(y_1) &= \exp\left[-\frac{1}{2}\left(\frac{y_1 - c_{13}}{d_{13}}\right)^2\right],\\
\mu_{F_1^4}(y_1) &= \exp\left[-\frac{1}{2}\left(\frac{y_1 - c_{14}}{d_{14}}\right)^2\right],\\
\mu_{F_2^1}(y_2) &= \exp\left[-\frac{1}{2}\left(\frac{y_2 - c_{21}}{d_{21}}\right)^2\right],\\
\mu_{F_2^2}(y_2) &= \exp\left[-\frac{1}{2}\left(\frac{y_2 - c_{22}}{d_{22}}\right)^2\right],\\
\mu_{F_2^3}(y_2) &= \exp\left[-\frac{1}{2}\left(\frac{y_2 - c_{23}}{d_{23}}\right)^2\right],\\
\mu_{F_2^4}(y_2) &= \exp\left[-\frac{1}{2}\left(\frac{y_2 - c_{24}}{d_{24}}\right)^2\right],
\end{aligned} \tag{28}$$

where

$$\begin{aligned}
&c_{11} = c_{21} = 0,\\
&c_{12} = c_{22} = 0.4,\\
&c_{13} = c_{23} = 0.8,\\
&c_{14} = c_{24} = 1.2,\\
&d_{11} = d_{21} = d_{12} = d_{22} = d_{13} = d_{23} = d_{14} = d_{24} = 0.4.
\end{aligned} \tag{29}$$

Furthermore, 8 fuzzy rules of the following form comprise the fuzzy rule base:

$R^{(1)}$: if y_1 is F_1^j, then $f_1 = A_{1j}$ for $j = 1, 2, 3, 4$ and $l = 1, 2, 3, 4$.

$R^{(2)}$: if y_2 is F_2^j, then $f_2 = A_{2j}$ for $j = 1, 2, 3, 4$ and $l = 5, 6, 7, 8$.

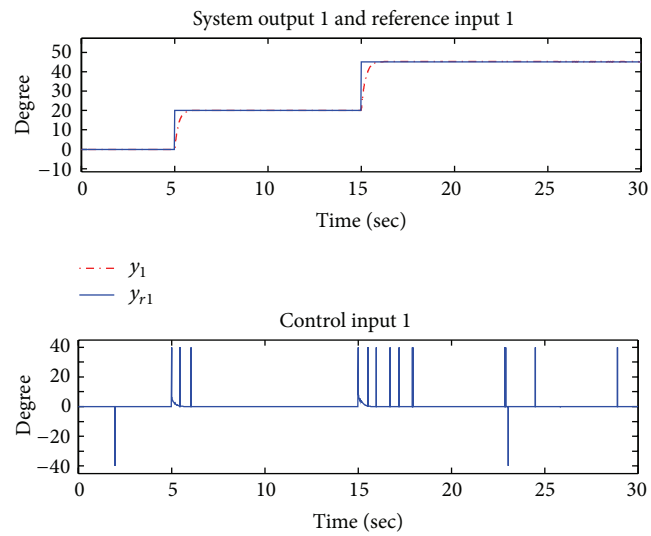


FIGURE 2: Time history of system output y_1 (the angle of attack, α), reference y_{r1}, and control input u_1.

Then, we obtain the following initial system parameters:

$$\begin{aligned}
A_{11} &= -0.1338,\\
A_{12} &= 0.5183,\\
A_{13} &= -1.0891,\\
A_{14} &= 1.546,\\
A_{21} &= -1.0328,\\
A_{22} &= 1.669,\\
A_{23} &= -0.0957,\\
A_{24} &= -0.3342,\\
B_c &= \begin{bmatrix} 16.5381 & -0.4429 \\ 1.1439 & 0.4489 \end{bmatrix}.
\end{aligned} \tag{30}$$

Finally, we design the following control gains:

$$\begin{aligned}
K_P &= \begin{bmatrix} -0.52 & 0 \\ 0 & -51 \end{bmatrix},\\
K_D &= \begin{bmatrix} -0.05 & 0 \\ 0 & -4.9 \end{bmatrix}.
\end{aligned} \tag{31}$$

The tracking performances of α and ϕ, together with the reference (or command), are presented in Figures 2 and 3. These figures show the responses with several step reference inputs. The disturbance is $w = [1, 1, 1, 1]^T \cdot \delta(t-2)$, a burst at $t = 2$ s.

From the simulation results, it is clear that the output tracks the desired command asymptotically with small transient errors, and the zero dynamics remain stable for all the simulated interval.

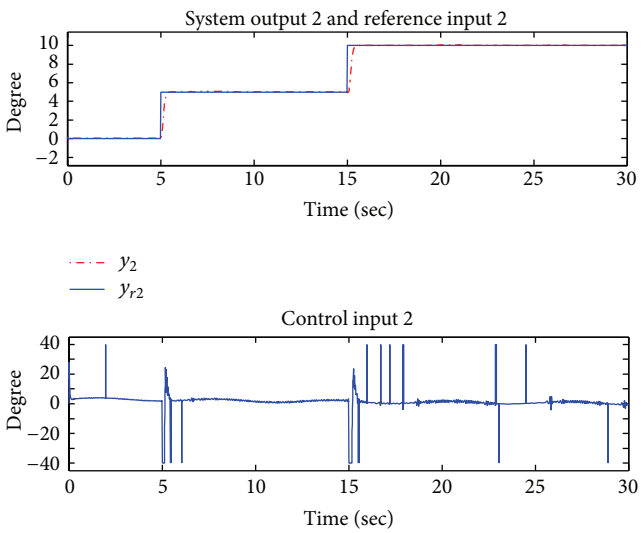


FIGURE 3: Time history of system output y_2 (the roll angle, ϕ), reference y_{r2}, and control input u_2.

5. Conclusion

We propose a two-level adaptive control scheme for nonlinear systems, such as the aircraft, which are MIMO and suffer from nonminimum phase phenomena. The control scheme is composed of an inner-level adaptive fuzzy PD control law and an outer-level supervisory control law. Importantly, the outer-level controller of the two-level scheme is designed based on a fuzzy model taking nonminimum phase phenomena and modeling error explicitly into account. Special care is taken of the nonminimum phase fuzzy subsets by restricting the magnitude of parameters in the singular-value decomposition operation.

The control strategy is much simpler and applicable to general MIMO, nonlinear, and nonminimum phase systems when compared with [3–5]. Simulation results of the application of the proposed control scheme on a five-degree-of-freedom nonlinear aircraft model verify its effectiveness.

Appendix

Derivation of the Condition of (17)

Consider

$$\begin{aligned}
J &= \int_0^{t_f} \left[e(t)^T \cdot \left(Q + P^T \cdot R^{-T} \cdot P\right) \cdot e(t) - \rho^2 \cdot e_{\text{mod}}^T \cdot e_{\text{mod}}\right] \cdot dt = e(0)^T \cdot P \cdot e(0) - e\left(t_f\right)^T \cdot P \cdot e\left(t_f\right) + \gamma_f^{-1} \cdot \text{trace}\left(\widetilde{A}_f(0)^T \cdot \widetilde{A}_f(0)\right) - \gamma_f^{-1} \cdot \text{trace}\left(\widetilde{A}_f\left(t_f\right)^T \cdot \widetilde{A}_f\left(t_f\right)\right) \\
&\quad + \int_0^{t_f} \left[e(t)^T \cdot \left(Q + P^T \cdot R^{-T} \cdot P\right) \cdot e(t) - \rho^2 \cdot e_{\text{mod}}^T \cdot e_{\text{mod}} + \dot{e}(t)^T \cdot P \cdot e(t) + e(t)^T \cdot P \cdot \dot{e}(t) + \gamma_f^{-1} \cdot \text{trace}\left(\dot{\widetilde{A}}_f(t)^T \cdot \widetilde{A}_f(t)\right) + \gamma_f^{-1} \cdot \text{trace}\left(\widetilde{A}_f(t)^T \cdot \dot{\widetilde{A}}_f(t)\right)\right] dt \\
&\le e(0)^T \cdot P \cdot e(0) + \gamma_f^{-1} \cdot \text{trace}\left(\widetilde{A}_f(0)^T \cdot \widetilde{A}_f(0)\right) \\
&\quad + \int_0^{t_f} \left\{e(t)^T \cdot \left(Q + P^T \cdot R^{-T} \cdot P\right) \cdot e(t) - \rho^2 \cdot e_{\text{mod}}^T \cdot e_{\text{mod}} + \left[\alpha \cdot e(t) + h_f(y) \cdot \widetilde{A}_f(t) - R^{-1} \cdot P \cdot e(t) + e_{\text{mod}}\right]^T \cdot P \cdot e(t) + e(t)^T \cdot P \cdot \left[\alpha \cdot e(t) + h_f(y) \cdot \widetilde{A}_f(t) - R^{-1} \cdot P \cdot e(t) + e_{\text{mod}}\right] + \gamma_f^{-1} \cdot \text{trace}\left(\dot{\widetilde{A}}_f(t)^T \cdot \widetilde{A}_f(t)\right) + \gamma_f^{-1} \cdot \text{trace}\left(\widetilde{A}_f(t)^T \cdot \dot{\widetilde{A}}_f(t)\right)\right\} dt = e(0)^T \cdot P \cdot e(0) + \gamma_f^{-1} \cdot \text{trace}\left(\widetilde{A}_f(0)^T \cdot \widetilde{A}_f(0)\right) + \int_0^{t_f} \left\{e(t)^T \cdot \left[\alpha^T \cdot P + P \cdot \alpha + Q + P^T \cdot R^{-T} \cdot P\right] \cdot e(t) - \rho^2 \cdot e_{\text{mod}}^T \cdot e_{\text{mod}} + \left[h_f(y) \cdot \widetilde{A}_f(t) - R^{-1} \cdot P \cdot e(t) + e_{\text{mod}}\right]^T \cdot P \cdot e(t) + e(t)^T \cdot P \cdot \left[h_f(y) \cdot \widetilde{A}_f(t) - R^{-1} \cdot P \cdot e(t) + e_{\text{mod}}\right] + \gamma_f^{-1} \cdot \text{trace}\left(\dot{\widetilde{A}}_f(t)^T \cdot \widetilde{A}_f(t)\right) + \gamma_f^{-1} \cdot \text{trace}\left(\widetilde{A}_f(t)^T \cdot \dot{\widetilde{A}}_f(t)\right)\right\} dt < e(0)^T \cdot P \cdot e(0) + \gamma_f^{-1} \cdot \text{trace}\left(\widetilde{A}_f(0)^T \cdot \widetilde{A}_f(0)\right) \\
&\quad + \int_0^{t_f} \left\{-\rho^{-2} e(t)^T \cdot P^T \cdot P \cdot e(t) - \rho^2 \cdot e_{\text{mod}}^T \cdot e_{\text{mod}} + e_{\text{mod}}^T \cdot P \cdot e(t) + e(t)^T \cdot P \cdot e_{\text{mod}} + \left[h_f(y) \cdot \widetilde{A}_f(t)\right]^T \cdot P \cdot e(t) + e(t)^T \cdot P \cdot h_f(y) \cdot \widetilde{A}_f(t) + \gamma_f^{-1} \cdot \text{trace}\left(\dot{\widetilde{A}}_f(t)^T \cdot \widetilde{A}_f(t)\right) + \gamma_f^{-1} \cdot \text{trace}\left(\widetilde{A}_f(t)^T \cdot \dot{\widetilde{A}}_f(t)\right)\right\} dt = e(0)^T \cdot P \cdot e(0) + \gamma_f^{-1} \cdot \text{trace}\left(\widetilde{A}_f(0)^T \cdot \widetilde{A}_f(0)\right) - \int_0^{t_f} \left[\rho \cdot e_{\text{mod}} - \rho^{-1} \cdot P \cdot e(t)\right]^T \cdot \left[\rho \cdot e_{\text{mod}} - \rho^{-1} \cdot P \cdot e(t)\right] dt \le e(0)^T \cdot P \cdot e(0) + \gamma_f^{-1} \cdot \text{trace}\left(\widetilde{A}_f(0)^T \cdot \widetilde{A}_f(0)\right).
\end{aligned} \tag{A.1}$$

This completes the proof.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to acknowledge the support of the Ministry of Science and Technology of the Republic of China and the former National Science Council, under Contracts NSC 100-2628-E-468-001, 101-2221-E-468-024, 102-2221-E-468-016, 102-2221-E-182-073, 103-2221-E-468-009-MY2, MOST 103-2221-E-182-045, and 104-2221-E-182-008-MY2; Asia University under Contracts 98-ASIA-02, 100-asia-35, and 101-asia-29; Chang Gung University; and Chang Gung Memorial Hospital, Taiwan, under Contracts CMRPD2C0052 and CMRPD2C0053.

References

[1] J. Liu and Q. Huang, "Robust H_∞ control for switched nonlinear system with multiple delays," *Mathematical Problems in Engineering*, vol. 2015, Article ID 852303, 10 pages, 2015.

[2] R.-E. Precup, R.-C. David, E. M. Petriu, S. Preitl, and M.-B. Radac, "Novel adaptive charged system search algorithm for optimal tuning of fuzzy controllers," *Expert Systems with Applications*, vol. 41, no. 4, pp. 1168–1175, 2014.

[3] S. S. Ge and C. Wang, "Adaptive neural control of uncertain MIMO nonlinear systems," *IEEE Transactions on Neural Networks*, vol. 15, no. 3, pp. 674–692, 2004.

[4] T.-H. S. Li and Y.-C. Huang, "MIMO adaptive fuzzy terminal sliding-mode controller for robotic manipulators," *Information Sciences*, vol. 180, no. 23, pp. 4641–4660, 2010.

[5] V. Nekoukar and A. Erfanian, "Adaptive fuzzy terminal sliding mode control for a class of MIMO uncertain nonlinear systems," *Fuzzy Sets and Systems*, vol. 179, pp. 34–49, 2011.

[6] T. Qian and S. Miao, "Discrete-time nonlinear control of VSC-HVDC system," *Mathematical Problems in Engineering*, vol. 2015, Article ID 929467, 11 pages, 2015.

[7] M. Erhard and H. Strauch, "Sensors and navigation algorithms for flight control of tethered kites," in *Proceedings of the European Control Conference (ECC '13)*, pp. 1–6, Zürich, Switzerland, July 2013.

[8] V. Dobrokhodov, E. Xargay, N. Hovakimyan, I. Kaminer, C. Cao, and I. M. Gregory, "Multicriteria analysis of an $L1$ adaptive flight control system," *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 227, no. 4, pp. 413–427, 2013.

[9] L. Benvenuti, M. D. Di Benedetto, and J. W. Grizzle, "Approximate output tracking for nonlinear non-minimum phase systems with an application to flight control," *International Journal of Robust and Nonlinear Control*, vol. 4, no. 3, pp. 397–414, 1994.

[10] A. Isidori and C. I. Byrnes, "Output regulation of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 35, no. 2, pp. 131–140, 1990.

[11] S. A. Al-Hiddabi, "Design of a flight control system for a non-minimum phase 5 DOF aircraft model," in *Proceedings of the 11th Mediterranean Conference on Control and Automation (MED '03)*, pp. 18–20, Rhodes Island, Greece, June 2003.

[12] Y.-Z. Chang and Z.-R. Tsai, "Supervised adaptive control of unknown nonlinear systems using fuzzily blended time-varying canonical model," in *New Trends in Applied Artificial Intelligence*, vol. 4570 of *Lecture Notes in Computer Science*, pp. 464–472, Springer, Berlin, Germany, 2007.

[13] Q. Gao, G. Feng, Y. Wang, and J. Qiu, "Universal fuzzy controllers based on generalized T-S fuzzy models," *Fuzzy Sets and Systems*, vol. 201, pp. 55–70, 2012.

[14] E. P. Klement and R. Mesiar, "A concept of universal fuzzy integrals," in *Proceedings of the Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS '12)*, pp. 1–4, Berkeley, Calif, USA, August 2012.

[15] Q. Gao, X.-J. Zeng, G. Feng, and Y. Wang, "Universal fuzzy models and universal fuzzy controllers based on generalized T-S fuzzy models," in *Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ '12)*, pp. 1–6, Brisbane, Australia, June 2012.

[16] S. S. Sastry and A. Isidori, "Adaptive control of linearizable systems," *IEEE Transactions on Automatic Control*, vol. 34, no. 11, pp. 1123–1131, 1989.

[17] H.-J. Uang and B.-S. Chen, "Robust adaptive optimal tracking design for uncertain missile systems: a fuzzy approach," *Fuzzy Sets and Systems*, vol. 126, no. 1, pp. 63–87, 2002.

[18] A. Stoorvogel, *The H_∞ Control Problem: A State Approach*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1992.

[19] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, Academic Press, London, UK, 1982.

[20] I. Rhee and J. L. Speyer, "A game theoretic approach to a finite-time disturbance attenuation problem," *IEEE Transactions on Automatic Control*, vol. 36, no. 9, pp. 1021–1032, 1991.

[21] J. Lévine, "Automatic flight control systems," in *Analysis and Control of Nonlinear Systems*, Mathematical Engineering, pp. 295–315, Springer, Berlin, Germany, 2009.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 434651, 7 pages
http://dx.doi.org/10.1155/2015/434651

Research Article

Synchronization and Antisynchronization for a Class of Chaotic Systems by a Simple Adaptive Controller

Ling Ren and Rongwei Guo

School of Science, Qilu University of Technology, Jinan 250353, China

Correspondence should be addressed to Rongwei Guo; rongwei_guo@163.com

Received 21 August 2015; Accepted 20 October 2015

Academic Editor: Jonathan N. Blakely

This paper investigates the synchronization and antisynchronization for a class of chaotic system. Firstly, a necessary and sufficient condition is proposed to synchronize and antisynchronization simultaneously for the chaotic systems. Secondly, two methods are obtained to realize coexistence of synchronization and antisynchronization in the chaotic systems, and the corresponding adaptive controllers are also given. Finally, two numerical examples with simulation verify the correctness and effectiveness of the obtained results.

1. Introduction

Since Lorenz firstly found the classical chaotic attractor in 1963 [1], as a most fascinating phenomenon in nonlinear dynamical system, chaos has been intensively studied over the past few decades; see [2, 3] and the references therein. It is well known that Pecora and Carroll firstly investigated the synchronization problem of chaotic systems in 1990 [4], and Ott et al. firstly presented a method to control chaotic systems successfully in 1990 in [5]. From then on, chaos control and chaos synchronization have received a great deal of attention in the area of nonlinear control as the significance of these two problems in both academic research and practical applications, and many important results were obtained; please refer to [6–11].

Up to date, several types of typical synchronization have been identified such as complete synchronization (CS), phase synchronization (PS), lag synchronization (LS), generalized synchronization (GS), anti-phase synchronization (AS), and projective synchronization (PS), and a variety of works have been done about the above problems; see [6, 7, 12–16] and the references therein. It is well known that the master system synchronizes the slave system which is equivalent to the error system that is asymptotically stable. That is to say, chaos synchronization is equivalent to the error system which is asymptotically stable. Similarly, the master system antisynchronizes the slave system which is equivalent to the sum system that is also asymptotically stable. From the view of control theory [17], in order to design a simple and physical controller, the following condition is necessary; that is, $e = 0$ is an equilibrium point of the unforced nominal error system $\dot{e} = f(y) - f(x)$, $\dot{x} = f(x)$, where $e = y - x$, and $E = 0$ is also an equilibrium point of the unforced nominal sum system $\dot{E} = f(y) + f(x)$, $\dot{x} = f(x)$, where $E = y + x$. Obviously, $e = 0$, that is, $y = x$, is an equilibrium point of the error system $\dot{e} = f(y) - f(x)$. Whereas, $E = 0$, that is, $y = -x$, is an equilibrium point of the error system $\dot{E} = f(y) + f(x)$ if and only if $f(-x) = -f(x)$. Thus, the antisynchronization problem is more complex than the synchronization problem. However, this necessary condition is not considered in the most of the existing works on antisynchronization of chaotic systems [14, 15]. Although the authors have solved the antisynchronization of chaotic systems successfully, the controllers that have been obtained are complex; that is, some terms in those controllers are needed to counteract the redundant terms which make $E = 0$ not the equilibrium point of the sum system $\dot{E} = f(y) + f(x)$. For example, $x_2z_2 + x_1z_1$ in u_2 of (14) counteracts the redundant term $-x_2z_2 - x_1z_1$ in error system (13), and $-x_1y_1 - x_2y_2$ in u_3 of (14) also does; for details please see [15].

It should be pointed out that most of the existing works focus on investigating the same kind synchronization in a given chaotic system; that is, all the states of the slave system have the same kind synchronization to the corresponding

states of the master system. For example, when we say that two systems are synchronized (or antisynchronized, or lagsynchronized, or something else) with each other, it means that each pair of the states between the interactive systems is complete synchronous (or antisynchronous, or something else). In [18], the authors firstly pointed out the coexistence and switching of anticipating synchronization and lag synchronization in an optical system. From then on, some important results have been obtained; see [19, 20]. However, there are no results which can give some conditions or algorithms to select what variables in the master chaotic system which can synchronize or antisynchronize the corresponding variables in the slave chaotic systems have been published so far. Therefore, the coexistence of synchronization and antisynchronization of a class of chaotic or hyperchaotic systems needs further research.

Motivated by the above two reasons, we investigate the synchronization and antisynchronization for a class of chaotic systems in this paper. Firstly, for a class of chaotic systems, we obtain a necessary and sufficient condition with which the master system can synchronize and antisynchronize the slave system simultaneously. Secondly, we give two methods to realize coexistence of synchronization and antisynchronization in the chaotic systems and design the corresponding adaptive controllers. Finally, two numerical examples with simulation verify the correctness and effectiveness of the obtained results.

2. Preliminary Knowledge

This paper studies the synchronization and antisynchronization for a class of chaotic systems by adaptive control method. In order to develop this paper, some assumption and definitions are introduced firstly.

Assumption 1 (see [17]). $x_e = 0$ is an equilibrium of the nonlinear system $\dot{x} = f(x)$; that is, $f(x_e) = 0$.

Remark 2. Assumption 1 is a basic assumption of the system control theory. Without loss of generality, if $x_e \neq 0$, we can obtain a new system $\dot{y} = f(y + x_e)$ whose equilibrium is $y_e = 0$ by making a coordinate transform $y = x - x_e$.

Definition 3 (see [12]). Consider the following chaotic system:

$$\dot{x} = f(x), \tag{1}$$

where $x \in \mathbb{R}^n$ is the state and $f(x)$ is a smooth nonlinear vector function.

Let system (1) be the master system; then the corresponding slave system with controller u is given as

$$\dot{y} = f(y) + u, \tag{2}$$

where $y \in \mathbb{R}^n$ is the state and $u = (u_1, u_2, \ldots, u_n)^T$ is the controller to be designed.

Let $e = y - x$, and the error system is described as

$$\dot{e} = f(y) - f(x) + u = F(x, e) + u. \tag{3}$$

We call master chaotic system (1) and slave system (2) reach complete synchronization if $\lim_{t\to\infty}\|e(t)\| = 0$.

Definition 4 (see [14]). Consider master system (1) and slave system (2), let $E = y + x$, and the sum system is described as

$$\dot{E} = f(y) + f(x) + u = G(x, E) + u. \tag{4}$$

We say master system (1) and slave (2) get antisynchronization if $\lim_{t\to\infty}\|E(t)\| = 0$.

Definition 5 (see [21]). Consider master system (1) and slave system (2). If error system (3) and sum system (4) can realize stabilization under the controllers $u = K(e)$ and $u = K(E)$, respectively, where $K(,)$ is a smooth function, we say master system (1) synchronizes and antisynchronizes slave system (2) simultaneously; that is, slave system (1) can synchronize and antisynchronize slave system (2) using a controller with the same form.

Definition 6 (see [18]). Consider master system (1) and slave system (2). If there exists a controller u satisfying $\lim_{t\to\infty} e_i = \lim_{t\to\infty} x_i - y_i = 0$ and $\lim_{t\to\infty} E_j = \lim_{t\to\infty} x_j + y_j = 0$, where $i \neq j$, and $i, j \in \Lambda = \{1, 2, \ldots, n\}$, we say master system (1) and slave system (2) can realize the coexistence of synchronization and antisynchronization. That is to say, some variables (e_i) get synchronization, while some variables (E_j) realize antisynchronization.

With the development of this paper, we introduce our previous result which can make the error system or the sum system reach stabilization.

Lemma 7 (see [13]). *Consider error system (3). If $e_i = 0$ and the remainder error system $\dot{e}_k = F_k(x, 0, e_k)$ is asymptotically stable, then the controlled error system is given as $\dot{e}_i = F_i(x, e) + k_1 e_i$, and $\dot{e}_k = F_k(x, e)$, where $i \neq k$, $i, k \in \Lambda$, and the feedback k_1 is updated according to the following update law:*

$$\dot{k}_1 = -\gamma \sum_{i\in\Lambda} e_i^2. \tag{5}$$

Similarly, for sum system (4), if $E_j = 0$ and the remainder sum system $\dot{E}_l = G_l(x, 0, E_l)$ is asymptotically stable, where $j \neq l$, $j, l \in \Lambda$, then the controlled sum system is given as $\dot{E}_j = G_j(x, E) + k_1 E_j$, and $\dot{E}_l = G_l(x, E)$, and the feedback k_1 is updated according to the following update law:

$$\dot{k}_1 = -\gamma \sum_{j\in\Lambda} E_j^2, \tag{6}$$

where γ is a positive constant.

3. Main Results

In this section, we firstly give a necessary and sufficient condition for a class of chaotic systems, by which we can determine whether master system (1) and slave system (2) realize synchronization and antisynchronization simultaneously or not.

Theorem 8. *Master system (1) and slave system (2) realize synchronization and antisynchronization simultaneously if and only if $f(-x) = -f(x)$; that is, $f(x)$ is the odd function.*

($\Rightarrow$) Since $f(-x) = -f(x)$, obviously, $e = 0$ and $E = 0$ are the equilibria of unforced nominal error system (3) and sum system (4), that is, $u = 0$, respectively. Then, if the controller $u = K(e)$ can stabilize error system (3), sum system (4) can be stabilized by the controller $u = K(E)$.

($\Leftarrow$) If there exist the controllers $u = K(e)$ and $u = K(E)$ stabilizing error systems (3) and (4), respectively, we can obtain that $e = 0$ and $E = 0$ are the equilibria of unforced nominal error system (3) and sum system (4), respectively. Then, $f(-x) = -f(x)$.

Remark 9. Although, the problem of synchronization and antisynchronization simultaneously of 4-dimension hyperchaotic system has been investigated in [22], no sufficient or necessary and sufficient condition for the general chaotic systems was proposed. Theorem 8 gives a necessary and sufficient condition for a class of chaotic systems.

If $f(x)$ is not an odd function, master system (1) and slave system (2) cannot realize synchronization and antisynchronization simultaneously according to Theorem 8. Under this condition, they can reach the coexistence of synchronization and antisynchronization. Then, we give two methods to realize the coexistence of synchronization and antisynchronization for a class of chaotic systems.

Firstly, consider master system (1) and slave system (2), and let

$$e = y - \alpha x, \tag{7}$$

where x, y, and $e \in \mathbb{R}^n$ and $\alpha = \mathrm{Diag}(\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n)$, $|\alpha_i| = 1$, $i \in \Lambda$. The error system is given as

$$\dot{e} = f(e + \alpha x) - f(x) + u. \tag{8}$$

Remark 10. Obviously, if $\alpha_i = 1$, $i \in \Lambda$, then master system (1) and slave system (2) reach complete synchronization. If some $\alpha_i = -1$, while some $\alpha_j = 1$, $i \neq j \in \Lambda$, we say master system (1) and slave system (2) realize the coexistence of synchronization and antisynchronization.

According to Assumption 1, we give the following conclusion.

Theorem 11. *Consider unforced nominal error system (8); that is, $u = 0$. If α satisfies the condition $f(\alpha x) - f(x) = 0$ and at least one $\alpha_i = -1$ and one $\alpha_j = 1$, where $i, j \in \Lambda$, master system (1) and slave system (2) can reach coexistence of synchronization and antisynchronization.*

Proof. According to Assumption 1, $e = 0$ should be the equilibrium of unforced nominal error system (8), and thus $f(\alpha x) - f(x) = 0$. And there exist at least one $\alpha_i = -1$ and one $\alpha_j = 1$, where $i, j \in \Lambda$, and master system (1) and slave system (2) can reach coexistence of synchronization and antisynchronization according to Definition 6. □

Remark 12. Theorem 11 gives a condition which can determine what variables in master system (1) can synchronize the corresponding variables in slave system (2), while other variables in master system (1) can antisynchronize the corresponding variables in slave system (2).

Inspired by the results in [16], we give another method to realize coexistence of synchronization and antisynchronization for a class of chaotic systems. First of all, master system (1) is rewritten as

$$\dot{x} = \begin{pmatrix} \dot{w} \\ \dot{z} \end{pmatrix} = f(x) = \begin{pmatrix} M(z)\, w \\ N(w, z) \end{pmatrix}, \tag{9}$$

where $w = (x_1, x_2, \ldots, x_r)^T$, $z = (x_{r+1}, x_{r+2}, \ldots, x_n)$, if $N(-w, z) = N(w, z)$; that is, $N(w, z)$ is an even function, and we can get the following conclusion.

Theorem 13. *Consider master system (1) and slave system (2), let $x = (w_m, z_m)^T \in \mathbb{R}^n$, $w_m \in \mathbb{R}^r$, $z_m \in \mathbb{R}^{n-r}$, and master system (1) can be written as*

$$\begin{pmatrix} \dot{w}_m \\ \dot{z}_m \end{pmatrix} = \begin{pmatrix} M(z_m)\, w_m \\ N(w_m, z_m) \end{pmatrix}. \tag{10}$$

The slave system is described as

$$\begin{pmatrix} \dot{w}_s \\ \dot{z}_s \end{pmatrix} = \begin{pmatrix} M(z_s)\, w_s \\ N(w_s, z_s) \end{pmatrix} + u, \tag{11}$$

where $y = (w_s, z_s)^T \in \mathbb{R}^n$, $w_s \in \mathbb{R}^r$, $z_s \in \mathbb{R}^{n-r}$, $u = (u_w, u_z)^T$ is the controller to be designed. If $N(-w, z) = N(w, z)$, master system (1) and slave system (2) realize the coexistence of synchronization and antisynchronization.

Proof. Let $E = w_m + w_s$, $e = z_s - z_m$, and then

$$\dot{E} = \dot{w}_m + \dot{w}_s = M(z_m)\, w_m + M(z_s)\, w_s + u_w, \tag{12}$$

$$\dot{e} = \dot{z}_s - \dot{z}_m = N(w_s, z_s) - N(w_m, z_m) + u_z. \tag{13}$$

Obviously, if $N(-w, z) = N(w, z)$, then $E = 0$, $e = 0$, are the equilibria of unforced nominal sum system (12) and error system (13), respectively. According to the nonlinear control theory, the controller designed can reach the coexistence of synchronization and antisynchronization of system (1) and (2). □

4. Illustrative Example

In this section, we give two numerical examples to illustrate how to use the results we obtained in this paper to realize the synchronization and antisynchronization simultaneously and the coexistence of synchronization and antisynchronization, respectively.

Example 14. Consider the following chaotic system [23]:

$$\begin{aligned} \dot{x}_1 &= px_2 - \frac{2p}{7}x_1^3 + \frac{p}{7}x_1, \\ \dot{x}_2 &= x_1 - x_2 + x_3, \\ \dot{x}_3 &= -qx_2, \end{aligned} \tag{14}$$

where $p = 10$ and $q = 100/7$, and system (14) is chaotic.

Obviously, system (14) satisfies Theorem 8; thus this system can realize the synchronization and antisynchronization simultaneously.

Let system (14) be the master system, and then the slave system is as follows:

$$\begin{aligned} \dot{y}_1 &= py_2 - \frac{2p}{7}y_1^3 + \frac{p}{7}y_1 + u_1, \\ \dot{y}_2 &= y_1 - y_2 + y_3 + u_2, \\ \dot{y}_3 &= -qy_2 + u_3, \end{aligned} \tag{15}$$

where $u = (u_1, u_2, u_3)^T$ is the controller to be designed.

Let $e = y - x$, and if $e_1 = 0$, the following remainder error system

$$\begin{aligned} \dot{e}_2 &= -e_2 + e_3, \\ \dot{e}_3 &= -qe_2 \end{aligned} \tag{16}$$

is asymptotically stable. According to Lemma 7, the controller is designed as $u = (k_1e_1, 0, 0)^T$, and $\dot{k}_1 = -e_1^2$.

By the similar procedure, the controller $u = (k_1E_1, 0, 0)^T$ can make master system (14) antisynchronizes slave system (15), where $\dot{k}_1 = -E_1^2$.

Next, we give numerical simulations. The initial values of master system (14) and slave system (15) are $x(0) = (1, -5, -4)^T$, $y(0) = (-2, 3, 1)^T$, and $k_1(0) = -1$. The simulation results are given in Figures 1 and 2, respectively.

Remark 15. From the results of numerical simulation, Figure 1 shows that system (14) and (15) reach synchronization, while Figure 2 shows that system (14) and (15) realize antisynchronization. Thus, the Chua chaotic system can realize the synchronization and antisynchronization simultaneously.

Example 16. Consider the following chaotic system [24]:

$$\begin{aligned} \dot{x}_1 &= (25\beta + 10)(x_2 - x_1), \\ \dot{x}_2 &= (28 - 35\beta)x_1 + (29\alpha - 1)x_2 - x_1x_3, \\ \dot{x}_3 &= -\frac{1}{3}(8 + \beta)x_3 + x_1x_2. \end{aligned} \tag{17}$$

The above system is called unified chaotic system, where $\beta \in [0, 1]$. If $\beta \in [0, 0.8)$, the system is generalized Lorenz system; if $\beta \in (0.8, 1]$, the system is generalized Chen system.

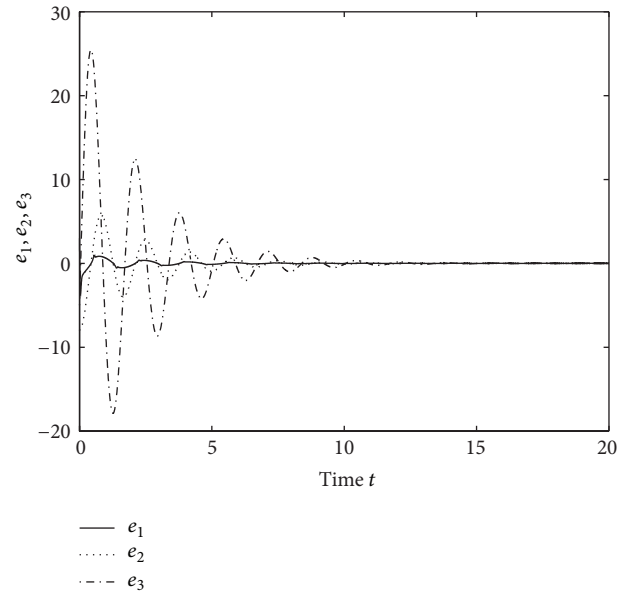


Figure 1: The response of the error system.

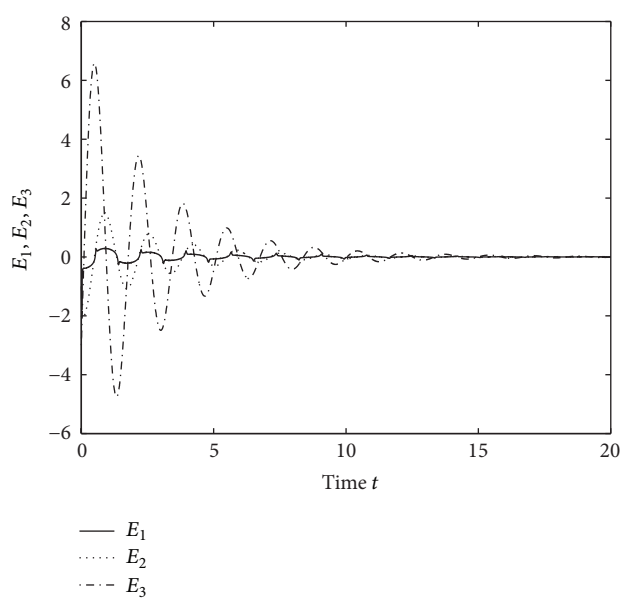


Figure 2: The response of the sum system.

Let system (17) be the master system, and then the slave system is described as

$$\begin{aligned} \dot{y}_1 &= (25\beta + 10)(y_2 - y_1) + u_1, \\ \dot{y}_2 &= (28 - 35\beta)y_1 + (29\beta - 1)y_2 - y_1y_3 + u_2, \\ \dot{y}_3 &= -\frac{1}{3}(8 + \beta)y_3 + y_1y_2 + u_3. \end{aligned} \tag{18}$$

Let $e_i = y_i - \alpha_i x_i$, where $|\alpha_i| = 1$, $i = 1, 2, 3$. Then the unforced nominal error system is described as follows:

$$\begin{aligned}
\dot{e}_1 &= (25\beta + 10)((y_2 - \alpha_1 x_2) - e_1), \\
\dot{e}_2 &= (28 - 35\beta) e_1 + (29\beta - 1) e_2 - e_1 e_3 - \alpha_3 x_3 e_1 \\
&\quad + \alpha_1 x_1 e_3 + (\alpha_1 \alpha_3 - \alpha_2) x_1 x_3, \\
\dot{e}_3 &= e_1 e_2 + \alpha_1 x_2 e_1 - \alpha_2 x_1 e_2 + (\alpha_1 \alpha_2 - \alpha_3) x_1 x_2 \\
&\quad - \frac{8 + \beta}{3} e_3.
\end{aligned} \tag{19}$$

According to Assumption 1, $e = 0$ is the equilibrium of the above unforced nominal error system (19), and the following algebraical equations

$$\begin{aligned}
\alpha_2 &= \alpha_1, \\
\alpha_1 \alpha_3 &= \alpha_2, \\
\alpha_1 \alpha_2 &= \alpha_3
\end{aligned} \tag{20}$$

should be satisfied.

By solving (20), we obtain two solutions: $\alpha_i = 1$, $i = 1, 2, 3$, or $\alpha_1 = \alpha_2 = -1$ and $\alpha_3 = 1$. If $\alpha_i = 1$, $i = 1, 2, 3$, which implies master system (17) and slave system (18) realize complete synchronization. If $\alpha_1 = \alpha_2 = -1$, $\alpha_3 = 1$, which implies that the first two variables of master system (17) antisynchronize the corresponding variables of slave system (18), while the third variable of master system (17) synchronizes the corresponding variables of slave system (18); that is, master system (17) and slave system (18) realize the coexistence of synchronization and antisynchronization.

Select $\alpha_1 = \alpha_2 = -1$, $\alpha_3 = 1$, and unforced nominal sum and error system (19) is given as

$$\begin{aligned}
\dot{E}_1 &= (25\beta + 10)(E_2 - E_1), \\
\dot{E}_2 &= (28 - 35\beta) E_1 + (29\beta - 1) E_2 - E_1 e_3 - x_3 E_1 \\
&\quad - x_1 e_3, \\
\dot{e}_3 &= E_1 E_2 - x_2 E_1 + x_1 E_2 - \frac{8 + \beta}{3} e_3.
\end{aligned} \tag{21}$$

It is easy to obtain that if $E_2 = 0$, then remainder sum and error system (21)

$$\begin{aligned}
\dot{E}_1 &= -(25\beta + 10) E_1, \\
\dot{e}_3 &= -x_2 E_1 - \frac{8 + \beta}{3} e_3
\end{aligned} \tag{22}$$

is asymptotically stable.

According to Lemma 7, forced sum and error system (21) is given as

$$\begin{aligned}
\dot{E}_1 &= (25\beta + 10)(E_2 - E_1), \\
\dot{E}_2 &= (28 - 35\beta) E_1 + (29\beta - 1) E_2 - E_1 e_3 - x_3 E_1 \\
&\quad - x_1 e_3 + k_1 E_2, \\
\dot{e}_3 &= E_1 E_2 - x_2 E_1 + x_1 E_2 - \frac{8 + \beta}{3} e_3;
\end{aligned} \tag{23}$$

that is, the controller is $u = (0, k_1 E_2, 0)^T$ and $\dot{k}_1 = -E_2^2$.

Next, for Example 16, we can obtain the same conclusion by using Theorem 13.

Let $w_m = (x_1, x_2)^T$, $z_m = x_3$, and master chaotic system (17) is rewritten as

$$\begin{aligned}
\dot{w}_m &= M(z_m) u_m = \begin{pmatrix} -(25\alpha + 10) & 0 \\ 0 & 25\alpha + 10 \end{pmatrix} w_m, \\
\dot{z}_m &= N(w_m, z_m) = -\frac{1}{3}(8 + \alpha) z_m + h(w_m),
\end{aligned} \tag{24}$$

where $h(w_m) = x_1 x_2$. Slave chaotic system (18) is also rewritten as

$$\begin{aligned}
\dot{w}_s &= M(z_s) u_s \\
&= \begin{pmatrix} -(25\alpha + 10) & 0 \\ 0 & 25\alpha + 10 \end{pmatrix} w_s + u_w, \\
\dot{z}_s &= N(w_s, z_s) = -\frac{1}{3}(8 + \alpha) z_s + h(w_s) + u_s,
\end{aligned} \tag{25}$$

where $h(w_s) = y_1 y_2$.

It is easy to see that master system (24) and slave system (25) satisfy the conditions of Theorem 13. Therefore, master system (24) and slave system (25) realize the coexistence of synchronization and antisynchronization.

Next, we give numerical simulations. The initial values of master system (17) and slave system (18) are $x(0) = (-2, 3, 4)^T$, $y(0) = (4, -1, -2)^T$, $\beta = 0.75$, and $k_1(0) = -1$. The simulation results are given in Figures 3 and 4, respectively.

Remark 17. From the results of numerical simulation, Figure 3 shows that sum and error system (23) is asymptotically stable, while Figure 4 shows that master system (17) and slave (18) realize coexistence of synchronization and antisynchronization.

5. Conclusions

In this paper, we have investigated the synchronization and antisynchronization for a class of chaotic systems. Firstly, a necessary and sufficient condition has been proposed, with which the master system can synchronize and antisynchronize the slave system simultaneously. Secondly, two methods have been obtained to realize coexistence of synchronization

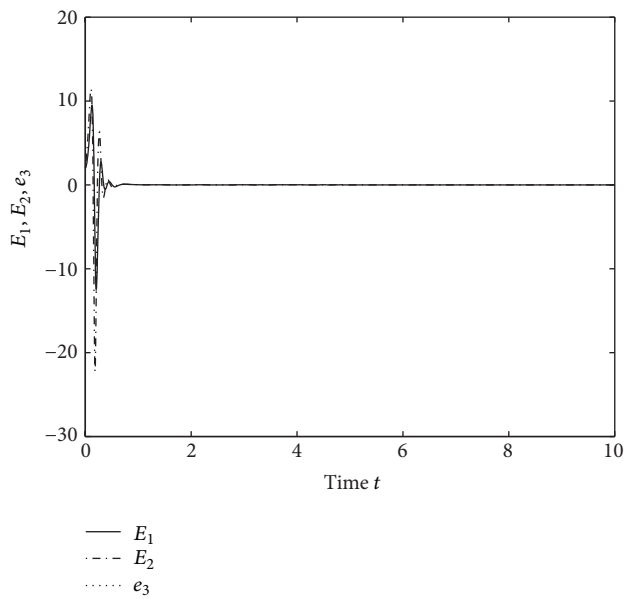


Figure 3: The response of sum and error system.

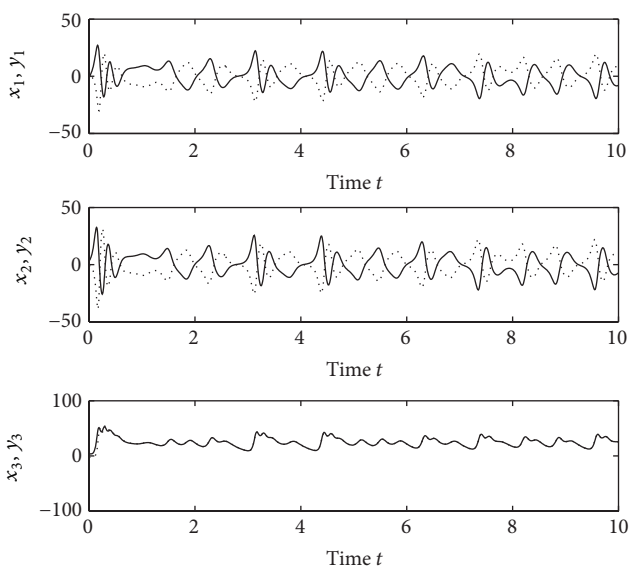


Figure 4: The response of master and slave system.

and antisynchronization for a class of chaotic systems, and the corresponding adaptive controllers have been also given. Two numerical examples with simulation have been used to verify the correctness and effectiveness of the obtained results.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by National Natural Science Foundation of China [61304133, 61305130, and 61374074], China Postdoctoral Science Foundation funded project [2013M541915, 2013M541912, and 2014T70638], and the Scientific Research Foundation of Shandong Province Outstanding Young Scientist Award [BS2013SF023].

References

[1] E. N. Lorenz, "Deterministic nonperiodic flow," *Journal of the Atmospheric Sciences*, vol. 20, no. 2, pp. 130–141, 1963.

[2] O. E. Rössler, "An equation for continuous chaos," *Physics Letters A*, vol. 57, no. 5, pp. 397–398, 1976.

[3] J.-M. Grandmont, "On endogenous competitive business cycles," *Econometrica*, vol. 53, no. 5, pp. 995–1045, 1985.

[4] L. M. Pecora and T. L. Carroll, "Synchronization in chaotic systems," *Physical Review Letters*, vol. 64, no. 8, pp. 821–824, 1990.

[5] E. Ott, C. Grebogi, and J. A. Yorke, "Controlling chaos," *Physical Review Letters*, vol. 64, no. 11, pp. 1196–1199, 1990.

[6] D. Auerbach, C. Grebogi, E. Ott, and J. A. Yorke, "Controlling chaos in high-dimensional systems," *Physical Review Letters*, vol. 69, no. 24, pp. 3479–3482, 1992.

[7] J. Sieber, O. E. Omel'Chenko, and M. Wolfrum, "Controlling unstable chaos: stabilizing chimera states by feedback," *Physical Review Letters*, vol. 112, no. 5, Article ID 054102, 2014.

[8] Z. Q. Zhang, S. Y. Xu, and B. Y. Zhang, "Asymptotic tracking control of uncertain nonlinear systems with unknown actuator nonlinearity," *IEEE Transactions on Automatic Control*, vol. 59, no. 5, pp. 1336–1341, 2014.

[9] Z. Q. Zhang, S. Y. Xu, and B. Y. Zhang, "Exact tracking control of nonlinear systems with time delays and dead-zone input," *Automatica*, vol. 52, pp. 272–276, 2015.

[10] Z. Q. Zhang and S. Y. Xu, "Observer design for uncertain nonlinear systems with unmodeled dynamics," *Automatica*, vol. 51, pp. 80–84, 2015.

[11] Z. Zhang and X.-J. Xie, "Asymptotic tracking control of uncertain nonlinear systems with unknown actuator nonlinearity and unknown gain signs," *International Journal of Control*, vol. 87, no. 11, pp. 2294–2311, 2014.

[12] D. Huang, "Simple adaptive-feedback controller for identical chaos synchronization," *Physical Review E*, vol. 71, no. 3, Article ID 037203, 2005.

[13] R. Guo, "A simple adaptive controller for chaos and hyperchaos synchronization," *Physics Letters A*, vol. 372, no. 34, pp. 5593–5597, 2008.

[14] S. Hammami, M. Benrejeb, M. Feki, and P. Borne, "Feedback control design for Rössler and Chen chaotic systems anti-synchronization," *Physics Letters A*, vol. 374, no. 28, pp. 2835–2840, 2010.

[15] M. Mossa Al-Sawalha, M. S. Noorani, and M. Al-dlalah, "Adaptive anti-synchronization of chaotic systems with fully unknown parameters," *Computers & Mathematics with Applications*, vol. 59, no. 10, pp. 3234–3244, 2010.

[16] D. L. Xu, W. L. Ong, and Z. G. Li, "Criteria for the occurrence of projective synchronization in chaotic systems of arbitrary

dimension," *Physics Letters A*, vol. 305, no. 3-4, pp. 167–172, 2002.

[17] E. A. Barbashin, *Introduction to the Theory of Stability*, Wolters-Noordhoff, Groningen, The Netherlands, 1970.

[18] L. Wu and S. Zhu, "Coexistence and switching of anticipating synchronization and lag synchronization in an optical system," *Physics Letters, Section A: General, Atomic and Solid State Physics*, vol. 315, no. 1-2, pp. 101–108, 2003.

[19] H.-T. Yau, "Synchronization and anti-synchronization coexist in two-degree-of-freedom dissipative gyroscope with nonlinear inputs," *Nonlinear Analysis. Real World Applications*, vol. 9, no. 5, pp. 2253–2261, 2008.

[20] Q. Zhang, J. Lü, and S. Chen, "Coexistence of anti-phase and complete synchronization in the generalized Lorenz system," *Communications in Nonlinear Science and Numerical Simulation*, vol. 15, no. 10, pp. 3067–3072, 2010.

[21] K. H. Wei, "The solution of a transcendental problem and its applications in simultaneous stabilization problems," *IEEE Transactions on Automatic Control*, vol. 37, no. 9, pp. 1305–1315, 1992.

[22] R.-W. Guo, "Simultaneous synchronization and anti-synchronization of two identical new 4D chaotic systems," *Chinese Physics Letters*, vol. 28, no. 4, Article ID 040205, 2011.

[23] M. T. Yassen, "Adaptive control and synchronization of a modified Chua's circuit system," *Applied Mathematics and Computation*, vol. 135, no. 1, pp. 113–128, 2003.

[24] Z. Q. Zhang, H. Shen, and J. L. Li, "Adaptive stabilization of uncertain unified chaotic systems with nonlinear input," *Applied Mathematics and Computation*, vol. 218, no. 8, pp. 4260–4267, 2011.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 787931, 12 pages
http://dx.doi.org/10.1155/2015/787931

Research Article

Adaptive Neural Control Based on High Order Integral Chained Differentiator for Morphing Aircraft

Zhonghua Wu,[1] Jingchao Lu,[1] Jahanzeb Rajput,[1] Jingping Shi,[1] and Wen Ma[2]

[1] *School of Automation, Northwestern Polytechnical University, Xi'an 710072, China*
[2] *Science and Technology on Aircraft Control Laboratory, AVIC Xi'an Flight Automatic Control Research Institute, Xi'an 710065, China*

Correspondence should be addressed to Zhonghua Wu; 463897575@qq.com

Received 28 July 2015; Accepted 17 September 2015

Academic Editor: Xinguang Zhang

This paper presents an adaptive neural control for the longitudinal dynamics of a morphing aircraft. Based on the functional decomposition, it is reasonable to decompose the longitudinal dynamics into velocity and altitude subsystems. As for the velocity subsystem, the adaptive control is proposed via dynamic inversion method using neural network. To deal with input constraints, the additional compensation system is employed to help engine recover from input saturation rapidly. The highlight is that high order integral chained differentiator is used to estimate the newly defined variables and an adaptive neural controller is designed for the altitude subsystem where only one neural network is employed to approximate the lumped uncertain nonlinearity. The altitude subsystem controller is considerably simpler than the ones based on backstepping. It is proved using Lyapunov stability theory that the proposed control law can ensure that all the tracking error converges to an arbitrarily small neighborhood around zero. Numerical simulation study demonstrates the effectiveness of the proposed strategy, during the morphing process, in spite of some uncertain system nonlinearity.

1. Introduction

With the development of morphing wing technology, the flight performance of an aircraft can be improved according to the current flight conditions [1–3]. The morphing aircraft are the flight vehicles that change their shape to either effectuate a change in mission or provide control authority for maneuvering [4, 5], without the use of discrete control surfaces or seams. Aircraft with morphing capability exhibit the distinct advantages of being able to fulfill multiple types of missions and to perform extreme maneuvers not possible with conventional aircraft [6, 7].

The field of morphing aircraft research is composed of a large array of interdisciplinary studies, including wing structure, actuation systems, aerodynamic modeling, non-rigid dynamics, and flight control [8]. A number of studies have focused on optimization of the actuator locations in the morphing structure units [9–11]. Other relative research work that involves the aeroelastics analysis is presented in [12]. The importance of the inertial forces and moments is studied in [13], with the goal of reducing the dynamics that must be dealt with in the flight control design. A methodology which is suitable for numerical calculation of the dynamic loads for a morphing aircraft is presented in [14]. In [15], linear parameter varying modeling is proposed for a folding wing morphing aircraft during the wing morphing process, whereas the longitudinal dynamic responses are numerically simulated based on the quasi-steady aerodynamic assumption.

Despite significant advances in the development of wing structure, actuation systems, and dynamic model, much work remains to be done to effectively control the morphing aircraft. The control system of a morphing aircraft must be capable of achieving consistent and robust performance meanwhile maintaining stability during large variations in the aircraft geometry, which may severely affect aerodynamic forces, moments of inertia, and center of mass.

For the disturbance rejection, a pair of linear controllers is synthesized for a linear input-varying morphing aircraft in [16]. A simple proportional state feedback control integrated with the eigenstructure assignment is proposed for the span-morphing aircraft in [17]. Based on a linear parameter

varying model, self-gain scheduled H_∞ controller is designed for the wing transition process in [18]. On the basis of varying linear parameter and classical methodology, a synthesized multiloop controller of a morphing unmanned aerial vehicle is formulated to guarantee a good performance subjected to large-scale geometrical shape changes in [19].

To cope with system uncertainties, adaptive control and neural network control techniques have been used for decades. For a linear morphing aircraft dynamic model, an indirect adaptive control method is designed in [6], which comprises the receding horizon optimal control law coupled with the modified sequential least squares parameter identification. In [20], a single network adaptive critic tracking controller design for a morphing aircraft is studied, wherein the set of initial weights of the neural network is determined by using a linear system model, which requires offline pretraining. Based on the concepts of feedback linearization, in [21], a combination of dynamic inversion and structured model reference adaptive control is used for the control of a morphing air vehicle. Typically a morphing aircraft exhibits highly nonlinear dynamics characteristics. Because of the morphing aircraft's design and flight condition, it is extremely sensitive to change in physics as well as aerodynamic parameters. Almost all controller designs discussed above are based on linear models. Moreover the input saturation (physical limitation in engine) has not been considered in any work, which usually appears in many practical systems and severely degrades the closed-loop performance [22].

As a powerful nonlinear technique, backstepping control has been used for control system designs with strict-feedback form, extensively. With conventional backstepping, a possible issue is the explosion of complexity. This is caused by the repeating differentiations of certain nonlinear functions. To efficiently handle the system uncertainty in each subsystem, RBFNN with the universal approximation capability is employed in [23, 24]. Since RBFNN is used, we need to take derivatives of those radial basis functions, which will further lead to heavier calculation burden in each step design. Recently, the dynamic surface control was employed to solve this problem and many research results were presented [25, 26]. However, the determination of virtual control terms during the backstepping design requires tedious and complex analysis. More than one neural network is taken for approximation whose complexity increases like the order of the controlled backstepping design.

The motivation of this paper is to present a nonlinear robust adaptive neural controller for the morphing aircraft based on high order integral chained differentiator to achieve stability in the sweeping process where both system uncertainty and input restrictions are considered. The contribution of this paper can be summarized as follows.

Firstly, a nonlinear longitudinal model is derived from a curved-fitted model, with the center of mass position, aerodynamic forces, and the moments of inertia being varied with respect to the morphing parameters. The longitudinal model is then decomposed into altitude and velocity subsystems.

Secondly, the highlight is that the altitude subsystem dynamics is transformed into normal-feedback formulation and a robust adaptive neural controller using HICD is designed where only one neural network is employed to approximate the lumped uncertain system nonlinearity. The controller is considerably simpler than the ones based on backstepping which requires tedious and complex analysis for their virtual control terms. This feature guarantees that the computational burden of the algorithm can be reduced. Moreover the algorithm is convenient for real-time implementation on flight computers. Meanwhile, the adaptive control is proposed for velocity subsystem and an additional compensation system is employed to deal with input constraints, which will help engine recover from input saturation rapidly.

Finally, the Lyapunov synthesis based on stability analysis is used to prove that all the signals in the closed systems are semiglobally uniformly ultimately bounded with tracking error converging to a close neighborhood of origin.

The rest of the paper is organized as follows: Section 2 introduces the model of the morphing aircraft and formulates the normal output-feedback form of the altitude and velocity subsystems of longitudinal dynamics of the morphing aircraft. Section 3 briefly describes the background theory of RBFNN. Section 4 presents the adaptive neural controller design and the stability analysis for altitude and velocity subsystems. The simulation results are presented and discussed in Section 5. Section 6 gives the concluding remarks and future works.

2. Problem Formulation

2.1. Morphing Aircraft Model. The control-oriented model of the longitudinal dynamics of a morphing aircraft considered in this study is based on Seigler [4, 5]. This model comprises five state variables (V, h, α, γ, and q) and two control inputs (δ_e, T), where V is the velocity, h is the altitude, α is angle of attack, γ is the flight path angle (FPA), and q is the pitch rate; δ_e and T represent elevator deflection and thrust force, respectively. Consider

$$\dot{V} = \frac{\left(-D + T\cos\alpha - mg\sin\gamma + F_{Ix}\right)}{m}, \tag{1}$$

$$\dot{h} = V\sin\gamma, \tag{2}$$

$$\dot{\gamma} = \frac{\left[L + T\sin\alpha - mg\cos\gamma - F_{Ikz}\right]}{(mV)}, \tag{3}$$

$$\dot{\alpha} = \frac{\left[-L - T\sin\alpha + mg\cos\gamma + F_{Iz}\right]}{(mV)} + q, \tag{4}$$

$$\dot{q} = -\frac{\dot{I}_y q}{I_y} + \frac{\left(-S_x g\cos\theta + M_A + TZ_T + M_{Iy}\right)}{I_y}, \tag{5}$$

$$\begin{aligned} F_{Ix} &= S_x\left(\dot{q}\sin\alpha + q^2\cos\alpha\right) + 2\dot{S}_x q\sin\alpha - \ddot{S}_x\cos\alpha, \\ F_{Iz} &= F_{Ikz} \\ &= S_x\left(\dot{q}\cos\alpha - q^2\sin\alpha\right) + 2\dot{S}_x q\cos\alpha + \ddot{S}_x\sin\alpha, \\ M_{Iy} &= S_x\left(\dot{V}\sin\alpha + V\dot{\alpha}\cos\alpha - Vq\cos\alpha\right), \end{aligned} \tag{6}$$

where D, L, and M_A represent drag force, lift force, and pitch moment, respectively; m, I_y, and g denote the mass of aircraft, moment of inertia about pitch axis, and gravity constant; F_{Ix}, F_{Iz}, F_{Ikz}, and M_{Iy} represent inertial force and moment caused by morphing process; Z_T is the position of engine in the body axis; S_x denotes the static moment distributed in the body axis of x; the related definitions are given as follows:

$$\begin{aligned}
S_x(\zeta) &\approx [2m_1 r_{1x} + m_3 r_{3x}], \\
Q &= \frac{1}{2\rho_h V^2}, \\
L &= C_L(\zeta) Q S_w(\zeta), \\
D &= C_D(\zeta) Q S_w(\zeta), \\
M_A &= C_m(\zeta) Q S_w(\zeta) c_A(\zeta), \\
C_L(\zeta) &= C_{L0}(\zeta) + C_{L\alpha}(\zeta)\alpha + C_{L\delta e}(\zeta)\delta_e \\
&\approx C_{L0}(\zeta) + C_{L\alpha}(\zeta)\alpha, \\
C_D(\zeta) &= C_{D0}(\zeta) + C_{D\alpha}(\zeta)\alpha + C_{D\alpha 2}(\zeta)\alpha^2, \\
C_m(\zeta) &= C_{m0}(\zeta) + C_{m\alpha}(\zeta)\alpha + C_{m\delta e}(\zeta)\delta_e \\
&\quad + \frac{C_{mq}(\zeta) q c_A(\zeta)}{(2V)},
\end{aligned} \tag{7}$$

where ζ represents the sweep angle, ρ_h denotes the air density, S_w is the wing surface, c_A represents the mean aerodynamic chord, and b is the wingspan. Q and M_A denote the dynamic pressure and pitch moment. C_L, C_D, and C_m are the total aerodynamic lift force coefficient, drag force coefficient, and pitching moment coefficient, respectively. m_1 and m_3 represent the mass of aircraft's wing and body. r_{1x} and r_{3x} denote the position of aircraft's wing and body in the aircraft-body coordinate frame.

We assume that the engine model can be expressed as follows [27].

(A) Engine Rate. The dynamics for the engine speed n is modeled by a first-order linear system with the time constant τ_n and the engine speed reference signal n_c as follows:

$$\dot{n} = -\frac{n}{\tau_n} + \frac{n_c}{\tau_n}. \tag{8}$$

(B) Thrust Force. The thrust force is generated by the propeller and can be expressed with dimensionless coefficients. The dimensionless thrust coefficient is

$$C_{FT}(J) = C_{FT1} + C_{FT2} J + C_{FT3} J^2 \tag{9}$$

with the ratio $J = V_T / D_T \pi n$, where the diameter of the propeller is D_T, the engine speed is n, and the airspeed is V_T. Here we assume that V_T is equal to V. The thrust force is computed as shown below:

$$T = \rho_h n^2 D_T^2 C_{FT}(J). \tag{10}$$

Remark 1. It is important to point out that r_{1x}, r_{3x}, I_y, c_A, S_w, b, C_L, C_D, and C_m are associated with sweep angle ζ in the morphing process. Their functional relationships will be shown later in Section 5.

2.2. System Transformation

(A) Altitude Subsystem. The tracking error of the altitude is defined as $\tilde{h} = h - h_d$. Furthermore, the altitude command is transformed into the desired flight path angle (FPA). The demand of flight path angle is generated as [22]

$$\gamma_d = \arcsin\left[\frac{\left(-k_h \tilde{h} - k_I \tilde{h} + \dot{h}_d\right)}{V}\right]. \tag{11}$$

If $k_h > 0$ and $k_I > 0$ are chosen appropriately and the FPA is controlled to follow γ_d, then the altitude error is regulated to zero exponentially.

Remark 2. Since the control problem considered in this paper only takes into account cruise trajectories and does not consider the aggressive maneuvering, the thrust $T \sin\alpha$ can be neglected since it is generally much smaller than the lift. In order to transform the altitude subsystem into strict-feedback form, F_{Ikz} in (3) is regarded as an unmodeled term.

Define $X = [x_1, x_2, x_3]^T$, $x_1 = \gamma$, $x_2 = \theta$, $x_3 = q$, $\theta = \alpha + \gamma$, $u = \delta_e$; the strict-feedback forms of equations of the altitude (3)–(5) are rewritten as

$$\begin{aligned}
\dot{x}_1 &= f_1(x_1) + g_1(x_1) x_2, \\
\dot{x}_2 &= f_2(x_1, x_2) + g_2(x_1, x_2) x_3, \\
\dot{x}_3 &= f_3(x_1, x_2, x_3) + g_3(x_1, x_2, x_3) u,
\end{aligned} \tag{12}$$

where

$$\begin{aligned}
f_1(x_1) &= \left(\frac{1}{mV}\right)(Q S_w C_{L0} - mg\cos\gamma), \\
f_2(x_1, x_2) &= 0, \\
g_1(x_1) &= \frac{Q S_w C_{L\alpha}}{mV},
\end{aligned}$$

$$f_3(x_1,x_2,x_3)=\frac{\left[QS_wc_A\left(C_{m0}+C_{m\alpha}\alpha+C_{mq}qc_A/(2V)\right)-S_xg\cos\theta-\dot{I}_yq+TZ_T+M_{Iy}\right]}{I_y},$$
$$g_2(x_1,x_2)=1,$$
$$g_3(x_1,x_2,x_3)=\frac{QS_wc_AC_{m\delta e}}{I_y}. \tag{13}$$

Assumption 3. f_1, f_3, f_V, g_1, g_3, and g_V are unknown smooth functions; we assume that there exist positive constants $\overline{g}_{i1}$, $\overline{g}_{i2}$, $\overline{g}_{V1}$, and $\overline{g}_{V2}$ such that $\overline{g}_{i1} \geq g_i(\cdot) \geq \overline{g}_{i2}$, $i = 1,3$, $\overline{g}_{V1} \geq g_V \geq \overline{g}_{V2}$. There also exist constants g_{1d} and g_{3d} such that $g_{1d} \geq |\dot{g}_1|$, $g_{3d} \geq |\dot{g}_3|$. Meanwhile, in this paper, we assume that all the system states can be measured and there is no time-delay in the signal transmission.

Lemma 4 (high order integral chained differentiator [28]). *Suppose the function $\overline{\varsigma}(t)$ and its first $n-1$ derivatives are bounded. Consider the following linear system:*

$$\begin{aligned}\dot{\varsigma}_1&=\varsigma_2\\ \dot{\varsigma}_2&=\varsigma_3\\ &\vdots\\ \dot{\varsigma}_n&=-\frac{a_{f1}}{\chi^n}\left(\varsigma_1-\overline{\varsigma}(t)\right)-\frac{a_{f2}}{\chi^{n-1}}\varsigma_2\cdots\frac{a_{fn}}{\chi}\varsigma_n,\end{aligned} \tag{14}$$

where χ is a small positive constant and parameters a_{f1} to a_{fn} are chosen such that the polynomial $\overline{s}^n + a_{fn}\overline{s}^{n-1} + \cdots + a_{f2}\overline{s} + a_{f1} = 0$ is Hurwitz. Then

$$\lim_{\chi\to 0}\overline{\varsigma}_i=\varsigma^{(i-1)}(t). \tag{15}$$

In the following, we show that original system (12) can be transformed into the normal form with respect to the newly defined state variables. Let $z_1 = x_1$ and $z_2 = \dot{z}_1 = f_1 + g_1x_2$. The derivative of z_2 with respect to time is formulated as

$$\begin{aligned}\dot{z}_2&=\frac{\partial f_1}{\partial x_1}\dot{x}_1+\frac{\partial g_1}{\partial x_1}\dot{x}_1x_2+g_1\dot{x}_2\\ &=\left(\frac{\partial f_1}{\partial x_1}+\frac{\partial g_1}{\partial x_1}x_2\right)(f_1+g_1x_2)+g_1f_2+g_1g_2x_3\\ &=a_2(x_1,x_2)+b_2(x_1,x_2)x_3,\end{aligned} \tag{16}$$

where $a_2(x_1,x_2) = (\partial f_1/\partial x_1 + (\partial g_1/\partial x_1)x_2)(f_1 + g_1x_2) + g_1f_2$, $b_2(x_1,x_2) = g_1g_2$.

Similarly, let $z_3 = \dot{z}_2 = a_2 + b_2x_3$ and its time derivative is induced by

$$\begin{aligned}\dot{z}_3&=\sum_{i=1}^{2}\frac{\partial a_2}{\partial x_i}\dot{x}_i+\sum_{i=1}^{2}\frac{\partial b_2}{\partial x_i}\dot{x}_ix_3+b_2\dot{x}_3\\ &=\sum_{i=1}^{2}\left(\frac{\partial a_2}{\partial x_i}+\frac{\partial b_2}{\partial x_i}x_3\right)(f_i+g_ix_{i+1})+b_2(f_3+g_3u)\\ &=a_3(x_1,x_2,x_3)+b_3(x_1,x_2,x_3)u,\end{aligned} \tag{17}$$

where $a_3 = \sum_{i=1}^{2}(\partial a_2/\partial x_i + (\partial b_2/\partial x_i)x_3)(f_i + g_ix_{i+1}) + b_2f_3$ and $b_3 = g_1g_2g_3$.

As a result, strict-feedback system (12) can be described as the following normal output form with respect to the newly defined state variables z_1, z_2, and z_3:

$$\begin{aligned}\dot{z}_1&=z_2,\\ \dot{z}_2&=z_3,\\ \dot{z}_3&=a_3+b_3u,\\ y&=z_1=x_1.\end{aligned} \tag{18}$$

(B) Velocity Subsystem. With the modeling uncertainties and external disturbance existing, the uncertain nonlinear model can be formulated as

$$\begin{aligned}\dot{V}&=\left[f_{V0}(X_V)+\Delta f_V\right]+g_V(X_V)T+d_V\\ &=f_{V0}(X_V)+g_V(X_V)T+\Delta_V,\end{aligned} \tag{19}$$

where $f_V(X_V) = f_{V0}(X_V) + \Delta f(X_V)$, $g_V = (1/m)\cos\alpha$, $X_V = [x_1, x_2, x_3, V]$. $f_{V0}(X_V)$ is the nominal parts of $f_V(X_V)$; Δf_V is the unknown system uncertainties of $f_V(X_V)$; $d_V(X_V)$ is the external disturbance and $\Delta_V = \Delta f_V(X_V) + d_V$ is the lump of system uncertainty.

Remark 5. It should be noted that a_3, b_3 are totally unknown and need to be approached by NN in the subsequent developments. For the newly defined states z_1, z_2, and z_3, an HICD will be introduced to estimate them. From Assumption 3, it is also noted that there exist constants $\overline{b}_3 > 0$ and $b_{3d} > 0$ such that $b_3 \geq \overline{b}_3$ and $b_{3d} > |\dot{b}_3|$.

3. Neural Networks

In many references of robust adaptive control of uncertain nonlinear systems, the RBFNNs are usually employed as approximate model terms for the unknown nonlinear and continuous function terms using their inherent approximation capabilities [25]. As a class of linearly parameterized NNs, RBFNNs are adopted to approximate the unknown and continuous function $H(X_{\text{in}}) : R^q \rightarrow R$ which can be written as follows:

$$H(X_{\text{in}}) = \widehat{w}^T \Phi(X_{\text{in}}) + \varepsilon, \tag{20}$$

where $X_{\text{in}} \in R^q$ is an input vector of NN, $\widehat{w} \in R^p$ is a weight vector of the NN, $\Phi(X_{\text{in}}) = [\phi_1(X_{\text{in}}), \phi_1 \cdots \phi_p(X_{\text{in}})]^T \in R^p$ is a basis function, ε is the approximation error which satisfies $|\varepsilon| \leq \varepsilon_{\text{re}}$, and ε_{re} is a bounded unknown parameter.

In general, an RBFNN can smoothly approximate any continuous function $H(X_{\text{in}})$ over the compact $\Omega_{X_{\text{in}}} \in R^q$ to any arbitrary accuracy as

$$H(X_{\text{in}}) = w^{*T} \Phi(X_{\text{in}}) + \varepsilon^*, \tag{21}$$

where w^* is the optimal weight value and ε^* is the smallest approximation error. The Gaussian basis function is written in the form of

$$\phi_i(X_{\text{in}}) = \exp\left[-\frac{(X_{\text{in}} - c_i)^T (X_{\text{in}} - c_i)}{m_i^2}\right], \quad i = 1, 2, \ldots, p, \tag{22}$$

where c_i and m_i are the center and width of the neural cell of the ith hidden layer.

Remark 6. There exists an RBFNN in the form of (21) and an optimal parameter vector w^* such that $|H(X_{\text{in}}) - w^{*T}\Phi(X_{\text{in}})| = |\varepsilon^*| < \varepsilon_{\text{re}}$. ε_{re} denotes the supremum of the reconstruction error that is inevitably generated. In what follows, the estimation of w^* is denoted as $\widehat{w}$.

4. Control Design and Stability Analysis

It is easy to note that h is mainly related to δ_e and V is mainly affected by T. Therefore, the dynamics can be decoupled into altitude and velocity subsystem and we design the altitude and velocity controller separately. The structure of the proposed control scheme is presented in Figure 1.

4.1. Adaptive Neural Controller for Altitude Subsystem. The control objective of system (12) is to design an adaptive neural controller, which makes $\gamma \rightarrow \gamma_d$, and therefore $h \rightarrow h_d$, while keeping all the signals involved bounded.

The following controller design is mainly based on the scheme in [29–31]. Vectors Y_d, E and a filtered tracking error s_γ are then defined as follows:

$$Y_d = [y_d, \dot{y}_d, \ddot{y}_d]^T, \tag{23}$$

$$E = Z - Y_d, \tag{24}$$

$$s_\gamma = \left(\frac{d}{dt} + \lambda\right)^2 E = \left[\Lambda^T \;\; 1\right] E, \tag{25}$$

$$e = y - y_d = z_1 - y_d, \tag{26}$$

where $Z = [z_1 \;\; z_2 \;\; z_3]^T$, $\Lambda = [\lambda^2 \;\; 2\lambda]^T$ with $\lambda > 0$.

By employing a high order integral chained differentiator, the estimation of $Z = [z_1 \;\; z_2 \;\; z_3]^T$ is acquired as $\widehat{Z} = [\varsigma_1 \;\; \varsigma_2 \;\; \varsigma_3]^T$. According to the discussion in [28], there exist positive constant ε_h and t^* such that $\forall t > t^*$

$$\left|\widehat{Z} - Z\right| \leq \varepsilon_h. \tag{27}$$

The estimations of E and s_γ using (14) are denoted as given below:

$$\begin{aligned} \widehat{E} &= \widehat{Z} - Y_d, \\ \widehat{s}_\gamma &= \left[\Lambda^T \;\; 1\right] \widehat{E}. \end{aligned} \tag{28}$$

Based on (25), the derivative of s_γ with respect to time can be expressed as

$$\begin{aligned} \dot{s}_\gamma &= \left[0 \;\; \Lambda^T\right] E + \left(y^{(3)} - y_d^{(3)}\right) \\ &= a_3 + b_3 u - y_d^{(3)} + \left[0 \;\; \Lambda^T\right] E \\ &= a_3 + b_3 u + \widehat{v} - \left[0 \;\; \Lambda^T\right] \widetilde{E}, \end{aligned} \tag{29}$$

where $\widehat{v} = -y_d^{(3)} + \left[0 \;\; \Lambda^T\right] \widehat{E}$, $\widetilde{E} = \widehat{E} - E = \widehat{Z} - Z$.

Define

$$u_{ad}^*(X_A, \widehat{v}) = \frac{(a_3 + \widehat{v})}{b_3}. \tag{30}$$

u_{ad}^* is approximated by RBFNN as

$$\begin{aligned} u_{\text{RBF}} &= \widehat{w}_A^T \Phi(X_A), \\ X_A &= \left[X^T, \widehat{v}\right], \end{aligned} \tag{31}$$

where $\widehat{w}_A$ is the estimation of the optimal parameter vector w_A^*, $\widetilde{w}_A = \widehat{w}_A - \widehat{w}_A^*$.

Substituting the unknown s_γ with $\widehat{s}_\gamma$, we determine the control input as follows:

$$u = -k\widehat{s}_\gamma - \widehat{w}_A^T \Phi(X_A). \tag{32}$$

The update law for $\widehat{w}$ is determined as

$$\dot{\widehat{w}}_A = \gamma_A \left(\widehat{s}_\gamma \Phi(X_A) - \sigma_s(\widehat{w}_A) \widehat{w}_A\right), \tag{33}$$

$$\sigma_s(\widehat{w}_A) = \begin{cases} \dfrac{c_\Phi}{\varepsilon_w}, & \text{if } |\widehat{w}_A| > \varepsilon_w \\ 0, & \text{otherwise,} \end{cases} \tag{34}$$

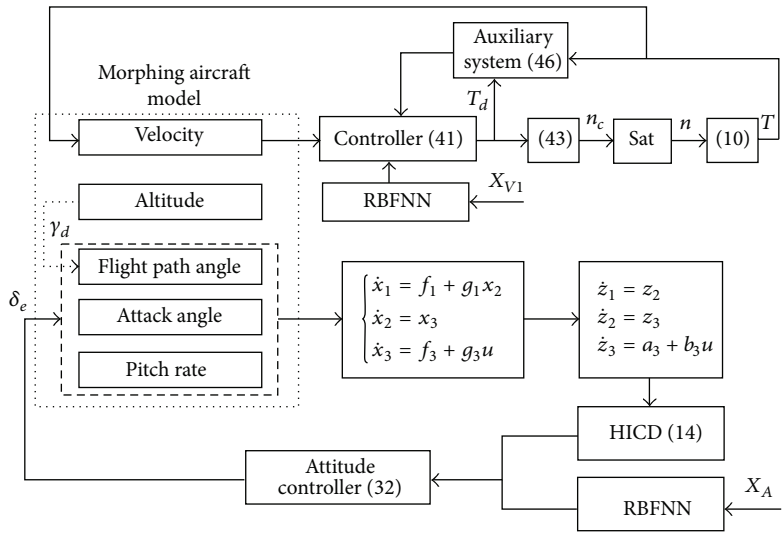


Figure 1: Control scheme.

where ε_W, c_Φ are positive design constants, $\|\Phi(X_A)\| < c_\Phi$, and γ_A denotes the positive learning rate.

Theorem 7. *Consider the adaptive system consisting of (12) under Assumption 3, controller (32) with HICD (14), and adaptive law (33). The filtered error s_γ and $\widetilde{\omega}_A$ are semiglobally uniformly ultimately bounded.*

Proof. Consider the Lyapunov function candidate $L = 1/(2b_3)s_\gamma^2 + 1/(2\gamma_A)\widetilde{\omega}_A^T\widetilde{\omega}_A$. Taking the time derivation of L, we get

$$\begin{aligned}
\dot{L} &= \frac{s_\gamma \dot{s}_\gamma}{b_3} - \frac{\dot{b}_3 s_\gamma^2}{2b_3^2} + \frac{\widetilde{\omega}_A^T \dot{\widehat{\omega}}_A}{\gamma_A} = \frac{1}{b_3} s_\gamma \Big(a_3 + b_3 u + \widehat{v} \\
&\quad - \begin{bmatrix} 0 & \Lambda^T \end{bmatrix} \widetilde{E}\Big) - \frac{\dot{b}_3 s_\gamma^2}{2b_3^2} + \widetilde{\omega}_A^T \left(\widehat{s}_\gamma \Phi - \sigma_s(\widehat{\omega}_A)\widehat{\omega}_A\right) \\
&= \frac{1}{b_3} s_\gamma \Big(a_3 + b_3 u - b_3 u_{ad}^* + b_3 u_{ad}^* + \widehat{v} - \begin{bmatrix} 0 & \Lambda^T \end{bmatrix} \widetilde{E}\Big) \\
&\quad - \frac{\dot{b}_3 s_\gamma^2}{2b_3^2} + \widehat{s}_\gamma \widetilde{\omega}_A^T \Phi - \sigma_s(\widehat{\omega}_A)\widetilde{\omega}_A^T \widehat{\omega}_A = \frac{1}{b_3} s_\gamma \Big(-k b_3 \widehat{s}_\gamma \\
&\quad + b_3 \left(u_{ad}^* - \widehat{\omega}_A^T \Phi\right) - \begin{bmatrix} 0 & \Lambda^T \end{bmatrix} \widetilde{E}\Big) - \frac{\dot{b}_3 s_\gamma^2}{2b_3^2} + s_\gamma \widetilde{\omega}_A^T \Phi \\
&\quad - \left(s_\gamma - \widehat{s}_\gamma\right)\widetilde{\omega}_A^T \Phi - \sigma_s(\widehat{\omega}_A)\widetilde{\omega}_A^T \widehat{\omega}_A = s_\gamma \Bigg(-k s_\gamma \\
&\quad + k\left(s_\gamma - \widehat{s}_\gamma\right) + \left(u_{ad}^* - w_A^{*T}\Phi + w_A^{*T}\Phi - \widehat{\omega}_A^T \Phi\right) \\
&\quad - \frac{\begin{bmatrix} 0 & \Lambda^T \end{bmatrix} \widetilde{E}}{b_3}\Bigg) - \frac{\dot{b}_3 s_\gamma^2}{2b_3^2} + s_\gamma \widetilde{\omega}_A^T \Phi - \left(s_\gamma - \widehat{s}_\gamma\right)\widetilde{\omega}_A^T \Phi \\
&\quad - \sigma_s(\widehat{\omega}_A)\widetilde{\omega}_A^T \widehat{\omega}_A = -k s_\gamma^2 + s_\gamma \times \Bigg(-k \begin{bmatrix} \Lambda^T & 1 \end{bmatrix} \widetilde{E} \\
&\quad + \left(u_{ad}^* - w_A^{*T}\Phi\right) - \widetilde{\omega}_A^T \Phi - \frac{\begin{bmatrix} 0 & \Lambda^T \end{bmatrix} \widetilde{E}}{b_3}\Bigg) - \frac{\dot{b}_3 s_\gamma^2}{2b_3^2} \\
&\quad + s_\gamma \widetilde{\omega}_A^T \Phi + \widetilde{\omega}_A^T \begin{bmatrix} \Lambda^T & 1 \end{bmatrix} \widetilde{E}\Phi - \sigma_s(\widehat{\omega}_A)\widetilde{\omega}_A^T \widehat{\omega}_A \\
&\le -\left(k - \frac{b_{3d}}{2\overline{b}_3^2}\right) s_\gamma^2 + \left|s_\gamma\right| \left(k c_{\lambda 1} \varepsilon_h + \varepsilon_{\mathrm{re}} + \frac{c_{\lambda 2}\varepsilon_h}{\overline{b}_3}\right) \\
&\quad + \widetilde{\omega}_A^T \begin{bmatrix} \Lambda^T & 1 \end{bmatrix} \widetilde{E}\Phi - \sigma_s(\widehat{\omega}_A)\widetilde{\omega}_A^T \widehat{\omega}_A.
\end{aligned} \tag{35}$$

Considering the following facts,

$$\begin{aligned}
\widetilde{\omega}_A^T \begin{bmatrix} \Lambda^T & 1 \end{bmatrix} \widetilde{E}\Phi &\le \frac{1}{8} c_\Phi k_s \|\widetilde{\omega}_A\|^2 + \frac{2}{k_s} c_\Phi \left\|\begin{bmatrix} \Lambda^T & 1 \end{bmatrix} \widetilde{E}\right\|^2 \\
&= \frac{1}{8} c_\Phi k_s \|\widetilde{\omega}_A\|^2 + \frac{2}{k_s} c_\Phi \mu_1^2 \\
2\widetilde{\omega}_A^T \widehat{\omega}_A &= \|\widetilde{\omega}_A\|^2 + \|\widehat{\omega}_A\|^2 - \|w_A^*\|^2 \\
&\ge \|\widetilde{\omega}_A\|^2 - \|w_A^*\|^2,
\end{aligned} \tag{36}$$

we have

$$\begin{aligned}
\dot{L} &\le -\left(k - \frac{b_{3d}}{2\overline{b}_3^2} - \frac{1}{2} C_1\right) s_\gamma^2 \\
&\quad - \left(\frac{1}{2}\frac{c_\Phi}{\varepsilon_w} - \frac{1}{8} c_\Phi k_s\right) \|\widetilde{\omega}_A\|^2 + \frac{2}{k_s} c_\Phi \mu_1^2 \\
&\quad + \frac{1}{2}\frac{c_\Phi}{\varepsilon_w} \|w_A^*\|^2 + \frac{1}{2} \le -\rho L + C,
\end{aligned} \tag{37}$$

where $c_{\lambda 1} = \| \left[\Lambda^T \quad 1\right] \|$, $c_{\lambda 2} = \| \left[0 \quad \Lambda^T\right] \|$, $\mu_1 = \left[\Lambda^T \quad 1\right] \widetilde{E}$, $|\sigma_s(\widehat{w})| \le c_\Phi/\varepsilon_w$, $k_s > 0$, $\|\Phi\| \le c_\Phi$, $|kc_{\lambda 1}\varepsilon_h + \varepsilon_{re} + c_{\lambda 2}\varepsilon_h/\bar{b}_3| = C_1$. ρ and C are given by

$$\rho := \min\left\{\left(k - \frac{b_{3d}}{2\bar{b}_3^2} - \frac{1}{2}C_1\right), \left(\frac{1}{2}\frac{c_\Phi}{\varepsilon_w} - \frac{1}{8}c_\Phi k_s\right)\right\},$$
$$C := \frac{2}{k_s}c_\Phi\mu_1^2 + \frac{1}{2}\frac{c_\Phi}{\varepsilon_w}\left\|w_A^*\right\|^2 + \frac{1}{2}. \tag{38}$$

To ensure the closed-loop stability, the corresponding design parameters should be chosen such that $k - b_{3d}/(2\bar{b}_3^2) - (1/2)C_1 > 0$ and $(1/2)(c_\Phi/\varepsilon_w) - (1/8)c_\Phi k_s > 0$.

According to (37), we have $0 \le L \le C/\rho + [L(0) - C/\rho]e^{-\rho t}$. From (37), we can know that L is convergent; that is, $\lim_{x\to\infty} L = C/\rho$. It can be shown that the filtered signal s_γ and $\widetilde{w}_A$ are semiglobally uniformly bounded. □

Remark 8. (1) The switching function $\sigma_s(\widehat{w}_A)$ is adopted so that the RBFNN can retain the learned information, which is based on a novel σ switching scheme. The adopted switching scheme prevents the loss of information, if ε_w is chosen sufficiently large value such that $\varepsilon_w > |\widehat{w}_A|$ while guaranteeing the boundness of $|\widehat{w}_A|$.

(2) It should be noted that, in this paper, only one RBFNN is employed to approximate the lumped uncertain nonlinear function in the altitude subsystem which highlights the simplicity of our proposed controller. However, at least two RBFNNs need to be used in the backstepping scheme, in [25], which require large computational burden. It is also demonstrated that control law and stability analysis is considerably simpler than the previous backstepping-based algorithms.

4.2. Adaptive Controller for Velocity Subsystem. Define

$$\widetilde{V} = V - V_d. \tag{39}$$

Its time derivative is

$$\dot{\widetilde{V}} = \dot{V} - \dot{V}_d = f_{V0} + g_V T + \Delta_V - \dot{V}_d. \tag{40}$$

By employing an RBFNN $\widehat{w}_V^T\Phi_V(X_{V1})$ to approximate unknown uncertainty Δ_V, we have

$$T_d = \frac{\left[-k_{pV}\widetilde{V} - k_{IV}\int_0^t\left(\widetilde{V} - V_e\right)d\tau - f_{V0} - \widehat{w}_V^T\Phi_V\left(X_{V1}\right) + \dot{V}_d\right]}{g_V}, \tag{41}$$

where $X_{V1} = [V, V_d, \widetilde{V}]$ and k_{pv}, k_{Iv} are the positive design parameters; V_e is the compensatory term which will be defined as follows. T_d represents the desired thrust force.

Equations (9) and (10) are rearranged so as to solve n_d in the following equation:

$$n_d^2\left(C_{FT1}\rho_h D_T^4\right) + n_d\left(\frac{C_{FT2}\rho_h D_T^3 V}{\pi}\right) + \frac{C_{FT3}\rho_h D_T^2 V^2}{\pi^2} - T_d = 0. \tag{42}$$

In order to solve (42) at each sampling time, V is assumed to be constant during the sampling period. Then

$$n_d = \frac{\left(c_{n1}V + \sqrt{c_{n2}V^2 + c_{n3}T_d}\right)}{c_{n4}}, \tag{43}$$

$$n_c = \begin{cases} n_{\max}, & n_d \ge n_{\max} \\ n_d, & n_d \le n_{\max}, \end{cases} \tag{44}$$

where $c_{n1} = -C_{FT2}\rho_h D_T^3/\pi$, $c_{n2} = (C_{FT2}^2 - 4C_{FT1}C_{FT3})\rho_h^2 D_T^6/\pi^2$, $c_{n3} = 4C_{FT1}\rho_h D_T^4$, and $c_{n4} = 2C_{FT1}\rho_h D_T^4$ are the intermediate variables. n_c is the actual engine speed; $n_{\max}$ is the upper limit of n_d.

Define

$$\widetilde{V}_e = \widetilde{V} - V_e, \tag{45}$$

$$\dot{V}_e = -k_{pV}V_e + g_V\left(T - T_d\right),$$
$$V_e(0) = 0. \tag{46}$$

The update law of $\widehat{w}_V$ is determined as

$$\dot{\widehat{w}}_V = \eta_V\left(\widetilde{V}_e\Phi_V\left(X_{V1}\right) - \sigma_V\widehat{w}_V\right), \tag{47}$$

where σ_V is a positive design constant and $\widetilde{w}_V = \widehat{w}_V - w_V^*$; (46) indicates the auxiliary system used to compensate the engine speed saturation.

The derivatives of $\widetilde{V}$ and $\widetilde{V}_e$ with respect to time, $\dot{\widetilde{V}}$ and $\dot{\widetilde{V}}_e$, can be expressed as

$$\begin{aligned}\dot{\widetilde{V}} &= \dot{V} - \dot{V}_d = f_{V0} + g_V T + w_V^{*T}\Phi_V + \varepsilon_V - \dot{V}_d \\ &= f_{V0} + g_V T_d + g_V\left(T - T_d\right) + \widehat{w}_V^T\Phi_V - \widetilde{w}_V^T\Phi_V \\ &\quad + \varepsilon_V - \dot{V}_d \\ &= -k_{pV}\widetilde{V} - k_{IV}\int_0^t\widetilde{V}_e d\tau + g_V\left(T - T_d\right) \\ &\quad - \widetilde{w}_V^T\Phi_V + \varepsilon_V, \\ \dot{\widetilde{V}}_e &= -k_{pV}\widetilde{V}_e - k_{IV}\int_0^t\left(\widetilde{V}_e\right)d\tau - \widetilde{w}_V^T\Phi_V + \varepsilon_V.\end{aligned} \tag{48}$$

Theorem 9. *Consider the adaptive system comprising (19), velocity subsystem controller (41) with adaptive law (47), and auxiliary system (46). $\widetilde{V}_e$ and $\widetilde{w}_V$ are semiglobally uniformly bounded.*

Proof. Consider the Lyapunov candidate function

$$L_V(t) = \frac{1}{2}\widetilde{V}_e^2 + \frac{k_{IV}}{2}\int_0^t\widetilde{V}_e^2 d\tau + \frac{1}{2\eta_V}\widetilde{w}_V^T\widetilde{w}_V. \tag{49}$$

Its time derivative is

$$\begin{aligned}\dot{L}_V &= \widetilde{V}_e\dot{\widetilde{V}}_e + k_{IV}\widetilde{V}_e\int_0^t \widetilde{V}_e d\tau + \frac{\widetilde{w}_V^T\dot{\widetilde{w}}_V}{\eta_V}\\ &= \widetilde{V}_e\left[-k_{pV}\widetilde{V}_e - k_{IV}\int_0^t\left(\widetilde{V}_e\right)d\tau - \widetilde{w}_V^T\Phi_V + \varepsilon_V\right]\\ &\quad + k_{IV}\widetilde{V}_e\int_0^t\widetilde{V}_e d\tau + \widetilde{w}_V^T\left(\widetilde{V}_e\Phi_V - \sigma_V\widehat{w}_V\right)\\ &= -k_{pV}\widetilde{V}_e^2 + \widetilde{V}_e\varepsilon_V - \sigma_V\widetilde{w}_V^T\widehat{w}_V.\end{aligned} \tag{50}$$

Considering the following fact,

$$\begin{aligned}2\widetilde{w}_V^T\widehat{w}_V &= \|\widetilde{w}_V\|^2 + \|\widehat{w}_V\|^2 - \|w_V^*\|^2 \geq \|\widetilde{w}_V\|^2 - \|w_V^*\|^2\\ \widetilde{V}_e\varepsilon_V &\leq \frac{1}{2}\left(\widetilde{V}_e^2 + \varepsilon_V^2\right),\end{aligned} \tag{51}$$

we have the following inequality:

$$\begin{aligned}\dot{L}_V &\leq -\left(k_{pV} - \frac{1}{2}\right)\widetilde{V}_e^2 - \frac{1}{2}\sigma_V\|\widetilde{w}_V\|^2\\ &\quad + \frac{1}{2}\left(\varepsilon_V^2 + \sigma_V\|w_V^*\|^2\right) \leq -\rho_V L_V + C_V,\end{aligned} \tag{52}$$

where ρ_V and C_V are given by $\rho_V := \min\{(k_{pV} - 1/2), \sigma_V/2\}$ and $C_V := \{1/2\varepsilon_V^2 + \sigma_V/2\|w_V^*\|^2\}$.

To ensure the closed-loop stability, the corresponding design parameters k_{pV}, σ_V should be chosen such that $k_{pV} - 1/2 > 0$, $\sigma_V > 0$. According to (52), it can be shown that the signals $\widetilde{V}_e$ and $\widetilde{w}_V$ are semiglobally uniformly bounded. □

Remark 10. In this section, the dynamic inversion control based on RBFNN is proposed for velocity subsystem with input saturation constraints. To handle the input saturation, auxiliary design system (46) is introduced to analyze the effect of saturation constraint and the auxiliary variable V_e is used to design the adaptive law. It is apparent that the constrained control T produced by the designed control command T_d can guarantee the closed-loop system's stability.

5. Numerical Simulation

In this section, the performance of the developed control strategy applied to the longitudinal model of the morphing aircraft is verified by means of simulations. The aircraft model parameters are shown in Table 1. Neural network $\widehat{w}_A^T\Phi(X_A)$ with input vector $X_A = [x_1, x_2, x_3, \widehat{v}]^T$ contains 50 nodes with centers c_{1i} $(i = 1\cdots 50)$ evenly spaced in $[-15^\circ, 15^\circ] \times [-15^\circ, 15^\circ] \times [-15^\circ, 15^\circ] \times [-15^\circ, 15^\circ]$ and widths m_{1i} $(i = 1\cdots 50) = 1$; neural network $\widehat{w}_V^T\Phi(X_{V1})$ with input vector $X_{V1} = [V, V_d, \widetilde{V}_e]^T$ contains 10 nodes with centers c_{2i} $(i = 1\cdots 10)$ evenly spaced in $[10, 50] \times [10, 50] \times [-50, 50]$ and widths m_{2i} $(i = 1\cdots 10) = 5$. The initial condition is set as $X_0 = [\gamma_0, \theta_0, q_0, h_0, V_0] = [0, 0.99512^\circ, 0, 1000\,\text{m}, 30\,\text{m/s}]$, $w_A(0) = 0$, and $w_V(0) = 0$. Control and HICD parameters are set as $k_h = 0.5$, $k_I = 0.01$, $k = 0.025$, $\gamma_A = 0.02$, $\varepsilon_w = 10$, $c_\Phi = 20$, and $\lambda = 5$; $k_{pv} = 5$, $k_{Iv} = 10$, $\eta_V = 10$, and $\sigma_V = 0.01$; $a_{f1} = 10$, $a_{f2} = 10$, $a_{f3} = 10$, and $\chi = 0.04$. Reference commands are smoothened via several second-order filters shown in (53) below. The engine speed saturation $n_{\max}$ which is set at 4900 RPM is deliberately tightened to explore the capability of the designed controller in adhering to the limits. Consider

Table 1: Morphing aircraft parameters for different configurations.

Parameters	$\zeta = 0^\circ$	$\zeta = 30^\circ$	$\zeta = 45^\circ$
$S/(\text{m}^2)v$	1.6040	1.168	0.958
$c_A/(\text{m})$	0.4874	0.411	0.416
$b/(\text{m})$	3.3494	2.981	2.503
$I_y/(\text{kg}\cdot\text{m})$	6.4929	7.882	8.606

$$\begin{aligned}\frac{h_d}{h_{d0}} &= \frac{0.64}{s^2 + 1.6s + 0.64},\\ \frac{V_d}{V_{d0}} &= \frac{1}{s^2 + 2s + 1},\\ \frac{\zeta_d}{\zeta_{d0}} &= \frac{1}{s^2 + 4s + 4}.\end{aligned} \tag{53}$$

Choosing $\zeta = 0^\circ, 5^\circ, \ldots, 45^\circ$ as the 10 reference points, the longitudinal aerodynamic parameters for different variation configurations can be computed through computational fluid dynamics (CFD). Then the aerodynamic parameters of the morphing aircraft during wing-transforming process can be linearly interpolated by those of static configurations with the help of MATLAB:

$$\begin{aligned}C_{L0} &= 0.0042\zeta^3 - 0.1374\zeta^2 - 0.0516\zeta + 0.2291,\\ c_A &= 0.2054\zeta^2 - 0.2520\zeta + 0.4874,\\ C_{L\alpha} &= -1.1264\zeta^3 - 0.4351\zeta^2 + 0.3816 + 4.592,\\ b &= -1.4599\zeta^2 + 0.0644\zeta + 3.3494,\\ C_{D0} &= -0.0024\zeta^3 + 0.0045\zeta^2 + 0.0022\zeta + 0.021,\\ C_{D\alpha} &= -0.0310\zeta^2 - 0.0458\zeta + 0.109,\\ C_{D\alpha 2} &= -1.2990\zeta^4 + 1.8282\zeta^3 - 0.7039\zeta^2 - 0.0258\zeta\\ &\quad + 1.097,\\ S &= -0.8271\zeta + 1.6040,\\ C_{m\alpha} &= 9.6542\zeta^3 - 6.5395\zeta^2 - 6.1887\zeta - 1.5909,\\ C_{m0} &= 0.4239\zeta^2 - 0.4462\zeta^2 - 0.0365,\\ C_{m\delta e} &= -0.1624\zeta^2 - 0.9376\zeta - 0.7889,\\ I_y &= -4.9021\zeta^3 + 6.5774\zeta^2 + 0.5500\zeta + 6.4929,\\ C_{mq} &= 41.4537\zeta^3 - 50.4868\zeta^2 - 9.7741\zeta - 10.673.\end{aligned} \tag{54}$$

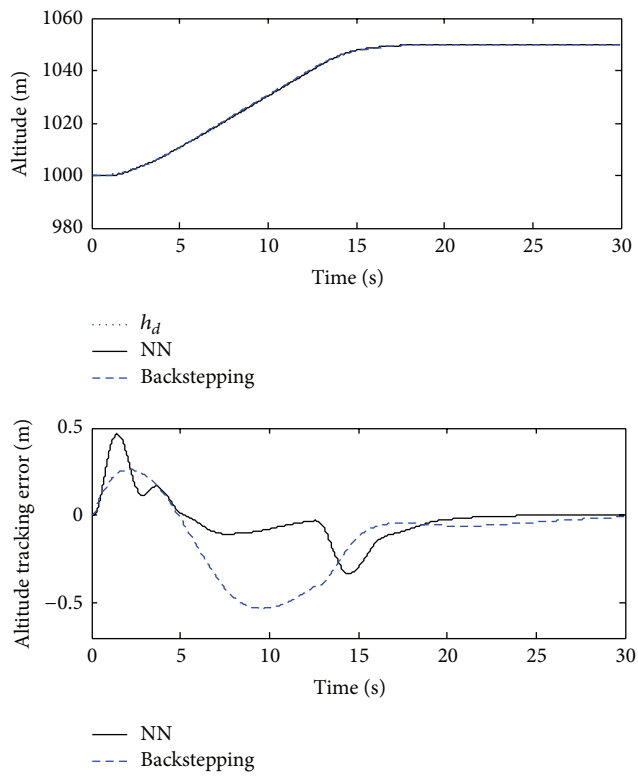


Figure 2: Altitude tracking.

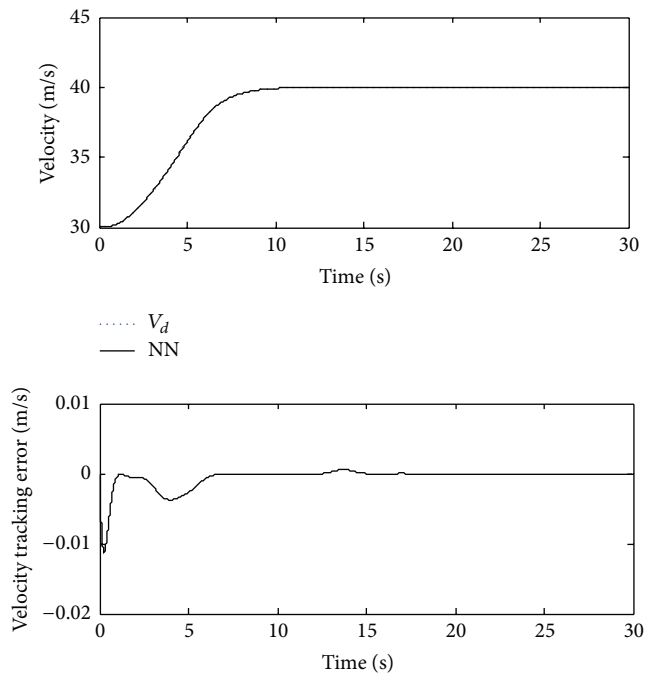


Figure 3: Velocity tracking.

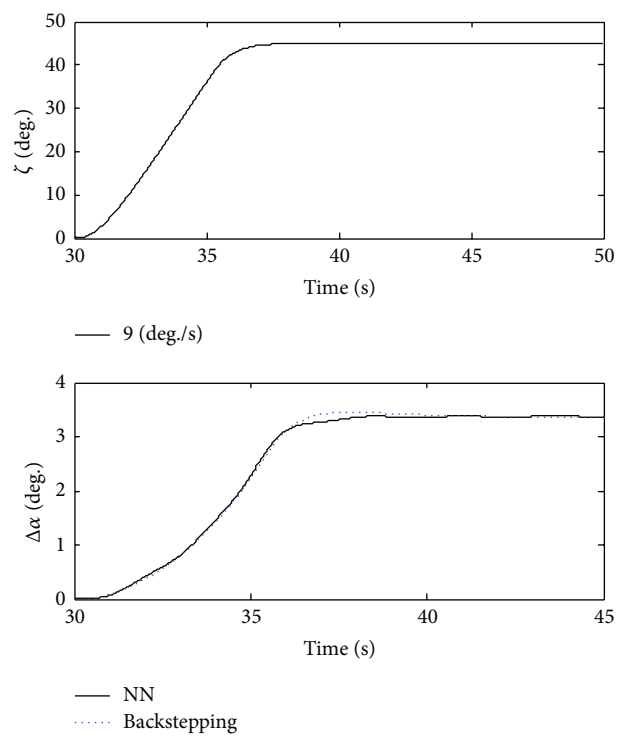


Figure 4: Sweep signal and angle of attack.

Due to the complex nonlinear aerodynamic of the morphing aircraft, the aerodynamics is not modeled precisely, the same as it appears in the actual flight conditions. Thus it is significant for the controller to have the ability to provide stability in spite of modeling errors due to unmodeled dynamics and plant parameter variations. To demonstrate the robustness of the proposed control scheme, 20% aerodynamic uncertainties are taken into account. The following two scenario simulations are employed to test the performance of the proposed controller in handling with aerodynamic uncertainty and input constraints compared with backstepping controller designed in the altitude subsystem.

Scenario 1. (A) The altitude h_d and velocity V_d reference commands are generated to make the aircraft climb from 1000 m to 1050 m and accelerate from 30 m/s to 40 m/s in 20 s, where the engine speed saturation is not considered. The simulation results of the tracking output are shown in Figures 2 and 3 ("NN" denotes the simulation results based on adaptive NN controller in this paper and "backstepping" represents the backstepping method in [25]). It can be observed that the system outputs h and V on the basis of NN and backstepping follow the desired trajectory of h_d and V_d well. The altitude tracking error of NN is smaller than the one based on backstepping. These simulation results show that good tracking performance can be obtained under the proposed adaptive NN control.

(B) To illustrate the effectiveness of the proposed adaptive NN control further, the sweep reference signal taking place at 30 s is generated to make the aircraft sweep from 0° to 45° at the rate of 9°/s. The simulation results are shown in Figures 4–7. It is clear that the velocity is almost constant, during the sweeping process, and the altitude which decreases about 0.32 m based on adaptive NN which is better than backstepping method decreases about 1.75 m. They can both

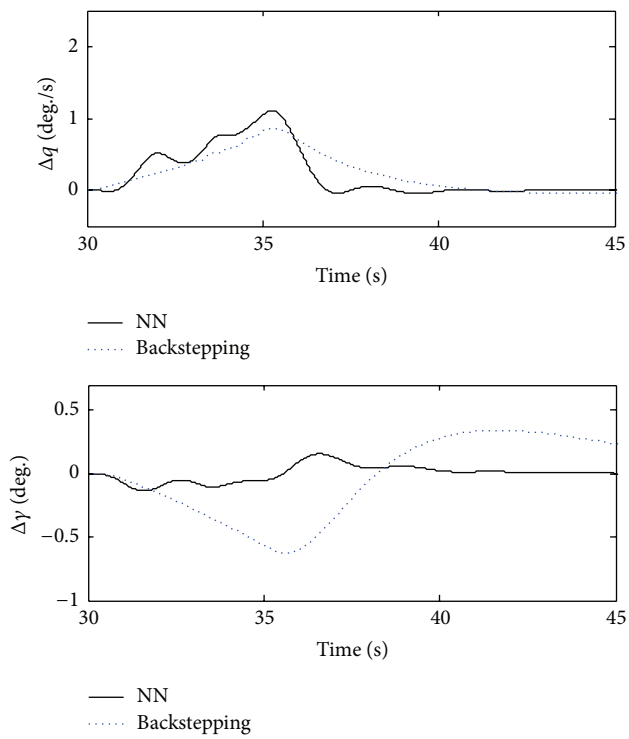


Figure 5: Pitch rate and FPA.

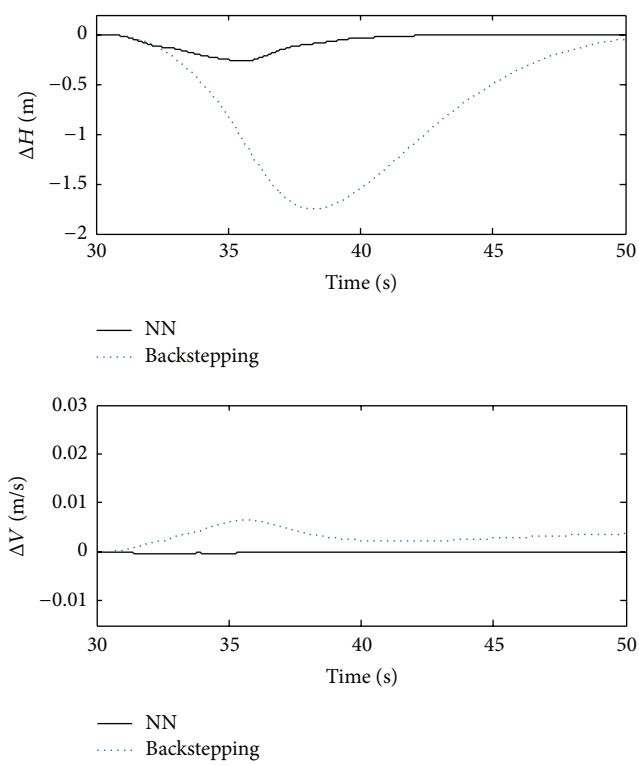


Figure 6: Altitude and velocity variation.

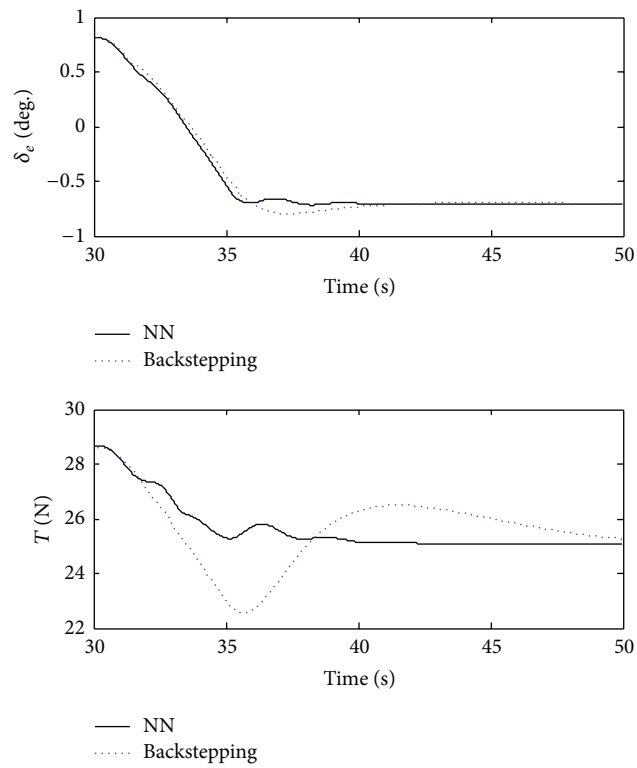


Figure 7: Elevator deflection and thrust.

converge within 20 s after the wing finishes sweeping. Since the wing area decreases after it sweeps, the angle of attack will increase to achieve a new trim point. In addition, the changes in elevator deflection and thrust are both within acceptable ranges. It can be concluded that the adaptive neural controller, in this paper, can accommodate different wing shapes that result in drastically changing plant dynamic and guarantee the flight more steady compared with backstepping method.

Scenario 2 (engine speed saturation). To illustrate the effectiveness of the auxiliary system, the reference commands are similar to Scenario 1(A), and the engine speed saturation $n_{\max}$ is set at 4900 RPM. The simulation results are shown in Figures 8 and 9. Due to engine speed saturation, it is obvious to observe that the velocity tracking errors are different between the used (V_1 with 0.5 m/s to the maximum) and unused (V_2 with 0.8 m/s to the maximum) additional system. As shown in Figure 9, the engine speed recovers from saturation in 9 s for V_1 which is better than V_2 which recovers in 15 s. These simulation results show that good tracking performance can be obtained under the proposed additional system.

6. Conclusions and Future Works

A robust adaptive neural controller based on high order integral chained differentiator is developed for the nonlinear longitudinal model of a morphing aircraft, where aerodynamic uncertainty and engine input constraint are taken into

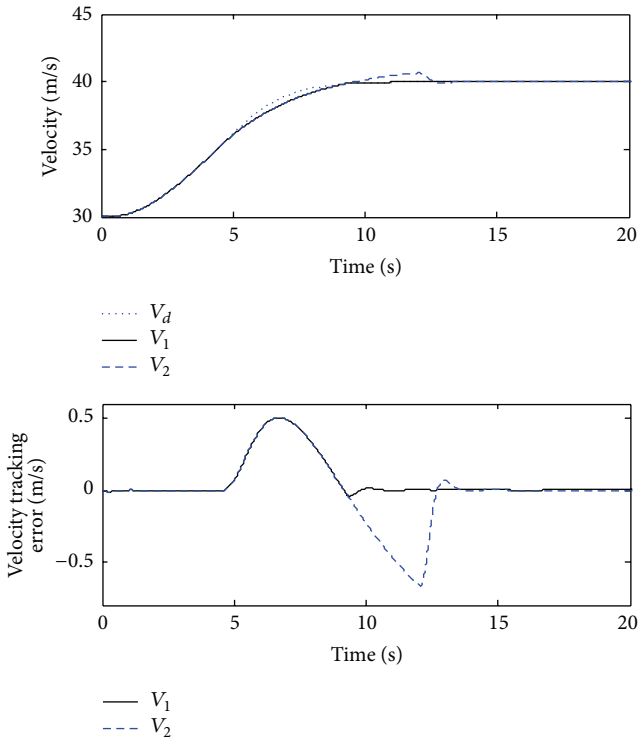


Figure 8: Velocity tracking error with input saturation.

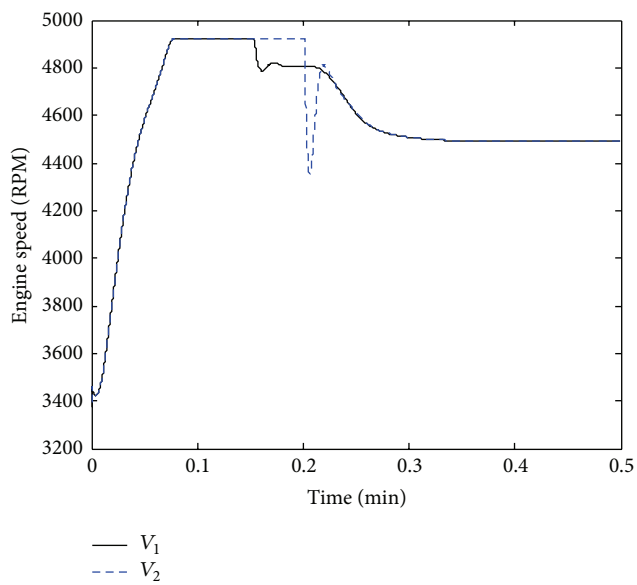


Figure 9: Engine speed.

consideration. The altitude controller is viewed as the output-feedback control problem with one NN to approximate the lumped uncertain nonlinearity while another adaptive NN controller is designed for the velocity subsystem. The problem posed by engine input constraints is overcome by additional systems. The filtered tracking error is proved to be guaranteed zero semiglobally and all the signals are uniformly bounded. The performance of the presented method is verified by simulations, from which we can deduce that the good performance has been ensured.

For future work, we will analyze how minimal parameter learning technique can be implemented on morphing aircraft in order to reduce the computation burden further. Also it is important to do research on theoretical analysis deeply for the system with input nonlinearity and time-delay where it is still an open problem for this scheme.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to express their sincere thanks to anonymous reviewers for their helpful suggestions for improving the technique note. This work is partially supported by the Natural Science Foundation of China (Grant no. 61374032) and Aeronautical Science Foundation of China (Grant no. 20140753012).

References

[1] J. Bowman, B. Sanders, B. Cannon, J. Kudva, and S. Joshi, "Development of next generation morphing aircraft structures," in *Proceedings of the 48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Honolulu, Hawaii, USA, April 2007.

[2] A. R. Rodriguez, "Morphing aircraft technology survey," in *Proceedings of the 5th AIAA Aerospace Sciences Meeting and Exhibit*, Reno, Nev, USA, January 2007.

[3] L. Yuping and H. Zhen, "A survey of morphing aircraft control systems," *Acta Aeronautica et Astronautica Sinica*, vol. 30, no. 10, pp. 1906–1911, 2009.

[4] T. M. Seigler, *Dynamics and control of morphing aircraft [Ph.D. thesis]*, 2005.

[5] T. M. Seigler, D. A. Neal, J.-S. Bae, and D. J. Inman, "Modeling and flight control of large-scale morphing aircraft," *Journal of Aircraft*, vol. 44, no. 4, pp. 1077–1087, 2007.

[6] N. Gandhi, A. Jha, J. Monaco, T. Seigler, D. Ward, and D. Inman, "Intelligent control of a morphing aircraft," in *Proceedings of the 46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, Honolulu, Hawaii, USA, April 2007.

[7] S. Barbarino, O. Bilgen, R. M. Ajaj, M. I. Friswell, and D. J. Inman, "A review of morphing aircraft," *Journal of Intelligent Material Systems and Structures*, vol. 22, no. 9, pp. 823–877, 2011.

[8] R. Shi and W. Wan, "Analysis of flight dynamics for large-scale morphing aircraft," *Aircraft Engineering and Aerospace Technology*, vol. 87, no. 1, pp. 38–44, 2015.

[9] B. O'Grady, *Multi-objective optimization of a three cell morphing wing substructure [Ph.D. thesis]*, University of Dayton, Dayton, Ohio, USA, 2010.

[10] T. Johnson, M. Frecker, M. Abdalla, Z. Gurdal, and D. Lindner, "Nonlinear analysis and optimization of diamond cell morphing wings," *Journal of Intelligent Material Systems and Structures*, vol. 20, no. 7, pp. 815–824, 2009.

[11] J. J. Joo and B. Sanders, "Optimal location of distributed actuators within an in-plane multi-cell morphing mechanism," *Journal of Intelligent Material Systems and Structures*, vol. 20, no. 4, pp. 481–492, 2009.

[12] G. R. Andersen, D. L. Cowan, and D. J. Piatak, "Aeroelastic modeling, analysis and testing of a morphing wing structure," in *Proceedings of the 48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, pp. 359–373, Honolulu, Hawaii, USA, April 2007.

[13] T. M. Seigler, D. A. Neal, and D. J. Inman, "Dynamic modeling of large-scale morphing aircraft," in *Proceedings of the 47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, pp. 3668–3678, Newport, RI, USA, May 2006.

[14] B. Obradovic and K. Subbarao, "Modeling of flight dynamics of morphing-wing aircraft," *Journal of Aircraft*, vol. 48, no. 2, pp. 391–402, 2011.

[15] T. Yue, L. Wang, and J. Ai, "Longitudinal linear parameter varying modeling and simulation of morphing aircraft," *Journal of Aircraft*, vol. 50, no. 6, pp. 1673–1681, 2013.

[16] K. Boothe, K. Fitzpatrick, and R. Lind, "Controllers for disturbance rejection for a linear input-varying class of morphing aircraft," in *Proceedings of the 46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics & Materials Conference*, Austin, Tex, USA, April 2005.

[17] S. Beaverstock, R. A. M. Friswell, W. Dettmer, R. de Breuker, and N. Werter, "Effect of span-morphing on the longitudinal flight stability and control," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Boston, Mass, USA, August 2013.

[18] T. Yue, L. Wang, and J. Ai, "Gain self-scheduled H_∞ control for morphing aircraft in the wing transition process based on an LPV model," *Chinese Journal of Aeronautics*, vol. 26, no. 4, pp. 909–917, 2013.

[19] D. H. Baldelli, D.-H. Lee, R. S. Sánchez Peñal, and B. Cannon, "Modeling and control of an aeroelastic morphing vehicle," *Journal of Guidance, Control, and Dynamics*, vol. 31, no. 6, pp. 1687–1699, 2008.

[20] W. G. Nobleheart, S. L. Geethalakshmi, A. Chakravarthy, and J. Steck, "Single Network Adaptive Critic (SNAC) architecture for optimal tracking control of a morphing aircraft during a pull-up maneuver," in *Proceedings of the AIAA Guidance, Navigation, and Control (GNC) Conference*, Boston Mass, USA, August 2013.

[21] J. Valasek, A. Lampton, and M. Marwaha, "Morphing unmanned air vehicle intelligent shape and flight control," in *Proceedings of the AIAA Infotech@ Aerospace Conference and AIAA Unmanned... Unlimited Conference*, Seattle, Wash, USA, April 2009.

[22] Q. Zong, F. Wang, B. Tian, and R. Su, "Robust adaptive dynamic surface control design for a flexible air-breathing hypersonic vehicle with input constraints and uncertainty," *Nonlinear Dynamics*, vol. 78, no. 1, pp. 289–315, 2014.

[23] W. Chenliang and L. Yan, "Adaptive dynamic surface control for linear multivariable systems," *Automatica*, vol. 46, no. 10, pp. 1703–1711, 2010.

[24] Y.-J. Liu, C. L. P. Chen, G.-X. Wen, and S. Tong, "Adaptive neural output feedback tracking control for a class of uncertain discrete-time nonlinear systems," *IEEE Transactions on Neural Networks*, vol. 22, no. 7, pp. 1162–1167, 2011.

[25] M. Chen, G. Tao, and B. Jiang, "Dynamic surface control using neural networks for a class of uncertain nonlinear systems with input saturation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 9, pp. 2086–2097, 2014.

[26] S. Luo, S. Wu, Z. Liu, and H. Guan, "Wheeled mobile robot RBFNN dynamic surface control based on disturbance observer," *ISRN Applied Mathematics*, vol. 2014, Article ID 634936, 9 pages, 2014.

[27] G. J. Ducard, *Fault-tolerant Flight Control and Guidance Systems: Practical Methods for Small Unmanned Aerial Vehicles*, Springer, 2009.

[28] J. Liu and X. Wang, *Advanced Sliding Mode Control for Mechanical Systems*, Springer, 2012.

[29] B. Xu, D. X. Gao, and S. X. Wang, "Adaptive neural control based on HGO for hypersonic flight vehicles," *Science China Information Sciences*, vol. 54, no. 3, pp. 511–520, 2011.

[30] J.-H. Park, S.-H. Kim, and C.-J. Moon, "Adaptive neural control for strict-feedback nonlinear systems without backstepping," *IEEE Transactions on Neural Networks*, vol. 20, no. 7, pp. 1204–1209, 2009.

[31] S. N. Huang, K. K. Tan, and T. H. Lee, "Further results on adaptive control for a class of nonlinear systems using neural networks," *IEEE Transactions on Neural Networks*, vol. 14, no. 3, pp. 719–721, 2003.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 413072, 14 pages
http://dx.doi.org/10.1155/2015/413072

Research Article

Optimal Investment and Consumption for an Insurer with High-Watermark Performance Fee

Lin Xu,[1] Hao Wang,[1] and Dingjun Yao[2]

[1] *School of Mathematics and Computer Science, Anhui Normal University, Wuhu, Anhui 241000, China*
[2] *School of Finance, Nanjing University of Finance and Economics, Nanjing, Jiangsu 210023, China*

Correspondence should be addressed to Lin Xu; xulinahnu@gmail.com

Received 18 August 2015; Accepted 5 October 2015

Academic Editor: Xinguang Zhang

The optimal investment and consumption problem is investigated for an insurance company, which is subject to the payment of high-watermark fee from profit. The objective of insurance company is to maximize the expected cumulated discount utility up to ruin time. The consumption behavior considered in this paper can be viewed as dividend payment of the insurance company. It turns out that the value function of the proposed problem is the viscosity solution to the associated HJB equation. The regularity of the viscosity is discussed and some asymptotic results are provided. With the help of the smooth properties of viscosity solutions, we complete the verification theorem of the optimal control policies and the potential applications of the main result are discussed.

1. Introduction

Investment and consumption are of great importance in the study of finance and financial engineering. This is due to the fact that investment and consumption not only are the key topic of financial agents but also provide idea and method of deriving equilibrium price of financial derivatives (cf. Shreve and Soner [1]). Applications of stochastic optimal control to management and financial problems were developed from the 1970s, especially after the papers by Merton [2, 3] on portfolio selection. The model and results of Merton were then extended by many authors; for example, see Zariphopoulou [4], Øksendal and Sulem [5], and Fleming and Pang [6]. These problems are also studied in the monograph by Karatzas and Shreve [7]. The decision makers associated with optimal investment and consumption problems that appeared in aforementioned papers stand on the perspective of financial firms or investment bank, and the business income of decision makers comes from proper construction of portfolio positions. Usually, it assumes that the financial market is frictionless: no transaction cost and no tax payment, money that can be infinitely divided, no restrictions on short or long positions, and so forth. More recently, there is also a large, more recent literature related to the investment in markets with frictions. A transaction cost is a typical example. For example, Davis and Norman [8] studied portfolio selection problem with transaction cost, which uses the variance of the portfolio positrons as the risk measure; Janeček and Sîrbu [9] studied the future trading problem with transaction cost. Shreve and Soner [1] investigated optimal consumption and investment with transaction cost; the optimization goal therein is to maximize expected cumulated discounted utility in an infinite time horizon. Whalley and Wilmott [10] studied optimal hedging model with transaction costs. Previously mentioned papers are just a few examples of the growing literature on the topic; for more progress on this topic, readers are also referred to the works of Cvitanić and Karatzas [11], Liu and Loewenstein [12], Korn [13], and Obizhaeva and Wang [14].

Among all kinds of transaction cost, the high gain tax payment or high-watermark fee has attracted many attentions recently. The high-watermark fee is taken as the following rule: whenever the maximum up to today, the so-called high-watermark, exceeds the previously attained historic maximum, the fixed proportion of the profit (relative to the previous maximum) is charged by the fund manager. In the early 1980s, Stiglitz [15] discussed the possibility and necessity of charging high gain tax from investment income. The past two decades have witnessed an increasing attention to the

research of optimal control problem with high-watermark fee. For example, Dammon et al. [16] investigated optimal investment and consumption problem with capital gain taxes, Goetzmann et al. [17] studied the hedge fund management when charging high gain tax, and Guasoni and Wang [18] studied high-watermark and separation of private problems. As it was shown previously, investment and consumption problem is an important topic for insurer and also a key topic of insurance mathematics or financial mathematics. Thus, it is natural to consider optimal investment and consumption problem for an insurer when high-watermark fees are charged. Up to now, seldom insurance company considers consumption problems. Thus, to discuss optimal consumption for an insurer seems weird. However, an analogue problem in insurance company is the dividend payment or pension payment problem. For example, Højgaard and Taksar [19] studied reinsurance and dividend with transaction costs, Cairns [20] studied optimal pension fund schedule problem, Zhu [21] took both investment and dividend into account in searching for optimal policies, and He and Liang [22] investigated pension schedule and asset allocation problem. For other works on investment and consumption problem related to insurance affair or partially related to insurance affairs, see Bielecki and Pliska [23], Dai and Yi [24], and Young [25].

In the paper of Janeček and Sîrbu [9], the optimal investment and consumption problems for a fund manager on infinite time horizon are considered when the fund manager is subject to the high-watermark fee from investment. However, the model considered there is not suitable for an insurance company because the latter one has premium income and claims with addition to the investment profit (or underlying risk). This paper contributes to bridging this gap. The surplus process of the insurer is specified by a classical risk model and the insurer has the chance to invest into risky asset and risk-free bond market. Whenever the profit of the insurer attains a new maximum, the high-watermark fee is taken as a kind of gain tax. The goal of the insurer is to choose optimal investment and consumption policies before ruin occurs. We will point out that although it seems weird to allow the insurance company to make consumption policies, dividend payment is a very common decision policy for the managers of insurance company. Thus, the consumption framework considered in this paper can be regarded as a kind of dividend payment of insurer. The progress achieved in this paper can be summarized as follows. The optimization problem considered in this paper is relevant to a jump diffusion process. Thus, the associated HJB equation contains an integration part, which brings us some difficulties in proving the smooth properties of the solution to HJB equation. Similarly, it brings us difficulties in proving that the value function is the viscosity solution to HJB equation with integration part. Following the idea presented in Janeček and Sîrbu [9], we prove that the value function is smooth on its domain. We obtain the value function of an insurer without high-watermark fee, which is not considered in other literature. We obtain a verification theorem, which means that the viscosity solution to HJB equation is indeed the value function. Due to the natural connections between the viscosity solution to HJB equation and the numerical algorithm to the stochastic control problem, the main result obtained in this paper is useful in the design of the numerical approximating method of the related HJB equation.

The rest of this paper is organized as follows. In Section 2, the model and problem are presented and efforts are made to transform the model such that dynamic programming principle and HJB equation method are applicable. In Section 3, the definition of viscosity solution to a kind of second-order partial integrodifferential equation is given, and the value function is proved to be the viscosity solution of the associated HJB equation. In Section 4, by employing the same method of Janeček and Sîrbu [9], the viscosity solution is proved to be smooth on certain domain. The properties of viscosity on some singular point are also discussed. Section 5 presents a verification theorem, which asserts that the solution to HJB equation is indeed the value function and the corresponding feedback control replicates the optimal realization of the insurer. Section 6 concludes the main contributions of this paper and potential applications of our results.

2. Model and Problem

2.1. Classical Risk Model and Its Diffusion Approximation. In this subsection, we briefly introduce the classical risk model of an insurance company and its diffusion approximation. The content presented here will be helpful for our later investigation. Classical risk model for an insurer is (cf. Grandell [26])

$$dU_t = adt - d\left(\sum_{i=1}^{N_t} Y_i\right), \tag{1}$$

where a is the constant premium income rate and $\{N_t;\ t \geq 0\}$ denotes the number of claims that arrived up to time t, which is assumed to be a homogeneous Poisson process with intensity $\lambda_0 > 0$. The individual claims $Y_1, Y_2, \ldots$ are assumed to be a sequence of independent and identically distributed (i.i.d.) positive random variables (r.v.s.) with common distribution function $G(y)$ and finite expectation and satisfy $G(0) = 0$. In addition, it assumes that $\{N_t;\ t \geq 0\}$ and $\{Y_i;\ i = 1, 2, \ldots\}$ are mutually independent. For notation convenience, we denote by $\{S_t;\ t \geq 0\}$ the aggregate claim process; that is, $S_t = \sum_{i=1}^{N_t} Y_i$. We denote that $\mathbb{E}Y_i = \mu$ and $\mathrm{Var}(Y_i) = \sigma^2$. Main topic associated with classical risk model is the ruin probability; in mathematics, it is $\mathbb{P}(\tau < \infty)$, where $\tau = \inf\{t \geq 0, U_t < 0\}$ is known as "ruin time" in risk theory. There are many methods to study the ruin time and ruin probability, such as renewal method prompted by Feller and martingale approach introduced by Gerber (cf. Grandell [26]). Another idea is to approximate the classical risk model by some stochastic process with good statistical properties, such as Gaussian process. This is so-called diffusion approximation of classical risk model; see Chapter 1 of Grandell [26].

2.2. A General Model of Profits from Dynamic Investment in a Hedge Fund. Now, suppose that the insurance company invests in a risky fund with a share or unit price F_t at time t. If the insurance company chooses to hold θ_t capital in the

fund at time t and no fees of any kind are imposed, then the accumulated profit at time t, denoted by P_t, evolves as

$$dP_t = \theta_t \frac{dF_t}{F_t} + adt - d\left(\sum_{i=1}^{N(t)} Y_i\right) = \theta_t \frac{dF_t}{F_t} + adt - dS_t, \quad 0 \le t < \infty, \qquad P_0 = 0. \tag{2}$$

Remark 1. The assumption $P_0 = 0$ seems unreasonable from practice; however, we want to compare our model with the model studied in Janeček and Sîrbu [9], so we made such an assumption. In later discussion, the initial surplus of the insurer is assumed to be $x > 0$.

Denote by $\{M_t, t \ge 0\}$ the maximum profit process; that is,

$$M_t \triangleq \sup_{0 \le s \le t} P_s. \tag{3}$$

Assume now that the manager tracks the high-watermark fee once the insurance company achieves new maximum of profit; the rule is as follows: anytime the high-watermark increases, $\lambda > 0$ percentage of this increase is paid to the fund manager. More precisely, the insurance company pays $\lambda \Delta M_t = \lambda(M_{t+\Delta t} - M_t)$ to the manager in the interval $[t, t + \Delta t]$. Under such a high-watermark fee taking rule, the evolution equation for the profit P_t is revised as

$$\begin{aligned} dP_t &= \theta_{t-} \frac{dF_t}{F_t} + adt - dS_t - \lambda dM_t, \quad P_0 = 0, \\ M_t &= \sup_{0 \le s \le t} P_s. \end{aligned} \tag{4}$$

Suppose that the insurance company has an initial maximum profit i ($i \ge 0$); the profits of the insurance company will be taxed when P reaches value i and will not be taxed before P reaches at least value i. Then, for any given $i \ge 0$, the dynamic of P is given by

$$\begin{aligned} dP_t &= \theta_{t-} \frac{dF_t}{F_t} + adt - dS_t - \lambda dM_t, \quad P_0 = 0, \\ M_t &= \sup_{0 \le s \le t} (P_s \vee i). \end{aligned} \tag{5}$$

A similar representation appears in the appendix of Guasoni and Wang [18], where an optimization problem related to maximizing utility of the fund manager is studied, which is opposed to the utility of the investor in our case. However, their state equation is similar to (5), so we resort to the same pathwise representation.

Proposition 2. *Assume that the share/unit prices process F is a continuous and strictly positive semimartingale, and the predictable processes $\{\theta_t; t \ge 0\}$ are such that the accumulated profit process corresponding to the trading strategy θ, in case no profit fees are imposed, namely,*

$$I_t = at - S_t + \int_0^t \theta_{u-} \frac{dF_u}{F_u}, \quad 0 \le t < \infty, \tag{6}$$

is well defined. Then (5) has a unique solution, which can be represented pathwise by

$$P_t = I_t - \frac{\lambda}{1+\lambda} \sup_{0 \le s \le t} [I_s - i]^+, \quad 0 \le t < \infty, \tag{7}$$

$$M_t = i + \frac{1}{1+\lambda} \max_{0 \le s \le t} [I_s - i]^+, \quad 0 \le t < \infty. \tag{8}$$

Proof. Note that $P_0 = I_0 = 0$, $M_0 = i$; (5) can be rewritten as

$$(P_t - i) + \lambda \sup_{0 \le s \le t} [P_s - i]^+ = (I_t - i), \quad 0 \le t < \infty. \tag{9}$$

Taking the positive part and the supremum on both sides, it follows that

$$\begin{aligned} (1+\lambda)(M_t - i) &= (1+\lambda) \sup_{0 \le s \le t} [P_s - i]^+ \\ &= \sup_{0 \le s \le t} [I_s - i]^+, \quad 0 \le t < \infty. \end{aligned} \tag{10}$$

Replacing (10) into (5), we finish the proof of uniqueness.

By checking that the process in (7) is a solution of (5), more precisely,

$$\begin{aligned} dP_t &= d\left(I_t - \frac{\lambda}{1+\lambda} \sup_{0 \le s \le t} [I_s - i]^+\right) \\ &= dI_t - \frac{\lambda}{1+\lambda} d\left(\sup_{0 \le s \le t} [I_s - i]^+\right) \\ &= \theta_{t-} \frac{dF_t}{F_t} + at - S_t - \lambda dM_t. \end{aligned} \tag{11}$$

This completes the proof. □

2.3. Optimal Investment and Consumption in a Special Model. Assume that the insurance company starts with initial capital $x > 0$ and only additional investment opportunity is the money market paying zero interest rate. The insurance company is given the intimal high-watermark $i \ge 0$ for its profits. We assume that the insurance company consumes at a rate $\gamma_t > 0$ per unit of time. Consumption can be made either from the money market account or from accumulated profit. Denote by

$$C_t \triangleq \int_0^t \gamma_s ds, \quad 0 \le t < \infty, \tag{12}$$

the accumulated consumption process and by $X_t^{\theta,\gamma}$ the wealth process of an insurer associated with decision policy (θ, γ). Since the money market pays zero interest rate, the wealth $X_t^{\theta,\gamma}$ is formulated as

$$X_t^{\theta,\gamma} = x + P_t - C_t, \quad 0 \le t \le \tau^{\theta,\gamma}, \tag{13}$$

where $\tau^{\theta,\gamma} \triangleq \inf\{s \ge 0 : X_s^{\theta,\gamma} \le 0\}$ is the first time that the wealth goes below zero. In actuarial theory, τ is referred to as the "ruin time." In later discussion, for notation ease, we drop the symbol θ, γ in $\tau^{\theta,\gamma}$.

If consumption is taken into account, the high-watermark of the insurance company's profit can be represented as

$$\begin{aligned} M_t^{\theta,\gamma} &= \sup_{0\le s\le t}\left[\left(X_s^{\theta,\gamma}+C_s-x\right)\vee i\right]^+ \\ &= i+\sup_{0\le s\le t}\left[\left(X_s^{\theta,\gamma}+C_s\right)-k\right]^+, \end{aligned} \tag{14}$$

where $k \triangleq x+i \ge x > 0$. In this situation, wealth evolves as

$$\begin{aligned} dX_t^{\theta,\gamma} &= \theta_{t-}\frac{dF_t}{F_t}+adt-\gamma_t dt-dS_t-\lambda dM_t^{\theta,\gamma}, \\ &\qquad X_0^{\theta,\gamma}=x, \\ M_t^{\theta,\gamma} &= i+\sup_{0\le s\le t}\left[\left(X_s^{\theta,\gamma}+\int_0^s\gamma_u\,du\right)-k\right]^+, \\ &\qquad 0\le t\le\tau. \end{aligned} \tag{15}$$

So far, this is a general model of investment/consumption in a hedge fund, which is also a good model of taxation. In what follows, we focus on a simple but important case, where the fund share/unit price $\{F_t,\ t\ge 0\}$ evolves as a geometric Brownian motion; that is,

$$\frac{dF_t}{F_t}=\alpha dt+\sigma dW_t,\quad 0\le t<\infty, \tag{16}$$

where $(W_t)_{0\le t<\infty}$ is a standard Brownian motion defined on the filtered probability space $(\Omega,\mathscr{F},(\mathscr{F}_t)_{0\le t<\infty},\mathbb{P})$. With this notation, (15) becomes

$$\begin{aligned} dX_t^{\theta,\gamma} &= \left(\theta_t\alpha-\gamma_t+a\right)dt+\theta_t\sigma dW_t-dS_t-\lambda dM_t^{\theta,\gamma}, \\ &\qquad X_0^{\theta,\gamma}=x, \\ M_t^{\theta,\gamma} &= i+\sup_{0\le s\le t}\left[\left(X_s^{\theta,\gamma}+\int_0^s\gamma_u du\right)-k\right]^+, \\ &\qquad 0\le t\le\tau. \end{aligned} \tag{17}$$

In order to use dynamic programming, we want to represent the control problem using a state process of minimal dimension. What is more, since we want to apply the HJB equation method, it is necessary to embed our state process into a Markovian system. As usual, the wealth X has to be a part of the state. But using (X,M) as state is not a possibility, since M does not contain the information on past consumption, just that on past profits. Copying the method of Janeček and Sîrbu [9], we observed that the fee is being paid as soon as the current profit $P_t^{\theta,\gamma}=X_t^{\theta,\gamma}+C_t-x$ (current wealth plus accumulated consumption plus aggregate claim minus income minus initial wealth) hits the high-watermark $M_t^{\theta,\gamma}=i+\sup_{0\le s\le t}[X_s+C_s-k]^+$. In other words, fees are paid whenever

$$X_t^{\theta,\gamma}+C_t-k=\sup_{0\le s\le t}\left[\left(X_s+C_s\right)-k\right]^+, \tag{18}$$

which is the same as $X^{\theta,\gamma}=K^{\theta,\gamma}$ for

$$\begin{aligned} K_t^{\theta,\gamma} &\triangleq k+\sup_{0\le s\le t}\left[\left(X_s^{\theta,\gamma}+C_s\right)-k\right]^+-C_t \\ &= \sup_{0\le s\le t}\left[\left\{X_s^{\theta,\gamma}+C_s\right\}\vee k\right]-C_t\ge X_t^{\theta,\gamma}. \end{aligned} \tag{19}$$

We now choose as state process the two-dimensional process (X,K) which satisfies $X\le K$ and is reflected whenever $X=K$. The controlled state process (X,K) follows the evolution

$$\begin{aligned} dX_t^{\theta,\gamma} &= \left(\theta_t\alpha-\gamma_t+a\right)dt+\theta_t\sigma dW_t-dS_t \\ &\quad -\lambda\left(dK_t^{\theta,\gamma}+\gamma_t dt\right),\quad X_0=x, \\ K_t^{\theta,\gamma} &= \sup_{0\le s\le t}\left[\left\{X_s^{\theta,\gamma}+\int_0^s\gamma_u du\right\}\vee k\right]-\int_0^t\gamma_u du, \\ &\qquad 0\le t\le\tau. \end{aligned} \tag{20}$$

Equation (20) is implicit, as is (5). The pathwise representation in Proposition 2 can be easily translated into a pathwise solution (20). More precisely, we have Proposition 3, and the proof of Proposition 3 is similar to Proposition 2; we omit it here.

Proposition 3. *Assume that the predictable process (θ,γ) satisfies*

$$\mathbb{P}\left(\int_0^t\left(\left|\theta_u\right|^2+\gamma_u\right)du<\infty,\ \forall 0\le t\le\tau\right)=1. \tag{21}$$

Denote

$$\begin{aligned} Z_t &= \int_0^t\theta_u\left(\alpha du+\sigma dW_u\right)+at-S_t, \\ C_t &= \int_0^t\gamma_u du,\quad 0\le t\le\tau. \end{aligned} \tag{22}$$

Suppose that the accumulated profit process corresponding to the trading strategy θ is well defined. Then

$$\begin{aligned} X_t^{\theta,\gamma} &= x+Z_t-C_t-\frac{\lambda}{1+\lambda}\sup_{0\le s\le t}\left[Z_s-i\right]^+, \\ &\qquad 0\le t\le\tau, \\ K_t^{\theta,\gamma} &= k+\frac{1}{1+\lambda}\sup_{0\le s\le t}\left[Z_s-i\right]^+-C_t,\quad 0\le t\le\tau. \end{aligned} \tag{23}$$

The high-watermark is computed as

$$M_t^{\theta,\gamma}=i+\frac{1}{1+\lambda}\sup_{0\le s\le t}\left[Z_s-i\right]^+,\quad 0\le t\le\tau. \tag{24}$$

Fix an initial capital $x>0$ and an initial high-watermark of profits $i\ge 0$. Recall that $k\triangleq x+i\ge x$. An investment/consumption strategy (θ,γ) is called admissible with respect to the initial date (x,k) if it satisfies integrability conditions (21); the consumption stream is positive ($\gamma_t\ge 0$). We denote by $\mathscr{A}(x,k)$ the set of all admissible strategies at (x,k).

We consider a concave utility function $U : (0,\infty) \to \mathbb{R}$. So we can define the expected utility from consumption $\mathbb{E}[\int_0^\tau e^{-\beta t}U(\gamma_t)dt]$. The object of this paper is to research the optimal investment/consumption amounts (θ,γ) related to following optimization problem (for each fixed (x,k)):

$$v(x,k) \triangleq \sup_{(\theta,\gamma)\in\mathscr{A}(x,k)} \mathbb{E}\left[\int_0^\tau e^{-\beta t}U(\gamma_t)\,dt\right], \quad 0 \le x \le k. \tag{25}$$

One should note that once the initial surplus is negative, that is, $x < 0$, it immediately followed by

$$v(x,k) = 0, \quad \forall x < 0. \tag{26}$$

Function v defined above is called the value function. We further assume that the utility function U has the particular form

$$U(\gamma) = \frac{\gamma^{1-p}}{1-p}, \quad \gamma > 0, \tag{27}$$

for some $p > 0$, $p \neq 1$, where p is called the relative risk aversion coefficient.

Using the controls (θ,γ) the insurance company controls the process (X,K) in (21) which is restricted to the domain $0 < x \le k$ and is reflected on the diagonal $x = k$ in the direction given by the vector

$$\tilde{r} \triangleq \begin{pmatrix} -\lambda \\ 1 \end{pmatrix}. \tag{28}$$

So, state (20) can be rewritten as

$$d\begin{pmatrix} X_t^{\theta,\gamma} \\ K_t^{\theta,\gamma} \end{pmatrix} = \begin{pmatrix} (\theta_t\alpha - \gamma_t + a)\,dt + \theta_t\sigma dW_t - dS_t \\ -\gamma_t dt \end{pmatrix} + \tilde{\gamma} dM_t^{\theta,\gamma}, \quad 0 \le t \le \tau, \tag{29}$$

where

$$\int_0^t 1_{\{X_s^{\theta,\gamma} \neq K_s^{\theta,\gamma}\}} dM_s = 0. \tag{30}$$

Denote the continuous part of process $(X_t^{\theta,\gamma}, K_t^{\theta,\gamma})$ by $(X_t^{c,\theta,\gamma}, K_t^{c,\theta,\gamma})$; that is,

$$d\begin{pmatrix} X_t^{c,\theta,\gamma} \\ K_t^{c,\theta,\gamma} \end{pmatrix} = \begin{pmatrix} (\theta_t\alpha - \gamma_t + a)\,dt + \theta_t\sigma dW_t \\ -\gamma_t dt \end{pmatrix} + \tilde{\gamma} dM_t, \theta, \gamma, \quad 0 \le t \le \tau. \tag{31}$$

The main goal of the present paper is to analyze the impact of fees on the investment/consumption strategies and the main method in this paper relies on HJB equation. One should note that, with the introduction of process $K_t^{\theta,\gamma}$, we embed our model into a Markov system, which enables further discussion.

3. Dynamic Programming and HJB Equation

Now, in order to obtain the HJB equation as follows, we will use the dynamic programming principle; see Proposition 4. The proof of Proposition 4 is similar to the one in Azcue and Muler [27] and we omit the proof here.

Proposition 4. *Suppose that $(\hat{\theta},\hat{\gamma})$ is an optimal control. Then one has*

$$\begin{aligned} v(x,k) &= \mathbb{E}\left[\int_0^{h\wedge\tau} e^{-\beta s}U(\hat{\gamma}_s)\,ds + e^{-\beta(h\wedge\tau)}v\left(X_{h\wedge\tau}^{\hat{\theta},\hat{\gamma}}, K_{h\wedge\tau}^{\hat{\theta},\hat{\gamma}}\right)\right], \\ v\left(X_\tau^{\hat{\theta},\hat{\gamma}}, K_\tau^{\hat{\theta},\hat{\gamma}}\right) &= 0. \end{aligned} \tag{32}$$

If $v(x,n)$ is smooth enough, by Itô's Lemma, we have

$$\begin{aligned} e^{-\beta(h\wedge\tau)}v\left(X_{h\wedge\tau}^{\hat{\theta},\hat{\gamma}}, K_{h\wedge\tau}^{\hat{\theta},\hat{\gamma}}\right) &= v(x,k) \\ &+ \int_0^{h\wedge\tau}\left(-\beta e^{-\beta s}v\right)ds + \int_0^{h\wedge\tau} e^{-\beta s}\frac{\partial v}{\partial x}dX_s^c \\ &+ \int_0^{h\wedge\tau} e^{-\beta s}\frac{\partial v}{\partial k}dK_s^c + \frac{1}{2}\int_0^{h\wedge\tau} e^{-\beta s}\frac{\partial^2 v}{\partial x^2}dX_s^c dX_s^c + \frac{1}{2} \\ &\cdot \int_0^{h\wedge\tau} e^{-\beta s}\frac{\partial^2 v}{\partial k^2}dK_s^c dK_s^c + \frac{1}{2} \\ &\cdot \int_0^{h\wedge\tau} e^{-\beta s}\frac{\partial^2 v}{\partial x\partial k}dX_s^c dK_s^c + \int_0^{h\wedge\tau}\int_0^\infty e^{-\beta s} \\ &\cdot \left[v\left(X_{s-}^{\hat{\theta},\hat{\gamma}} - Y, K_{s-}^{\hat{\theta},\hat{\gamma}}\right) - v\left(X_{s-}^{\hat{\theta},\hat{\gamma}}, K_{s-}^{\hat{\theta},\hat{\gamma}}\right)\right]\overline{K}(dY, ds), \end{aligned} \tag{33}$$

where $\overline{N}$ is the Poisson random measure on $[0,\tau]\times[0,\infty)$ defined by

$$\overline{K} = \sum_{n\ge1}\delta(T_k, Y_k). \tag{34}$$

Denote by $\mathscr{L}^{\theta,\gamma}v(\cdot,\cdot)$ (associated with (θ,γ)) the second-order partial differential operator with the form of

$$\mathscr{L}^{\theta,\gamma}v = (\theta_t\alpha - \gamma_t + a)\,v_x + \frac{1}{2}\theta_t^2\sigma^2 v_{xx} - \gamma_t v_k, \tag{35}$$

where v_x, v_{xx}, v_k are the first, the second, and the first partial derivatives with respect to x and k, respectively. Then, by compensating (33) with

$$\lambda_0\int_0^{h\wedge\tau}\int_0^\infty e^{-\beta s}\left[v\left(X_{s-}^{\hat{\theta},\hat{\gamma}} - Y, K_{s-}^{\hat{\theta},\hat{\gamma}}\right) - v\left(X_{s-}^{\hat{\theta},\hat{\gamma}}, K_{s-}^{\hat{\theta},\hat{\gamma}}\right)\right]dG(Y)\,ds, \tag{36}$$

(33) can be rewritten as

$$\begin{aligned} e^{-\beta(h\wedge\tau)} v\left(X_{h\wedge\tau}^{\widehat{\theta},\widehat{\gamma}}, K_{h\wedge\tau}^{\widehat{\theta},\widehat{\gamma}}\right) &= v(x,k) + \int_0^{h\wedge\tau} e^{-\beta s}\left(-\beta v\right. \\ &\quad \left. + \mathscr{L}^{\widehat{\theta},\widehat{\gamma}} v\right)\left(X_s^{\widehat{\theta},\widehat{\gamma}}, K_s^{\widehat{\theta},\widehat{\gamma}}\right) ds \\ &\quad + \int_0^{h\wedge\tau} e^{-\beta s} \widehat{\theta}_s \sigma v_x\left(X_s^{\widehat{\theta},\widehat{\gamma}}, K_s^{\widehat{\theta},\widehat{\gamma}}\right) dW_s \\ &\quad + \int_0^{h\wedge\tau} e^{-\beta s}\left(-\lambda v_x + v_k\right)\left(X_s^{\widehat{\theta},\widehat{\gamma}}, K_s^{\widehat{\theta},\widehat{\gamma}}\right) dM_s \\ &\quad + \int_0^{h\wedge\tau}\int_0^{\infty} e^{-\beta s}\left[v\left(X_{s-}^{\widehat{\theta},\widehat{\gamma}} - Y, K_{s-}^{\widehat{\theta},\widehat{\gamma}}\right)\right. \\ &\quad \left. - v\left(X_{s-}^{\widehat{\theta},\widehat{\gamma}}, K_{s-}^{\theta,\gamma}\right)\right]\left[\overline{K}(dYds) - \lambda_0 dG(Y)\, ds\right] \\ &\quad + \lambda_0 \int_0^{h\wedge\tau}\int_0^{\infty} e^{-\beta s}\left[v\left(X_{s-}^{\widehat{\theta},\widehat{\gamma}} - Y, K_{s-}^{\widehat{\theta},\widehat{\gamma}}\right)\right. \\ &\quad \left. - v\left(X_{s-}^{\widehat{\theta},\widehat{\gamma}}, K_{s-}^{\widehat{\theta},\widehat{\gamma}}\right)\right] dG(Y)\, ds. \end{aligned} \tag{37}$$

Since

$$\begin{aligned} &\int_0^{h\wedge\tau}\int_0^{\infty} e^{-\beta s}\left[v\left(X_{s-}^{\widehat{\theta},\widehat{\gamma}} - Y, K_{s-}^{\widehat{\theta},\widehat{\gamma}}\right) - v\left(X_{s-}^{\widehat{\theta},\widehat{\gamma}}, K_{s-}^{\widehat{\theta},\widehat{\gamma}}\right)\right] \\ &\quad \cdot \left[\overline{K}(dY, ds) - \lambda_0 dG(Y)\, ds\right] \end{aligned} \tag{38}$$

is a martingale (see [28, page 63]), it follows that

$$\begin{aligned} 0 &= \mathbb{E}\left[\int_0^{h\wedge\tau} e^{-\beta s}\left[-\beta v\left(X_s^{\widehat{\theta},\widehat{\gamma}}, K_s^{\widehat{\theta},\widehat{\gamma}}\right) + \mathscr{L}^{\widehat{\theta},\widehat{\gamma}} v\left(X_s^{\widehat{\theta},\widehat{\gamma}}, K_s^{\widehat{\theta},\widehat{\gamma}}\right) + U(\widehat{\gamma}_s)\right] ds\right] + \mathbb{E}\left[\int_0^{h\wedge\tau} e^{-\beta s}\left(-\lambda v_x + v_k\right)\left(X_s^{\widehat{\theta},\widehat{\gamma}}, K_s^{\widehat{\theta},\widehat{\gamma}}\right) dM_s\right] \\ &\quad + \lambda_0 \mathbb{E}\left[\int_0^{h\wedge\tau}\int_0^{\infty} e^{-\beta s}\left[v\left(X_{s-}^{\widehat{\theta},\widehat{\gamma}} - Y, K_{s-}^{\widehat{\theta},\widehat{\gamma}}\right) - v\left(X_{s-}^{\widehat{\theta},\widehat{\gamma}}, K_{s-}^{\widehat{\theta},\widehat{\gamma}}\right)\right] dG(Y)\, ds\right]. \end{aligned} \tag{39}$$

With boundary condition (26), dividing proceeding equation by h and sending h to 0, we can formally write the HJB function:

$$\begin{aligned} &(\lambda_0 + \beta) v(x,k) - \sup_{(\theta,\gamma)\in\mathscr{A}(x,k)}\left\{U(\gamma) + \mathscr{L}^{\theta,\gamma} v(x,k)\right. \\ &\quad \left. + \lambda_0 \int_0^x v(x - Y, k)\, dG(Y)\right\} = 0, \\ &\qquad\qquad \text{for } x \geq 0,\ k > x, \end{aligned} \tag{40}$$

$$\lambda v_x(x,x) - v_k(x,x) = 0, \quad \text{for } x > 0. \tag{41}$$

Boundary condition (41) comes from the fact that the wealth process will reflect whenever $X = K$ with direction $\widetilde{\gamma}$ and the gradient of $v(x,k)$ is perpendicular to $\widetilde{\gamma}$ at $x = k$.

If we can find a smooth solution for the HJB, then the optimal consumption will actually be given in feedback form by

$$\widehat{\gamma}(x,k) = I\left(v_x(x,k) + v_n(x,k)\right), \tag{42}$$

where $I \triangleq (U')^{-1}$ is the inverse of marginal utility. In addition, we expect the optimal amount invested in the fund to be given by

$$\widehat{\theta}(x,k) = -\frac{\alpha}{\sigma^2} \cdot \frac{v_x(x,k)}{v_{xx}(x,k)}. \tag{43}$$

Usually, it is difficult to justify the smoothness of value function or the existence of classical solution to the HJB equation that appeared in a control problem. The theory of viscosity principally provides us with a way to analyze our problem (cf. Crandall et al. [29]). To proceed our discussion, we need the following alternative expressions of dynamic programming principle; the readers are referred to Pham [30]. In the sequel, we denote by $\mathscr{T}_{0,\tau}$ the set of stopping times valued in $[0,\tau]$; then one has the following.

Proposition 5. *(1) For all $(\theta,\gamma) \in \mathscr{A}(x,k)$ and $\vartheta \in \mathscr{T}_{0,\tau}$,*

$$v(x,k) \geq \mathbb{E}\left[\int_0^{\vartheta} e^{-\beta s} U(\gamma_s)\, ds + e^{-\beta\vartheta} v\left(X_\vartheta^{\theta,\gamma}, K_\vartheta^{\theta,\gamma}\right)\right]. \tag{44}$$

(2) For all $\varepsilon > 0$, there exists $(\theta,\gamma) \in \mathscr{A}(x,k)$ such that, for all $\vartheta \in \mathscr{T}_{0,\tau}$,

$$\begin{aligned} &v(x,k) - \varepsilon \\ &\quad \leq \mathbb{E}\left[\int_0^{\vartheta} e^{-\beta s} U(\gamma_s)\, ds + e^{-\beta\vartheta} v\left(X_\vartheta^{\theta,\gamma}, K_\vartheta^{\theta,\gamma}\right)\right]. \end{aligned} \tag{45}$$

4. Value Function, Viscosity Solution, and Its Regularity

4.1. Value Function and Viscosity Solution. In order to introduce the concept of viscosity solutions, we first introduce some additional notations. Given a locally bounded function ω (i.e., for all $(x,k) \in (0,+\infty) \times (0,+\infty)$, there exists a compact neighborhood V of (x,k) such that ω is bounded on V); we define its upper-semicontinuous envelope ω^* and lower-semicontinuous envelope ω_* on $[0,+\infty) \times (0,+\infty)$ by

$$\begin{aligned} \omega^*(x,k) &= \limsup_{(x',k')\to(x,k)} \omega\left(x',k'\right), \\ \omega_*(x,k) &= \liminf_{(x',k')\to(x,k)} \omega\left(x',k'\right). \end{aligned} \tag{46}$$

Recall that ω^* (resp., ω_*) is the smallest (resp., largest) upper-semicontinuous function (u.s.c.) above (resp., lower-semicontinuous function (l.s.c.) below) ω on $(0,+\infty)\times(0,+\infty)$. Note that a locally bounded function ω on $(0,+\infty)\times(0,+\infty)$ is lower-semicontinuous (resp., upper-semicontinuous) if and only if $\omega=\omega_*$ on $(0,+\infty)\times(0,+\infty)$, and it is continuous if (and only if) $\omega=\omega^*=\omega_*$ on $(0,+\infty)\times(0,+\infty)$.

Remark 6. Here, the first and second partial derivatives with respect to x at $x=0$ mean the right partial derivatives.

Definition 7 (viscosity subsolution and supersolution). An u.s.c. function $\omega\in C$ is a viscosity subsolution of (40) iff for any test function $\psi\in C^{2,1}(0,+\infty)\times(0,+\infty)$; if $(\overline{x},\overline{k})$ is a global maximum point of $\omega^*-\psi$, then

$$\begin{aligned}&\beta\psi\left(\overline{x},\overline{k}\right)-\sup_{(\theta,\gamma)\in\mathscr{A}(\overline{x},\overline{k})}\left\{U\left(\gamma\right)+\mathscr{L}^{\theta,\gamma}\psi\left(\overline{x},\overline{k}\right)\right.\\&\quad\left.+\lambda_0\mathbb{E}\left[\psi\left(\overline{x}-Y,\overline{k}\right)-\psi\left(\overline{x},\overline{k}\right)\right]\right\}\le 0,\\&\lambda\psi_x\left(\overline{x},\overline{k}\right)-\psi_k\left(\overline{x},\overline{k}\right)\le 0.\end{aligned}\tag{47}$$

A l.s.c. function $\omega\in C$ is a viscosity supersolution of (40) iff for any test function $\varphi\in C^{2,1}((0,+\infty)\times(0,+\infty))$; if $(\overline{x},\overline{k})$ is a global minimum point of $\omega_*-\psi$, then

$$\begin{aligned}&\beta\varphi\left(\overline{x},\overline{k}\right)-\sup_{(\theta,\gamma)\in\mathscr{A}(\overline{x},\overline{k})}\left\{U\left(\gamma\right)+\mathscr{L}^{\theta,\gamma}\varphi\left(\overline{x},\overline{k}\right)\right.\\&\quad\left.+\lambda_0\mathbb{E}\left[\varphi\left(\overline{x}-Y,\overline{k}\right)-\varphi\left(\overline{x},\overline{k}\right)\right]\right\}\ge 0,\\&\lambda\varphi_x\left(\overline{x},\overline{k}\right)-\varphi_k\left(\overline{x},\overline{k}\right)\ge 0.\end{aligned}\tag{48}$$

Finally, ω is a viscosity solution of (40) if it is simultaneously a viscosity subsolution and supersolution.

In addition to Definition 7, there are three equivalent definitions on second-ordered Integro-differential partial differential equations; the readers who are interested in the proof of the equivalence of these definitions are referred to Benth et al. [31] or Barles and Imbert [32].

Theorem 8. *$v(x,k)$ is a viscosity solution of (40), where $v(x,k)$ was defined in (25).*

Proof. Let us prove firstly that v is a viscosity supersolution. Let $(\overline{x},\overline{k})\in(0,+\infty)\times(0,+\infty)$ and let $\varphi\in C^{2,1}((0,+\infty)\times(0,+\infty))$ be a test function such that

$$\begin{aligned}0&=\left(v_*-\varphi\right)\left(\overline{x},\overline{k}\right)\\&=\min_{(x,k)\in(0,+\infty)\times(0,+\infty)}\left(v_*-\varphi\right)(x,k).\end{aligned}\tag{49}$$

We further extend the domain of $\varphi(x,k)$ to $\mathbb{R}\times(0,+\infty)$ with the convention that $\varphi(x,k)=0$ for all $x<0$. One will see later that such extension does not prevent us from discussing our problem. By definition of $v_*(\overline{x},\overline{k})$, there exists a sequence (x_m,k_m) in $(0,+\infty)\times(0,+\infty)$, such that

$$\begin{aligned}\left(x_m,k_m\right)&\longrightarrow\left(\overline{x},\overline{k}\right),\\v\left(x_m,k_m\right)&\longrightarrow v_*\left(\overline{x},\overline{k}\right),\end{aligned}\tag{50}$$

when m goes to infinity. By the continuity of φ and by (49) we also have that

$$\zeta_m:=v\left(x_m,k_m\right)-\varphi\left(x_m,k_m\right)\longrightarrow 0,\tag{51}$$

when m goes to infinity.

Let $(\theta,\gamma)\in\mathscr{A}(x,k)$; we denote by $(X_s^{\theta,\gamma},K_s^{\theta,\gamma})$ the associated controlled process. Let τ_m^1 and τ_m^2 be the stopping times given by $\tau_m^1=\inf\{0\le s\le\tau:|X_s^{\theta,\gamma}(x_m,k_m)-x_m|\ge\eta\}$ and $\tau_m^2=\inf\{0\le s\le\tau:|K_s^{\theta,\gamma}(x_m,k_m)-k_m|\ge\eta\}$ in which $\eta>0$ is a fixed constant, and $\tau_m:=\tau_m^1\wedge\tau_m^2$. Let (h_m) be a strictly positive sequence such that

$$\begin{aligned}h_m&\longrightarrow 0,\\\frac{\zeta_m}{h_m}&\longrightarrow 0,\end{aligned}\tag{52}$$

when m goes to infinity. We apply the first part of the dynamic programming principle (44) for $v(x_m,k_m)$ to $\vartheta_m:=\tau_m\wedge h_m$ and get

$$\begin{aligned}&v\left(x_m,k_m\right)\\&\quad\ge\mathbb{E}\left[\int_0^{\vartheta_m}e^{-\beta s}U\left(\gamma_s\right)ds+e^{-\beta\vartheta_m}v\left(X_{\vartheta_m}^{\theta,\gamma},K_{\vartheta_m}^{\theta,\gamma}\right)\right].\end{aligned}\tag{53}$$

Equation (49) implies that $v\ge v_*\ge\varphi$. Thus

$$\begin{aligned}&\varphi\left(x_m,k_m\right)+\zeta_m\\&\quad\ge\mathbb{E}\left[\int_0^{\vartheta_m}e^{-\beta s}U\left(\gamma_s\right)ds+e^{-\beta\vartheta_m}\varphi\left(X_{\vartheta_m}^{\theta,\gamma},K_{\vartheta_m}^{\theta,\gamma}\right)\right].\end{aligned}\tag{54}$$

Applying Itô's formula to $e^{-\beta\vartheta_s}\varphi(X_{\vartheta_s}^{\theta,\gamma},K_{\vartheta_s}^{\theta,\gamma})$ between 0 and ϑ_m, we obtain

$$\begin{aligned}&\frac{\zeta_m}{h_m}+\mathbb{E}\left[\frac{1}{h_m}\int_0^{\vartheta_m}\left(\beta\varphi-U\left(\gamma_s\right)-\mathscr{L}^{\theta,\gamma}\varphi\right.\right.\\&\quad\left.\left.-\lambda_0\mathbb{E}\left[\varphi\left(X-Y,K\right)-\varphi\left(X,K\right)\right]\right)\left(X_s^{\theta,\gamma},K_s^{\theta,\gamma}\right)ds\right]\\&\quad+\mathbb{E}\left[\frac{1}{h_m}\int_0^{\vartheta_m}\left(\lambda\varphi_x-\varphi_k\right)\left(X_s^{\theta,\gamma},K_s^{\theta,\gamma}\right)dM_s\right]\ge 0,\end{aligned}\tag{55}$$

after noting that the stochastic integral term cancels out by taking expectations since the integrand is bounded. Since the random variable inside the expectation in (55) is bounded by a constant independent of m, we then obtain

$$\begin{aligned}&\left(\lambda_0+\beta\right)\varphi\left(x,k\right)-U\left(\gamma\right)-\mathscr{L}^{\theta,\gamma}\varphi\left(x,k\right)\\&\quad-\lambda_0\int_0^x\varphi\left(x-Y,k\right)dG\left(Y\right)\ge 0,\\&\lambda\varphi_x\left(x,k\right)-\varphi_k\left(x,k\right)\ge 0,\end{aligned}\tag{56}$$

when m goes to infinity by the dominated convergence theorem. We conclude from the arbitrariness of $(\theta,\gamma) \in \mathscr{A}(x,k)$. Thus we get (48).

It remains to prove that v is a viscosity subsolution. Let $(\overline{x},\overline{k}) \in (0,+\infty)\times(0,+\infty)$ and let $\psi \in C^{2,1}((0,+\infty)\times(0,+\infty))$ be a test function such that

$$0 = \left(v^* - \psi\right)\left(\overline{x},\overline{k}\right) = \max_{(x,k)\in R^2}\left(v^* - \psi\right)(x,k). \tag{57}$$

We will show the result by contradiction. Assume on the contrary that

$$\begin{aligned} &\beta\psi\left(\overline{x},\overline{k}\right) - \sup_{(\theta,\gamma)\in\mathscr{A}(x,k)}\Big\{U(\gamma) + \mathscr{L}^{\theta,\gamma}\psi\left(\overline{x},\overline{k}\right) \\ &\quad + \lambda_0\mathbb{E}\left[\psi\left(\overline{x}-Y,\overline{k}\right) - \psi\left(\overline{x},\overline{k}\right)\right]\Big\} > 0, \\ &\lambda\psi_x\left(\overline{x},\overline{k}\right) - \psi_k\left(\overline{x},\overline{k}\right) > 0. \end{aligned} \tag{58}$$

There exist $\eta > 0$ and $\varepsilon > 0$ such that

$$\begin{aligned} &\beta\psi\left(x',k'\right) - \sup_{(\theta,\gamma)\in\mathscr{A}(x,k)}\Big\{U(\gamma) + \mathscr{L}^{\theta,\gamma}\psi\left(x',k'\right) \\ &\quad + \lambda_0\mathbb{E}\left[\psi\left(x'-Y,k'\right) - \psi\left(x',k'\right)\right]\Big\} \geq \varepsilon, \\ &\lambda\psi_x\left(x',k'\right) - \psi_n\left(x',k'\right) \geq \varepsilon, \end{aligned} \tag{59}$$

for all $(x',k') \in B(\overline{x},\overline{k},\eta) = \{(x',k') \in (0,+\infty)\times(0,+\infty) : |\overline{x}-x'|^2 + |\overline{k}-k'|^2 < \eta\}$. By the definition of $v(\overline{x},\overline{k})$, there exists a sequence (x_m,k_m) taking values in $B(\overline{k},\eta)$ such that

$$\begin{aligned} &(x_m,k_m) \longrightarrow \left(\overline{x},\overline{k}\right), \\ &v(x_m,k_m) \longrightarrow v\left(\overline{x},\overline{k}\right), \end{aligned} \tag{60}$$

when m goes to infinity. By continuity of ψ and using (57), we also find that

$$\zeta_m := v(x_m,k_m) - \psi(x_m,k_m) \longrightarrow 0, \tag{61}$$

when m goes to infinity. Let (h_m) be a strictly positive sequence such that

$$\begin{aligned} &h_m \longrightarrow 0, \\ &\frac{\zeta_m}{h_m} \longrightarrow 0. \end{aligned} \tag{62}$$

Then, according to the second part of dynamic programming principle (45) and using (57), there is a sequence $(\theta_m,\gamma_m) \in \mathscr{A}(x_m,k_m)$ such that

$$\begin{aligned} &\psi(x_m,k_m) + \zeta_m - \frac{\varepsilon h_m}{2} \\ &\quad \leq \mathbb{E}\left[\int_0^{\vartheta_s} e^{-\beta s}U(\gamma_m)\,ds + e^{-\beta\vartheta_m}\psi\left(X_{\vartheta_m}^{\theta_m,\gamma_m}, K_{\vartheta_m}^{\theta_m,\gamma_m}\right)\right], \end{aligned} \tag{63}$$

in which we take $\vartheta_m = \tau_m' \wedge h_m$, $\tau_m' = \tau_m^3 \wedge \tau_m^4$, $\tau_m^3 = \inf\{0 \leq s \leq \tau : |X_s^{\widehat{\theta},\widehat{\gamma}}(x_m,k_m) - x_m| \geq \eta'\}$, $\tau_m^4 = \inf\{0 \leq s \leq \tau : |K_s^{\theta,\gamma}(x_m,k_m) - k_m| \geq \eta'\}$, and $0 < \eta' < \eta$. Since (x_m,k_m) converges to $(\overline{x},\overline{k})$, we can always assume that $B(x_m,k_m,\eta') \subset B(\overline{x},\overline{k},\eta)$. For $0 \leq s \leq \vartheta_m \leq \tau$, by applying Itô's formula to $e^{-\beta s}\psi(X_s^{\theta_m,\gamma_m}, K_s^{\theta_m,\gamma_m})$, we get

$$\begin{aligned} 0 \geq{}& \frac{\zeta_m}{h_m} - \frac{\varepsilon}{2} + \mathbb{E}\left[\frac{1}{h_m}\int_0^{\vartheta_m} L\left(X_s^{\theta_m,\gamma_m}, K_s^{\theta_m,\gamma_m}\right)ds\right] \\ &+ \mathbb{E}\left[\frac{1}{h_m}\int_0^{\vartheta_m}\left(\lambda\psi_x - \psi_k\right)\left(X_s^{\theta,\gamma}, K_s^{\theta,\gamma}\right)dM_s\right] \end{aligned} \tag{64}$$

with

$$\begin{aligned} L(x,k) ={}& \beta v(x,k) - U(\gamma) - \mathscr{L}^{\theta,\gamma}\psi(x,k) \\ &- \lambda_0\mathbb{E}\left[\psi(x-Y,k) - \psi(x,k)\right], \end{aligned} \tag{65}$$

after noting that the stochastic integral term cancels out by taking expectations since the integrand is bounded.

Moreover, noting that for $0 \leq s < \vartheta_m \leq \tau$

$$\begin{aligned} L\left(X_s^{\theta_m,\gamma_m}, K_s^{\theta_m,\gamma_m}\right) \geq{}& \beta v(x,k) - \sup_{(\theta,\gamma)\in\mathscr{A}(x,k)}\Big\{U(\gamma) \\ &+ \mathscr{L}^{\theta,\gamma}\psi(x,k) + \lambda_0\mathbb{E}\left[\psi(x-Y,k) - \psi(x,k)\right]\Big\} \\ &\geq \varepsilon, \end{aligned} \tag{66}$$

we find using (59) and (64) that

$$0 \geq \frac{\zeta_m}{h_m} - \varepsilon\left(\frac{1}{2} - \frac{1}{h_m}\mathbb{E}\left[\vartheta_m\right]\right) \tag{67}$$

since (see Pham [30, Page 38])

$$\begin{aligned} &\lim_{h_m\downarrow 0^+}\mathbb{E}\left[\sup_{s\in(0,\vartheta_m]}\left|X_s^{\widehat{\theta}_m,\widehat{\gamma}_m} - x_m\right|^2\right] = 0, \\ &\lim_{h_m\downarrow 0^+}\mathbb{E}\left[\sup_{s\in(0,\vartheta_m]}\left|K_s^{\widehat{\theta}_m,\widehat{\gamma}_m} - k_m\right|^2\right] = 0. \end{aligned} \tag{68}$$

By Chebyshev's inequality, we deduce that

$$\begin{aligned} \mathbb{P}\left[\tau_m' \leq h_m\right] \leq{}& \mathbb{P}\left[\sup_{s\in(0,h_m]}\left|X_s^{\widehat{\theta}_m,\widehat{\gamma}_m} - x_m\right| \geq \eta\right] \\ &\cdot \mathbb{P}\left[\sup_{s\in(0,h_m]}\left|K_s^{\widehat{\theta}_m,\widehat{\gamma}_m} - k_m\right| \geq \eta\right] \\ \leq{}& \frac{\mathbb{E}\left|\sup_{s\in(0,h_m]}\left|X_s^{\widehat{\theta}_m,\widehat{\gamma}_m} - x_m\right|^2\right|}{\eta^2} \\ &\cdot \frac{\mathbb{E}\left|\sup_{s\in(0,h_m]}\left|K_s^{\widehat{\theta}_m,\widehat{\gamma}_m} - k_m\right|^2\right|}{\eta^2} \longrightarrow 0, \end{aligned} \tag{69}$$

when h_m goes to zero, that is, when m goes to infinity. Moreover, since

$$\mathbb{E}\left[\vartheta_t\right] = \int_{\{\tau'_m > h_m\}} h_m d\mathbb{P} + \int_{\{\tau'_m \le h_m\}} \left(\tau'_m\right) d\mathbb{P} \tag{70}$$

we deduce that

$$\begin{aligned} h_m \mathbb{P}\left(\tau'_m > h_m\right) = h_m \mathbb{P}\left(\tau'_m > h_m\right) &= \int_{\{\tau'_m > h_m\}} h_m d\mathbb{P} \\ &\le \mathbb{E}\left[\vartheta_t\right] \\ &\le \int_{\{\tau'_m > h_m\}} h_m d\mathbb{P} + \int_{\{\tau'_m \le h_m\}} h_m d\mathbb{P} \\ &= h_m. \end{aligned} \tag{71}$$

So we obtain

$$\mathbb{P}\left[\tau'_m > h_m\right] \le \frac{1}{h_m} \mathbb{E}\left[\vartheta_m\right] \le 1. \tag{72}$$

This implies that $(1/h_m)\mathbb{E}[\vartheta_m]$ converges to 1 when h_m goes to zero. We thus get the desired contradiction by letting m go to infinity in (67).

So (47) holds and we complete the proof. □

4.2. Dimension Reduction and Regularity of Viscosity Solution. A key insight noted by Magill and Constantinides [33] and exploited in Davis and Norman [8] is that because of the homotheticity of power utility function (Proposition 3.3) the dimension of our control problem is ready to be reduced from two to one. In Janeček and Sîrbu [9], where the decision maker is assumed to be a hedge fund manager, such reduction is successful and with such reduction, the authors proved the regularity of the viscosity solution to the HJB equation associated with their control problem. In our problem we guess that the value function, also the viscosity solution to the HJB equation, resembles similar property. The following intuitive interpretation will help us to understand this point. In Section 2.1, it has been shown that the ruin probability of classical risk model can be approximated to a drifted Brownian motion with proper drift and diffusion coefficients. What is more, one can even try to approximate the distribution of the functional of the maximum process of classical risk model by diffusion process. So, if we replace the classical risk model by a proper drifted Brownian motion, then after some easy calculations, one can find that the corresponding HJB equation shares the same formulation with the one presented in Janeček and Sîrbu [9]. In this situation, it is natural to guess that the value function can be reduced from two to one. The main difference of the HJB equation of this paper is that there is an integral term in the HJB equation, however, after noting that the control process is stopped after stopping time $\tau^{\theta,\gamma}$, so we still hope that there is a possibility to reduce the viscosity solution from two to one. More precisely, we expect that

$$v(x,k) = x^{1-p} v\left(1, \frac{k}{x}\right) \triangleq x^{1-p} u(z) \quad \text{for } z \triangleq \frac{k}{x}. \tag{73}$$

In addition, instead of looking for the optimal amounts $\widehat{\theta}(x,k)$ and $\widehat{\gamma}(x,k)$ in (43) and (42) we look for the proportions

$$\widehat{c}(x,k) = \frac{\widehat{\gamma}}{x} = \frac{I\left(v_x(x,k) + v_k(x,k)\right)}{x}, \tag{74}$$

$$\widehat{\pi}(x,k) = \frac{\widehat{\theta}}{x} = -\frac{\alpha}{\sigma^2} \cdot \frac{x v_x(x,k)}{x^2 v_{xx}(x,k)}. \tag{75}$$

Since

$$\begin{aligned} v_k(x,k) &= u'(z) \cdot x^{-p}, \\ v_x(x,k) &= \left((1-p)u(z) - zu'(z)\right) \cdot x^{-p}, \\ v_{xx}(x,k) &= \left(-p(1-p)u(z) + 2pzu'(z) + z^2 u''(z)\right) \\ &\quad \cdot x^{-1-p}, \end{aligned} \tag{76}$$

it is followed that (40) and (41) can be reformulated as

$$\begin{aligned} \sup_{\gamma>0,\theta} \Big\{ &-\beta u + \frac{c^{1-p}}{1-p} + (\pi\alpha - c)\left[(1-p)u - zu'\right] - cu' \\ &+ \frac{1}{2}\pi^2\sigma^2\left(-p(1-p)u + 2pzu' + z^2 u''\right) \\ &+ \lambda_0 \chi(u(z)) \Big\} = 0, \end{aligned} \tag{77}$$

$$-\lambda(1-p)u(1) + (1+\lambda)u'(1) = 0, \tag{78}$$

where, for notation simplicity, we adopt $\chi(u(z))$ for $\int_0^x (x - Y, k) dG(y)$. We also expect that

$$\lim_{z\to\infty} u(z) = \frac{1}{1-p} c_0^{-p} \tag{79}$$

with c_0 given by (96) below; see (98).

The optimal investment proportion in (75) could therefore be expressed (if we can find a smooth solution for reduced HJB (77)) as

$$\widehat{\pi}(z) = \frac{\alpha}{p\sigma^2} \cdot \frac{(1-p)u - zu'}{(1-p)u + 2zu' - (1/p)z^2 u''}, \tag{80}$$

and the optimal consumption proportion $\widehat{c}$ in (74) would be given by

$$\widehat{c}(z) = \frac{\left(v_x + v_k\right)^{-1/p}}{x} = \left((1-p)u - (z-1)u'\right)^{-1/p}. \tag{81}$$

The following theorem asserts the regularity of the viscosity solution to (77) with boundary condition (78).

Theorem 9. *The function u is C^2 on $[1,\infty)$ and satisfies*

$$\begin{aligned} -p(1-p)u+2pzu'+z^2u''<0,\\ (1-p)u-(z-1)u'>0,\\ (1-p)u-zu'>0.\\ z>1. \end{aligned} \tag{82}$$

Moreover, it is a solution of the equation

$$\begin{aligned} &\sup_{c\geq 0,\pi}\mathscr{L}_{c,\pi}u \\ &\quad=-\beta u+\widetilde{V}\left((1-p)u-(z-1)u'\right)+\lambda_0\chi(u) \\ &\qquad-\frac{1}{2}\frac{\alpha^2}{\sigma^2}\frac{\left((1-p)u-zu'\right)^2}{-p(1-p)u+2pzu'+z^2u''},\quad z>1, \end{aligned} \tag{83}$$

$$-\lambda(1-p)u(1)+(1+\lambda)u'(1)=0, \tag{84}$$

where

$$\begin{aligned} \widetilde{V}(y)&=\begin{cases}\dfrac{p}{1-p}y^{(p-1)/p}, & y>0,\\ +\infty, & y\leq 0,\end{cases}\quad \text{for } p<1,\\ \widetilde{V}(y)&=\begin{cases}\dfrac{p}{1-p}y^{(p-1)/p}, & y\geq 0,\\ +\infty, & y<0,\end{cases}\quad \text{for } p>1. \end{aligned} \tag{85}$$

Proof. The proof is very similar to the one for Theorem 5.2 of Janeček and Sîrbu [9] more or less; we do not copy the steps here. One just needs to note that the HJB equation in this paper differs from the one in Janeček and Sîrbu [9] lies in $\chi(u)$; however, this term is not involved in the discussion of the regularity of viscosity. □

Remark 10. Although the jump term of insurer does not affect the smoothness of the value function of our control problem, due to the existence of such jump term, the value function and consequently the optimal policies will be highly influenced. This will be illustrated in the next section by partial analysis on the properties to the viscosity solution.

Theorem 9 claims the regularity of value function $v(x,k)$ when $x>0$. When $x=0$, the value function $v(0,k)$ is specified by the following theorem.

Theorem 11. *$v(0,k)$ satisfies*

$$\begin{aligned} &\beta v(0,k) \\ &\quad-\sup_{0\leq\gamma\leq a}\{U(\gamma)+(a-\gamma)v_0(0,k)-\gamma v_k(0,k)\} \\ &=0. \end{aligned} \tag{86}$$

Proof. If initial surplus of insurer $x=0$, then to invest any amount on risky market can be optimal since the diffusion property of the risky market will cause ruin to happen immediately (cf. Dufresne and Gerber [34]). So optimal investment for insurer is to invest 0 amount on risky market in a very small interval, and of course, the optimal consumption rate γ_t, which is to be determined, should not exceed the premium income rate, say a. Based on this analysis, the HJB function for value function at $x=0$ is reduced to

$$\begin{aligned} &\beta v(0,k) \\ &\quad-\sup_{0\leq\gamma\leq a}\{U(\gamma)+(a-\gamma)v_0(0,k)-\gamma v_k(0,k)\} \\ &=0. \end{aligned} \tag{87}$$

□

4.3. Asymptotic Properties of Value Function. In this section, we will have some asymptotic properties of value function.

Lemma 12. *$v(x,k)$ is bounded on $[0,\infty)\times(0,\infty)$.*

Proof. Revisit the definition of $v(x,k)$, suppose that at time t the wealth process of insurer is $X_t^{\theta,\gamma}$, and then obviously $\gamma_t\leq X_t^{\theta,\gamma}$, or else the ruin will take place, which cannot be the optimal policy for insurer. Thus, one can see that

$$v(x,k)\leq\mathbb{E}\left[\int_0^\tau e^{-\beta t}U\left(\widetilde{X}_t^{\theta,0}\right)\right], \tag{88}$$

where the wealth process $\widetilde{X}_t$, $0\leq t\leq\tau$, is the one under policy θ and $\gamma_t\equiv 0$. So, the policy that maximizes the ruin time τ will maximize $\mathbb{E}[\int_0^\tau e^{-\beta t}U(\widetilde{X}_t^{\theta,0})]$. Yang and Zhang [35] prove that a constant investment policy maximizes this amount. If the insurer adopts the constant investment policy, then the wealth process of insurer is

$$X_t=x+at-Z_t+C^*\ast(\alpha t+\sigma W_t),\quad t\geq 0, \tag{89}$$

where C^* is the constant investment policy. Then, if $p\leq 1$, it is easy to see that $\mathbb{E}[\int_0^\tau e^{-\beta t}U(\widetilde{X}_t^{C^*,0})]$ is bounded. If $p>1$, Protter [36] shows that

$$\mathbb{E}\left[\sup_{0\leq s\leq t}\left|X_s^{C^*,0}\right|^p\right]\leq\Gamma e^{\rho t}\left(1+x^p\right), \tag{90}$$

where Γ and ρ are constants depending on coefficients involved in the wealth process. Thus, by choosing a large enough β, it follows that

$$\mathbb{E}\left[\int_0^\tau e^{-\beta t}U\left(\widetilde{X}_t^{C^*,0}\right)\right]<\Gamma\int_0^\infty e^{-(\beta-\rho)t}x^{1-p}dt. \tag{91}$$

This indicates that $v(x,k)$ is bounded. □

Theorem 13. *For $\widehat{\pi}$ and $\widehat{c}$ that are defined by (74) and (75), one has*

$$\begin{aligned} \lim_{z\to\infty}\widehat{\pi}(z)&=\frac{\alpha}{p\sigma^2},\\ \lim_{z\to\infty}\widehat{c}(z)&=\left((1-p)u(\infty)\right)^{-1/p}, \end{aligned} \tag{92}$$

and $u(\infty) \triangleq \lim_{z\to\infty} u(z)$ *is determined by*

$$\beta u(\infty) - \chi(u(\infty)) = \widetilde{V}((1-p)u(\infty)), \tag{93}$$

where $\chi(u(\infty)) = \lim_{z\to\infty}\int_0^1 v(x-Y,k)\,dG(y) = u(\infty)\int_0^1 (x-Y)^{1-p}\,dG(y)$.

Proof. Note that $u(z) \in C^2[1,\infty)$ and $u(z) = v(1,z)$, by Theorem 9 one can prove that

$$\begin{gathered} zu'(z) \longrightarrow 0, \\ z^2u''(z) \longrightarrow 0, \\ z \longrightarrow \infty. \end{gathered} \tag{94}$$

Here we assume that the above limits exist; in fact, by repeating a similar discussion to the proof for Proposition 4.1 of Janeček and Sîrbu [9], such assumptions are guaranteed. By (80) and (81) and (94), we have (92) immediately. Let $z \to \infty$ in (83), and by (94) we have (93). □

Remark 14. (1) (the case when paying no fee $\lambda = 0$ and $\lambda_0 = 0$) This is the classical problem in Merton [2, 3] and can be solved in closed form. More precisely, for $\lambda = 0$, the optimal investment and consumption proportions are constant, which are given by

$$\pi_0 \triangleq \frac{\alpha}{p\sigma^2}, \tag{95}$$

$$c_0 \triangleq \frac{\beta}{p} - \frac{1}{2}\frac{1-p}{p^2}\cdot\frac{\alpha^2}{\sigma^2}. \tag{96}$$

The Merton value function (and solution of the HJB) equals

$$v_0(x,k) = \frac{1}{1-p}c_0^{-p}x^{1-p}, \quad 0 < x \le n. \tag{97}$$

It follows that for $\lambda = 0$

$$u_0(z) = \frac{1}{1-p_0}c_0^{-p}, \quad z \ge 1. \tag{98}$$

Since u_0 in (98) is constant, (95) and (96) are compatible with the feedback formulas (80) and (81).

As can be easily seen from above, for the case $0 < p < 1$, in order to obtain a finite value function, an additional constraint needs to be imposed on the parameters. This is equivalent to c_0 in (96) being strictly positive, which translates to the standing assumption

$$\beta > \frac{1}{2}\frac{1-p}{p}\cdot\frac{\alpha^2}{\sigma^2}, \quad \text{if } 0 < p < 1. \tag{99}$$

(2) When $\lambda = 0$, our model reduces to the case that an insurer would like to maximize his expected cumulative discount utility form consumption. To the best of our knowledge, this problem has not been addressed before. One may find that when $\lambda = 0$, it means that the insurer does not need to pay any high-watermark fee for the gain profit, which is equal to the case that the initial high-watermark of the insurer is infinity in the model studied in this paper. Denote by $m(x) \triangleq \sup_{\theta,\gamma>0}\mathbb{E}^x[\int_0^\tau U(\gamma_t)]$ the value function of the insurer who does not need to be subject to high-watermark fee; then

$$m(x) = v(x,\infty) = x^{1-p}v(1,\infty) = x^{1-p}u(\infty), \tag{100}$$

where $u(\infty)$ is specified by (93). This is also the value function for the insurer without high-watermark fee.

Comparing v_0x and $m(x)$, it is obvious that two functions share the same power formulation and differ on the constant term. These results indicate that there is no significant difference between the investment and consumption behavior between an insurance company and a hedge fund manager. This is not the first time that we observe such phenomenon; in fact, when we consider the optimal investment for maximizing the survival probability of an insurer (cf. Yang and Zhang [35]) or the one of a fund manager (cf. Browne [37]), the value function shares the same exponential form, which just differs on the constant term.

5. The Verification Theorem

Theorem 15 (the verification theorem). *Let* $\omega(x,k)$ *be a function in* $C^{2,1}((0,+\infty)\times(0,\infty))$ *and satisfy a quadratic growth condition; that is, there exists a constant* D *such that*

$$|\omega(x,k)| \le D\left(1+|x|^2+|k|^2\right). \tag{101}$$

(1) Suppose that

$$\begin{aligned} \beta\omega(x,k) - \sup_{(\theta,\gamma)\in\mathscr{A}(x,k)}&\left\{U(\gamma) + \mathscr{L}^{\theta,\gamma}\omega(x,k)\right. \\ &\left. + \lambda_0\int_0^x \omega(x-Y,k)\,dG(Y)\right\} \ge 0, \\ &\forall (x,k)\in(0,+\infty]\times(0,\infty), \end{aligned} \tag{102}$$

$$\lambda\omega_x(x,k) - \omega_n(x,n) \ge 0, \tag{103}$$

$$\begin{aligned} &\limsup_{t\to\tau} e^{-\beta t}\mathbb{E}\left[\omega\left(X_t^{\theta,\gamma}, K_t^{\theta,\gamma}\right)\right] \ge 0, \\ &\qquad (x,k)\in(0,+\infty)\times(0,\infty). \end{aligned} \tag{104}$$

Then $\omega \ge v$ *on* R^2.

(2) Suppose further that, for all $(x,k)\in(0,+\infty)\times(0,\infty)$, *there exists a measurable function* $(\widehat{\theta}(x,k),\widehat{\gamma}(x,k))$, $(x,k)\in(0,+\infty)\times(0,\infty)$, *value in* $\mathscr{A}$ *such that*

$$\begin{aligned} \beta\omega(x,n) - \sup_{(\theta,\gamma)\in\mathscr{A}(x,k)}&\left\{U(\gamma) + \mathscr{L}^{\theta,\gamma}\omega(x,k)\right. \\ &\left. + \lambda_0\int_0^x \omega(x-Y,k)\,dG(Y)\right\} = -\omega_t(x,k) \\ &- U(\widehat{\gamma}) - \mathscr{L}^{\widehat{\theta},\widehat{\gamma}}\omega(x,k) - \lambda_0\int_0^x \omega(x \\ &- Y,k)\,dG(Y) = 0, \end{aligned} \tag{105}$$

$$\lambda\omega_x(x,k) - \omega_n(x,k) = 0,$$

and SDE (20) admits a unique solution, denoted by $(X_s^{\widehat{\theta},\widehat{\gamma}}, K_s^{\widehat{\theta},\widehat{\gamma}})$*, given an initial condition* $X_0 = x$*, which satisfies*

$$\liminf_{t\to\tau} e^{-\beta t}\mathbb{E}\left[\omega\left(X_t^{\theta,\gamma}, K_t^{\theta,\gamma}\right)\right] \le 0, \tag{106}$$

and the process $\{(\widehat{\theta}(X_s^{\widehat{\theta},\widehat{\gamma}}, K_s^{\widehat{\theta},\widehat{\gamma}}), \widehat{\theta}(X_s^{\widehat{\theta},\widehat{\gamma}}, K_s^{\widehat{\theta},\widehat{\gamma}}))\}$ *that stops at* $\tau^{\theta,\gamma}$ *lies in* $\mathscr{A}(x,n)$*.*

Then

$$\omega = v, \quad on \ (0,\infty)\times(0,\infty), \tag{107}$$

and $(\widehat{\theta},\widehat{\gamma})$ *is an optimal Markovian control.*

Proof. (1) Since $\omega \in C^{2,1}((0,\infty)\times(0,\infty))$, we have for all $(x,k)\in(0,\infty)\times(0,\infty)$, $(\theta,\gamma)\in\mathscr{A}(x,k)$, $s\in(0,\tau]$, similar to (37); by Itô formula and taking the expectation, we have

$$\begin{aligned}\mathbb{E}\left[e^{-\beta(t\wedge\tau)}\omega\left(X_{t\wedge\tau}^{\theta,\gamma}, K_{t\wedge\tau}^{\theta,\gamma}\right)\right] &= \omega(x,k) + \mathbb{E}\left[\int_0^{t\wedge\tau} e^{-\beta u}\left(-\beta\omega\left(X_u^{\theta,\gamma}, K_u^{\theta,\gamma}\right) + \mathscr{L}^{\theta,\gamma}\omega\left(X_u^{\theta,\gamma}, K_u^{\theta,\gamma}\right)\right)du\right]\\ &\quad + \mathbb{E}\left[\lambda_0\int_0^{t\wedge\tau}\int_0^{\infty} e^{-\beta u}\left[\omega\left(X_{u-}^{\theta,\gamma}-Y, K_{u-}^{\theta,\gamma}\right) - \omega\left(X_{u-}^{\theta,\gamma}, K_{u-}^{\theta,\gamma}\right)\right]dG(Y)\,du\right]\\ &\quad + \mathbb{E}\int_0^{t\wedge\tau} e^{-\beta u}\left(-\lambda\omega_x + \omega_k\right)\left(X_u^{\theta,\gamma}, K_u^{\theta,\gamma}\right)dM_u.\end{aligned} \tag{108}$$

Since ω satisfies (102), we have

$$\begin{aligned}&-\beta\omega\left(X_u^{\theta,\gamma}, K_u^{\theta,\gamma}\right) + U(\gamma) + \mathscr{L}^{\theta,\gamma}\omega\left(X_u^{\theta,\gamma}, K_u^{\theta,\gamma}\right)\\ &\quad + \lambda_0\mathbb{E}\left[\omega(X-Y,K) - \omega(X,K)\right] \le 0,\\ &\qquad\qquad \forall(\theta,\gamma)\in\mathscr{A}(x,k),\\ &\lambda\omega_x(x,k) - \omega_k(x,k) \ge 0,\end{aligned} \tag{109}$$

and so

$$\begin{aligned}&\mathbb{E}\left[e^{-\beta(T\wedge\tau)}\omega\left(X_{T\wedge\tau}^{\theta,\gamma}, K_{T\wedge\tau}^{\theta,\gamma}\right)\right]\\ &\quad \le \omega(x,k) - \mathbb{E}\left[\int_0^{T\wedge\tau} e^{-\beta u}U(\gamma_u)\,du\right],\\ &\qquad\qquad \forall(\theta,\gamma)\in\mathscr{A}(x,k).\end{aligned} \tag{110}$$

We have

$$\left|\mathbb{E}\left[\int_0^{T\wedge\tau} e^{-\beta u}U(\gamma_u)\,du\right]\right| \le \int_0^{T\wedge\tau}\left|e^{-\beta u}U(\gamma_u)\right|du. \tag{111}$$

The right hand side of (111) is integrable by the integrability condition on $\mathscr{A}(x,k)$. According to (110), by sending t to τ, since ω satisfies a quadratic grown condition, we obtain by the dominated convergence theorem and by (104)

$$\begin{aligned}&0 \le \omega(x,k) - \mathbb{E}\left[\int_0^{T\wedge\tau} e^{-\beta u}U(\gamma_u)\,du\right],\\ &\qquad\qquad \forall(\theta,\gamma)\in\mathscr{A}(x,k).\end{aligned} \tag{112}$$

Since $(\theta,\gamma)\in\mathscr{A}(x,k)$ is arbitrary, we conclude that $\omega(x,k)\ge v(x,k)$, for all $(x,k)\in R^2$.

(2) We apply Itô's formula to $e^{-\beta u}\omega(X_u^{\widehat{\theta},\widehat{\gamma}}, K_u^{\widehat{\theta},\widehat{\gamma}})$ between 0 and t (after an eventual localization for removing the stochastic integral term in the expectation):

$$\begin{aligned}\mathbb{E}\left[e^{-\beta t}\omega\left(X_t^{\widehat{\theta},\widehat{\gamma}}, K_t^{\widehat{\theta},\widehat{\gamma}}\right)\right] &= \omega(x,k) + \mathbb{E}\left[\int_0^{T}\left(-\beta\omega\left(X_u^{\widehat{\theta},\widehat{\gamma}}, K_u^{\widehat{\theta},\widehat{\gamma}}\right) + \mathscr{L}^{\widehat{\theta},\widehat{\gamma}}\omega\left(X_u^{\widehat{\theta},\widehat{\gamma}}, K_u^{\widehat{\theta},\widehat{\gamma}}\right)\right)du\right]\\ &\quad + \mathbb{E}\left[\lambda_0\int_0^{T}\int_0^{\infty} e^{-\beta u}\left[\omega\left(X_{u-}^{\widehat{\theta},\widehat{\gamma}}-Y, K_{u-}^{\widehat{\theta},\widehat{\gamma}}\right) - \omega\left(X_{u-}^{\widehat{\theta},\widehat{\gamma}}, K_{u-}^{\widehat{\theta},\widehat{\gamma}}\right)\right]dG(Y)\,du\right]\\ &\quad + \mathbb{E}\int_0^{t} e^{-\beta u}\left(-\lambda\omega_x + \omega_k\right)\left(u, X_u^{\widehat{\theta},\widehat{\gamma}}, K_u^{\widehat{\theta},\widehat{\gamma}}\right)dM_u.\end{aligned} \tag{113}$$

Now, by definition of $(\widehat{\theta},\widehat{\gamma})$, we have

$$\begin{aligned}&\beta\omega(x,k) - U(\widehat{\gamma}) - \mathscr{L}^{\widehat{\theta},\widehat{\gamma}}\omega(x,k)\\ &\quad - \lambda_0\mathbb{E}\left[\omega(x-Y,K) - \omega(x,k)\right] = 0,\\ &\lambda\omega_x(x,k) - \omega_k(x,k) = 0,\end{aligned} \tag{114}$$

and so

$$\begin{aligned}&\mathbb{E}\left[e^{-\beta t}\omega\left(X_t^{\widehat{\theta},\widehat{\gamma}}, K_t^{\widehat{\theta},\widehat{\gamma}}\right)\right]\\ &\quad = \omega(x,k) - \mathbb{E}\left[\int_0^{t} e^{-\beta u}U(\widehat{\gamma}_u)\,du\right].\end{aligned} \tag{115}$$

By sending T to τ and from (106), we obtain

$$\omega(x,k) = \mathbb{E}\left[\int_0^\tau e^{-\beta u} U(\widehat{\gamma}_u)\, du\right] \le v(x,k), \tag{116}$$

and finally we obtain that $\omega = v$ with $(\widehat{\theta}, \widehat{\gamma})$ as an optimal Markovian control. So we complete the proof. □

6. Conclusions

In this paper, we study the optimal investment and consumption problem of an insurer, where the consumption of insurer can be regarded as a kind of dividend payment. Thus, the problem considered in this paper is of practical relevance and reasonable. By dynamic programming method, the associated HJB equation is derived and the value function is proved to be the viscosity solutions. This result enables us to apply the numerical scheme for PDE, especially for HJB equation in viscosity sense (cf. Soner [38]) to find the optimal investment and consumption policies and the value function.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors are very grateful to editors and anonymous advices and help, which improved this paper greatly. This paper was supported by Natural Science Foundation of China (11201006 and 11301303), Humanities and Social Sciences Project of the Ministry of Education of China (12YJC910012, 14YJA630088, and 15YJC910008), and Natural Science Foundation of the Jiangsu Higher Education Institutions of China (15KJB110009).

References

[1] S. E. Shreve and H. M. Soner, "Optimal investment and consumption with transaction costs," *The Annals of Applied Probability*, vol. 4, no. 3, pp. 609–692, 1994.

[2] R. C. Merton, "Lifetime portfolio selection under uncertainty: the continuous-time case," *The Review of Economics and Statistics*, vol. 51, no. 3, pp. 247–257, 1969.

[3] R. C. Merton, "Optimum consumption and portfolio rules in a continuous-time model," *Journal of Economic Theory*, vol. 3, no. 4, pp. 373–413, 1971.

[4] T. Zariphopoulou, "Optimal investment and consumption models with non-linear stock dynamics," *Mathematical Methods of Operations Research*, vol. 50, no. 2, pp. 271–296, 1999.

[5] B. Øksendal and A. Sulem, "Optimal consumption and portfolio with both fixed and proportional transaction costs," *SIAM Journal on Control and Optimization*, vol. 40, no. 6, pp. 1765–1790, 2002.

[6] W. H. Fleming and T. Pang, "A stochastic control model of investment, production and consumption," *Quarterly of Applied Mathematics*, vol. 63, no. 1, pp. 71–87, 2005.

[7] I. Karatzas and S. E. Shreve, *Methods of Mathematical Finance*, vol. 39, Springer, 1998.

[8] M. H. Davis and A. R. Norman, "Portfolio selection with transaction costs," *Mathematics of Operations Research*, vol. 15, no. 4, pp. 676–713, 1990.

[9] K. Janeček and M. Sîrbu, "Optimal investment with high-watermark performance fee," *SIAM Journal on Control and Optimization*, vol. 50, no. 2, pp. 790–819, 2012.

[10] A. E. Whalley and P. Wilmott, "An asymptotic analysis of an optimal hedging model for option pricing with transaction costs," *Mathematical Finance*, vol. 7, no. 3, pp. 307–324, 1997.

[11] J. Cvitanić and I. Karatzas, "Hedging and portfolio optimization under transaction costs: a martingale approach," *Mathematical Finance*, vol. 6, no. 2, pp. 133–165, 1996.

[12] H. Liu and M. Loewenstein, "Optimal portfolio selection with transaction costs and finite horizons," *Review of Financial Studies*, vol. 15, no. 3, pp. 805–835, 2002.

[13] R. Korn, "Portfolio optimisation with strictly positive transaction costs and impulse control," *Finance and Stochastics*, vol. 2, no. 2, pp. 85–114, 1998.

[14] A. A. Obizhaeva and J. Wang, "Optimal trading strategy and supply/demand dynamics," *Journal of Financial Markets*, vol. 16, no. 1, pp. 1–32, 2013.

[15] J. E. Stiglitz, "Some aspects of the taxation of capital gains," *Journal of Public Economics*, vol. 21, no. 2, pp. 257–294, 1983.

[16] R. M. Dammon, C. S. Spatt, and H. H. Zhang, "Optimal consumption and investment with capital gains taxes," *Review of Financial Studies*, vol. 14, no. 3, pp. 583–616, 2001.

[17] W. N. Goetzmann, J. E. Ingersoll Jr., and S. A. Ross, "High-water marks and hedge fund management contracts," *The Journal of Finance*, vol. 58, no. 4, pp. 1685–1718, 2003.

[18] P. Guasoni and G. Wang, "High-water marks and separation of private investments," *SSRN Electronic Journal*, 2012.

[19] B. Højgaard and M. Taksar, "Optimal proportional reinsurance policies for diffusion models with transaction costs," *Insurance: Mathematics and Economics*, vol. 22, no. 1, pp. 41–51, 1998.

[20] A. Cairns, "Some notes on the dynamics and optimal control of stochastic pension fund models in continuous time," *ASTIN Bulletin*, vol. 30, no. 1, pp. 19–56, 2000.

[21] J. Zhu, "Optimal dividend control for a generalized risk model with investment incomes and debit interest," *Scandinavian Actuarial Journal*, vol. 2013, no. 2, pp. 140–162, 2013.

[22] L. He and Z. Liang, "Optimal dynamic asset allocation strategy for ELA scheme of DC pension plan during the distribution phase," *Insurance: Mathematics and Economics*, vol. 52, no. 2, pp. 404–410, 2013.

[23] T. R. Bielecki and S. R. Pliska, "Risk sensitive asset management with transaction costs," *Finance and Stochastics*, vol. 4, no. 1, pp. 1–33, 2000.

[24] M. Dai and F. Yi, "Finite-horizon optimal investment with transaction costs: a parabolic double obstacle problem," *Journal of Differential Equations*, vol. 246, no. 4, pp. 1445–1469, 2009.

[25] V. R. Young, "Optimal investment strategy to minimize the probability of lifetime ruin," *North American Actuarial Journal*, vol. 8, no. 4, pp. 106–126, 2004.

[26] J. Grandell, *Aspects of Risk Theory*, Springer, Berlin, Germany, 1991.

[27] P. Azcue and N. Muler, "Optimal reinsurance and dividend distribution policies in the cramér-lundberg model," *Mathematical Finance*, vol. 15, no. 2, pp. 261–308, 2005.

[28] S. Watanabe and N. Ikeda, *Stochastic Differential Equations and Diffusion Processes*, Elsevier, 1981.

[29] M. G. Crandall, H. Ishii, and P.-L. Lions, "User's guide to viscosity solutions of second order partial differential equations," *Bulletin of the American Mathematical Society*, vol. 27, no. 1, pp. 1–67, 1992.

[30] H. Pham, *Continuous-time Stochastic Control and Optimization with Financial Applications*, vol. 1, Springer, 2009.

[31] F. E. Benth, K. H. Karlsen, and K. Reikvam, "Portfolio optimization in a Lévy market with intertemporal substitution and transaction costs," *Stochastics*, vol. 74, no. 3-4, pp. 517–569, 2002.

[32] G. Barles and C. Imbert, "Second-order elliptic integro-differential equations: viscosity solutions' theory revisited," *Annales de l'Institut Henri Poincare. Annales: Analyse Non Lineaire/Nonlinear Analysis*, vol. 25, no. 3, pp. 567–585, 2008.

[33] M. J. P. Magill and G. M. Constantinides, "Portfolio selection with transactions costs," *Journal of Economic Theory*, vol. 13, no. 2, pp. 245–263, 1976.

[34] F. Dufresne and H. U. Gerber, "Risk theory for the compound poisson process that is perturbed by diffusion," *Insurance Mathematics and Economics*, vol. 10, no. 1, pp. 51–59, 1991.

[35] H. Yang and L. Zhang, "Optimal investment for insurer with jump-diffusion risk process," *Insurance: Mathematics and Economics*, vol. 37, no. 3, pp. 615–634, 2005.

[36] P. E. Protter, *Stochastic Integration and Differential Equations: Version 2.1*, vol. 21, Springer, 2004.

[37] S. Browne, "Optimal investment policies for a firm with a random risk process: exponential utility and minimizing the probability of ruin," *Mathematics of Operations Research*, vol. 20, no. 4, pp. 937–958, 1995.

[38] H. M. Soner, "Optimal control of jump-markov processes and viscosity solutions," in *Stochastic Differential Systems, Stochastic Control Theory and Applications*, pp. 501–511, Springer, 1988.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 818353, 22 pages
http://dx.doi.org/10.1155/2015/818353

Research Article

Singularly Perturbation Method Applied To Multivariable PID Controller Design

Mashitah Che Razali,[1] Norhaliza Abdul Wahab,[1] P. Balaguer,[2] M. F. Rahmat,[1] and Sharatul Izah Samsudin[3]

[1] *Department of Control and Mechatronics Engineering, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia*
[2] *Department of Industrial Systems Engineering and Design, Jaume I University of Castello, 12080 Castello de la Plana, Spain*
[3] *Department of Industrial Electronics, Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka, 76100 Durian Tunggal, Melaka, Malaysia*

Correspondence should be addressed to Norhaliza Abdul Wahab; aliza@fke.utm.my

Received 8 April 2015; Accepted 4 June 2015

Academic Editor: Herve G. E. Kadji

Proportional integral derivative (PID) controllers are commonly used in process industries due to their simple structure and high reliability. Efficient tuning is one of the relevant issues of PID controller type. The tuning process always becomes a challenging matter especially for multivariable system and to obtain the best control tuning for different time scales system. This motivates the use of singularly perturbation method into the multivariable PID (MPID) controller designs. In this work, wastewater treatment plant and Newell and Lee evaporator were considered as system case studies. Four MPID control strategies, Davison, Penttinen-Koivo, Maciejowski, and Combined methods, were applied into the systems. The singularly perturbation method based on Naidu and Jian Niu algorithms was applied into MPID control design. It was found that the singularly perturbed system obtained by Naidu method was able to maintain the system characteristic and hence was applied into the design of MPID controllers. The closed loop performance and process interactions were analyzed. It is observed that less computation time is required for singularly perturbed MPID controller compared to the conventional MPID controller. The closed loop performance shows good transient responses, low steady state error, and less process interaction when using singularly perturbed MPID controller.

1. Introduction

Multivariable PID Control. Among the controller variety, PID becomes the controller that is most applied in a physical system [1]. The reason is that it has a characteristic that offers simplicity, clear functionality, and ease of use [2]. However, Ho et al. [3] reported that only one-fifth of PID control loops are in good condition. The others are not, where 30% of PID controllers are not able to perform well due to lack of tuning parameters, 30% due to the installation of a controller system operating manual, and 20% due to the use of default controller parameters.

In recent years, many researchers have paid attention to the MPID controller design for various systems such as in Industrial Scale-Polymerization Reactor [4], Coupled Pilot Plant Distillation Column [5], Narmada Main Canal [6], Quadruple-Tank Process [7], Boiler-Turbine Unit [8], and Wood-Berry Distillation Column [9]. A research by Kumar et al. [4] had proposed a synthesis method of PI controllers based on approximation of relative gain array (RGA) concept to multivariable process. The method was further improved by relative normalize gain array concept (RNGA). Controller based on RNGA concept provides a better performance than RGA concept. Both concepts use the nonstandard PID controller which requires Maclaurin series expansion [10]. In the work by Sarma and Chidambaram [5], PI/PID controllers based on Davison and Tanttu-Lieslehto method extended to nonsquare systems with right-half plane zero were applied. Results show that the Davison method gives better performance with less settling time than

Tanttu-Lieslehto method. However, both methods are not applicable for square system.

Essentially, there are many integral controllers that are designed for nonlinear system [11, 12]. However, most of the existing techniques do not guarantee the desired transient performances in the presence of nonlinear parameter variations and unknown external disturbances [13]. In a previous study, Martin and Katebi [14] had proposed Davison, Penttinen-Koivo, Maciejowski, and Combined method as a control tuning design for ship positioning. These controls strategies are based on PID controller which is used to control multivariable system. Due to the effectiveness and simplicity of those proposed controllers, Wahab et al. [15] had used those methods as tuning strategies for wastewater treatment plant (WWTP). The controllers were designed based on steady state of a linear system and static model inverse. The reliability of the proposed method was tested to a nonlinear WWTP. The response shows that good result was obtained from Davison until the Combined method. In the work by Balaguer et al. [16], a comparison between MPID controller with figure of merit controller was done for WWTP based on open-loop, closed-loop, and open-closed loop controller structure analysis. The MPID control tuning based on Davison and Penttinen-Koivo method was carried out by minimizing the residuals of both controllers obtained from the data. However, the dynamic nature of WWTP which involves ill condition characteristic causes difficulties in finding the optimum MPID tuning parameter. The system's behavior that involves slow and fast variables causes the control tuning strategies to not easily meet specification for multiple control variables at the same time.

A lot of approaches have been proposed to control multivariable system. Some of the approaches are able to deal with a high order multivariable system. However, a simple controller design has always become a desired controller where it is certainly can be accepted by the industry. By that, the required cost to run the system will be minimized as well. Realizing the simple controller design by other researchers [14–16], those methods were applied in this project and improved by adopting singularly perturbation method (SPM) to the controller designs by considering the dynamic matrix inverse.

Singularly Perturbed Multivariable Controller. Analysis and synthesis of singularly perturbed control have received much thoughtfulness over the past decades by many people from numerous arenas of studies [17–21]. Singularly perturbation method is able to decompose and simplify the higher order of the full order system into slow and fast subsystems [17, 22, 23], which are known as singularly perturbation system. Definitely, most of the control systems are dynamic, where the decomposition into stages is dictated by multitime scale. In this situation, the slow subsystem corresponds to the slowest phenomena and the fast subsystem corresponds to the fastest phenomena. It basically has two different parts of eigenvalue represented for slow and fast dynamic subsystems [24], where slow subsystem corresponds to small eigenvalue and fast subsystem corresponds to large eigenvalue [25].

This work is focused on the analysis of singularly perturbation system on two different case studies given. Singularly perturbed control of multivariable system is comprised of two steps. First, the multivariable system is decomposed into slow and fast subsystems. Then, the optimal composite singularly perturbed controller is designed [25–28]. There are many approaches that have been developed concerning the control of singularly perturbation system. The approaches use different conditions on the properties of the used functions, different assumptions, different theorems, and different lemma [13, 21, 29] which are specifically based on the systems behaviour.

In a study by Rabah and Aldhaheri [24], singularly perturbation system has been modelled by using a real Schur form method. It shows that any two-time scale system can be altered into the singularly perturbed form via a transformation into an order real Schur form (ORSF). It is based on two steps, transformation of matrix A into an ORSF using an orthogonal matrix and then application of balancing algorithm to an ORSF. Li and Lin [17] had addressed the composite fuzzy multivariable controller to nonlinear singularly perturbation system. The composite controller was obtained from the combination of slow and fast subsystems. It was tested to a DC motor driven inverted pendulum system and it provides realistic and satisfactory simulation results. Multivariable control by Kim et al. [30] used successive Galerkin approximation (SGA) method. This method causes the complexity in computations to increase with respect to the order of the system. Therefore, singularly perturbation method was adopted to decompose the original system into slow and fast subsystems. Result shows that the use of the method greatly reduces the computation complexity and it is more effective than the original SGA method.

To the best of author knowledge, there are two methods to obtain the singularly perturbation system, which are by analytical [21, 29–31] and linear analysis [32–35]. Singularly perturbation system obtained based on linear analysis is discussed and has been applied in this research. By exploiting the properties of singularly perturbation system to the dynamic matrix inverse of MPID control tuning methodology, an easy multivariable tuning method should be established. In Section 2, the time scale analysis is presented to determine the behavior of the system. Section 3 described the methods to obtain singularly perturbation system based on Naidu and Jian Niu method. The sequences of MPID controller based on Davison, Penttinen-Koivo, Maciejowski, and Combined methods are discussed in Section 4. Section 5 presented the optimization method which is based on particle swarm optimization (PSO). The case studies and the performance of the proposed methods for two case studies are investigated and discussed thoroughly in Sections 6 and 7. Finally, conclusions are given in Section 8.

2. Time Scale Analysis

To apply singularly perturbation method to the controller designs, the considered system must consist of a two-time scale characteristic. The two-time scale characteristic can be

determine by rearranging the eigenvalue of the system in increasing order which will give

$$e(A) = \{p_{s1}, \ldots, p_{sm}, p_{f1}, \ldots, p_{fn}\}, \tag{1a}$$

$$e(A_s) = \{p_{s1}, \ldots, p_{sm}\}, \tag{1b}$$

$$e(A_f) = \{p_{f1}, \ldots, p_{fn}\}, \tag{1c}$$

where $e(A)$, $e(A_s)$, and $e(A_f)$ are a total, slow, and fast eigenspectrum of the system, respectively. p_{s1} is a smallest eigenvalue of the slow eigenspectrum, p_{sm} is a largest eigenvalue of the slow eigenspectrum, p_{f1} is a smallest eigenvalue of the fast eigenspectrum, and p_{fn} is a largest eigenvalue of the fast eigenspectrum:

$$0 < |p_{s1}| < \cdots < |p_{sm}| < |p_{f1}| < \cdots < |p_{fn}|. \tag{2}$$

The system is said to possess a two-time scale characteristic, if the largest absolute eigenvalue of the slow eigenspectrum is much smaller than the smallest absolute eigenvalue of the fast eigenspectrum. This is proven by

$$\varepsilon = \frac{|p_{sm}|}{|p_{f1}|} \ll 1, \tag{3}$$

where ε is a measure of separation of time scales that represents an intrinsic property.

3. Singularly Perturbation Method (SPM) for MIMO System

Industrial processes possess "n" number of inputs and outputs variables, where interaction phenomena exist. Interaction phenomena that occur among the inputs and outputs variables of multivariable process cause great difficulties in MPID controller design. Usually, it is solved by tuning the most important loop whereas other loops are detuned by keeping the interactions of that loop adequate. To compensate the interaction phenomena, one of the loops is forced to perform poorly. This detuning method is far from the optimal [4]. In this work, to account for the interaction phenomena, instead of using the original process transfer function to the MPID controller design, that transfer function is rearranged by separating the slow and fast eigenvalues using SPM.

3.1. Naidu Method. In this section we propose a procedure for a separation of slow and fast subsystem. The considered linear equations for two-time scale continuous system with the output vector possessing two widely separated groups of eigenvalues are

$$\dot{x} = A_{11}x + A_{12}z + B_1u, \tag{4a}$$

$$\varepsilon\dot{z} = A_{21}x + A_{22}z + B_2u, \tag{4b}$$

$$y = C_1x + C_2z, \tag{4c}$$

where x and z are slow and fast variables in p and q dimension and y is a measured output. Matrices A_{ij}, B_i, and C_i are constant matrices of appropriate dimensions. Consider the problem as in (4a) to (4c). The system possesses a two-time scale property. Preliminary to separation of slow and fast subsystem, the system consists of m number of small eigenvalue (close to the origin) for slow subsystem and n number of fast eigenvalue (far from the origin) for the fast subsystem. The number of slow and fast eigenvalues needs to be identified based on eigenvalue location. Fast eigenvalue of the system is only essential during a short period of time. Then, it is insignificant and the characteristic of the system can be described by degenerating system known as a slow subsystem.

By letting $\varepsilon = 0$, slow subsystem is obtained as

$$\dot{x}_{\text{slow}} = A_{11}x_{\text{slow}} + A_{12}z_{\text{slow}} + B_1u_{\text{slow}}, \tag{5a}$$

$$0 = A_{21}x_{\text{slow}} + A_{22}z_{\text{slow}} + B_2u_{\text{slow}}, \tag{5b}$$

$$y_{\text{slow}} = C_1x_{\text{slow}} + C_2z_{\text{slow}}. \tag{5c}$$

By assuming A_{22} as a nonsingular matrix, (5b) becomes

$$z_{\text{slow}} = -A_{22}^{-1}(A_{21}x_{\text{slow}} + B_2u_{\text{slow}}). \tag{6}$$

Using equation (6) in (5a), $\dot{x}_{\text{slow}}$ is represented as

$$\dot{x}_{\text{slow}} = A_{\text{slow}}x_{\text{slow}} + B_{\text{slow}}u_{\text{slow}}, \tag{7}$$

where

$$\begin{aligned} A_{\text{slow}} &= A_{11} - A_{12}A_{22}^{-1}A_{21}, \\ B_{\text{slow}} &= B_1 - A_{12}A_{22}^{-1}B_2. \end{aligned} \tag{8}$$

Using (6) in (5c), y_{slow} is represented as

$$y_{\text{slow}} = C_{\text{slow}}x_{\text{slow}} + D_{\text{slow}}u_{\text{slow}}, \tag{9}$$

where

$$\begin{aligned} C_{\text{slow}} &= C_1 - C_2A_{22}^{-1}A_{21}, \\ D_{\text{slow}} &= -C_2A_{22}^{-1}B_2. \end{aligned} \tag{10}$$

To obtain fast subsystem, it can be assumed that the slow variables are constant at fast transients, where

$$\begin{aligned} x_{\text{slow}} &= x = \text{constant}, \\ \dot{z}_{\text{slow}} &= 0. \end{aligned} \tag{11}$$

From (4b) and (6),

$$\varepsilon(\dot{z} - \dot{z}_{\text{slow}}) = A_{22}(z - z_{\text{slow}}) + B_2(u - u_{\text{slow}}). \tag{12}$$

Let

$$\begin{aligned} z_{\text{fast}} &= (z - z_{\text{slow}}), \\ u_{\text{fast}} &= (u - u_{\text{slow}}), \\ y_{\text{fast}} &= (y - y_{\text{slow}}), \\ A_{\text{fast}} &= A_{22}, \\ B_{\text{fast}} &= B_2, \\ C_{\text{fast}} &= C_2. \end{aligned} \tag{13}$$

The fast subsystem is obtained as

$$\varepsilon \dot{z}_{\text{fast}} = A_{\text{fast}} z_{\text{fast}} + B_{\text{fast}} u_{\text{fast}}, \tag{14a}$$

$$y_{\text{fast}} = C_{\text{fast}} \left(z - z_{\text{slow}}\right), \tag{14b}$$

$$y_{\text{fast}} = C_{\text{fast}} z_{\text{fast}}. \tag{14c}$$

The composite system which consists of slow and fast subsystem is achieved using two-stage linear transformation which can be referred in an article written by Chang [36]:

$$A_{\text{spm}} = \begin{bmatrix} A_{\text{slow}} & Z_{12} \\ Z_{21} & A_{\text{fast}} \end{bmatrix}, \tag{15a}$$

$$B_{\text{spm}} = \begin{bmatrix} B_{\text{slow}} \\ B_{\text{fast}} \end{bmatrix}, \tag{15b}$$

$$C_{\text{spm}} = \left[C_{\text{slow}} \quad C_{\text{fast}}\right], \tag{15c}$$

$$D_{\text{spm}} = D, \tag{15d}$$

where

$$Z_{12} = \text{zeros}\left(m, 1\right), \tag{15e}$$

$$Z_{21} = \text{zeros}\left(1, m\right). \tag{15f}$$

The state space form of composite system is represented in (15a) to (15f).

3.2. Jian Niu Method. The two-time scale system can also be solved using other method. This section presents singularly perturbation method based on Jian Niu. In order to apply Jian Niu method, transfer function matrix should be transform into a state space model:

$$G\left(s\right) = \begin{bmatrix} G_{11}\left(s\right) & G_{12}\left(s\right) \\ G_{21}\left(s\right) & G_{22}\left(s\right) \end{bmatrix}. \tag{16}$$

To illustrate the two-time scale decomposition, (16) is considered. Equation (16) can have this form

$$\dot{x} = Ax + Bu, \tag{17a}$$

$$y = Cx, \tag{17b}$$

where

$$A = \text{diag}\left[A_{11}, A_{12}, A_{21}, A_{22}\right], \tag{17c}$$

$$B = \begin{bmatrix} B_{11} & 0 \\ 0 & B_{12} \\ B_{21} & 0 \\ 0 & B_{22} \end{bmatrix}, \tag{17d}$$

$$C = \begin{bmatrix} C_{11} & C_{12} & 0 & 0 \\ 0 & 0 & C_{21} & C_{22} \end{bmatrix} \tag{17e}$$

(A_{ij}, B_{ij}, C_{ij}) is a state space form of $G_{ij}(s)$. Equations (17a) and (17b) can be represented as

$$\begin{bmatrix} \varepsilon \dot{x}_1 \\ \varepsilon \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} \varepsilon A_{11} & 0 & 0 & 0 \\ 0 & \varepsilon A_{12} & 0 & 0 \\ 0 & 0 & A_{21} & 0 \\ 0 & 0 & 0 & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} \varepsilon B_{11} & 0 \\ 0 & \varepsilon B_{12} \\ B_{21} & 0 \\ 0 & B_{22} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \tag{18}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & 0 & 0 \\ 0 & 0 & C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix},$$

where ε is a very small positive constant. Equations (18) can be considered as

$$\dot{x} = \overline{A}_{11} x + \overline{A}_{12} z + \overline{B}_1 u, \tag{19a}$$

$$\varepsilon \dot{z} = \overline{A}_{21} x + \overline{A}_{22} z + \overline{B}_2 u, \tag{19b}$$

$$y = \overline{C}_1 x + \overline{C}_2 z + \overline{D} u. \tag{19c}$$

Equations (19a) to (19c) are the linear equations for two-time scale continuous system, similar just like (1a) to (1c) where

$$\overline{A}_{11} = \begin{bmatrix} A_{21} & 0 \\ 0 & A_{22} \end{bmatrix}, \tag{19d}$$

$$\overline{A}_{12} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \tag{19e}$$

$$\overline{A}_{21} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \tag{19f}$$

$$\overline{A}_{22} = \begin{bmatrix} \varepsilon A_{11} & 0 \\ 0 & \varepsilon A_{12} \end{bmatrix}, \tag{19g}$$

$$\overline{B}_1 = \begin{bmatrix} B_{21} & 0 \\ 0 & B_{22} \end{bmatrix}, \tag{19h}$$

$$\overline{B}_2 = \begin{bmatrix} \varepsilon B_{11} & 0 \\ 0 & \varepsilon B_{12} \end{bmatrix}, \tag{19i}$$

$$\overline{C}_1 = \begin{bmatrix} 0 & 0 \\ C_{21} & C_{22} \end{bmatrix}, \tag{19j}$$

$$\overline{C}_2 = \begin{bmatrix} C_{11} & C_{12} \\ 0 & 0 \end{bmatrix}, \tag{19k}$$

$$\overline{D} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{19l}$$

This method is discussed in several literatures [33, 35]. The slow subsystem is denoted as

$$\dot{x}_{\text{slow}} = A_{\text{slow}} x_{\text{slow}} + B_{\text{slow}} u_{\text{slow}}, \tag{20a}$$

$$y = C_{\text{slow}} x_{\text{slow}} + D_{\text{slow}} u_{\text{slow}}. \tag{20b}$$

And the fast subsystem is

$$\dot{z}_{\text{fast}} = A_{\text{fast}} z_{\text{fast}} + B_{\text{fast}} u_{\text{fast}}, \tag{21a}$$

$$y = C_{\text{fast}} z_{\text{fast}} + D_{\text{fast}} u_{\text{fast}}, \tag{21b}$$

where

$$A_{\text{slow}} = \overline{A}_{11} - \overline{A}_{12} \overline{A}_{22}^{-1} \overline{A}_{21}, \tag{22}$$

$$B_{\text{slow}} = \overline{B}_1 - \overline{A}_{12} \overline{A}_{22}^{-1} \overline{B}_2, \tag{23}$$

$$C_{\text{slow}} = \overline{C}_1 - \overline{C}_2 \overline{A}_{22}^{-1} \overline{A}_{21}, \tag{24}$$

$$D_{\text{slow}} = \overline{D} - \overline{C}_2 \overline{A}_{22}^{-1} \overline{B}_2, \tag{25}$$

$$A_{\text{fast}} = \overline{A}_{22}, \tag{26}$$

$$B_{\text{fast}} = \overline{B}_2, \tag{27}$$

$$C_{\text{fast}} = \overline{C}_2, \tag{28}$$

$$D_{\text{fast}} = \overline{D}, \tag{29}$$

$$G_{\text{slow}}(s) = C_{\text{slow}} \left(sI - A_{\text{slow}}\right)^{-1} B_{\text{slow}} + D_{\text{slow}}, \tag{30}$$

$$G_{\text{fast}}(s) = C_{\text{fast}} \left(sI - A_{\text{fast}}\right)^{-1} B_{\text{fast}} + D_{\text{fast}}. \tag{31}$$

The transfer functions for slow and fast subsystem are denoted by (30) and (31), respectively. The composite model is signified as a sum of slow and fast subsystem and a very little item $O(\varepsilon)$ [37]

$$G_{\text{composite}}(s) = G_{\text{slow}}(s) + G_{\text{fast}}(s) - D_{\text{slow}} + O(\varepsilon), \tag{32}$$

where

$$O(\varepsilon) = \begin{bmatrix} 0 & 0 \\ O_{21}(\varepsilon) & O_{22}(\varepsilon) \end{bmatrix}. \tag{33}$$

4. Multivariable PID Controller Design

Owing to the industrial process control involved with multivariable system, MPID controller design technique is necessary. It is a powerful control technique for solving coupling nonlinear system [38]. The conventional MPID controller designs technique is based on static inverse model [15]. This technique is difficult to obtain the desired control performance. Therefore, an enhancement is presented based on the dynamic inverse matrix and singularly perturbation method to the designs of MPID controller. Essentially, this enhancement has been considered in the previous work reported in [39] and it shows that the enhancement is able to control dynamic system where the output is able to track the set point change and produced less proses interaction. Nevertheless, it only considered that three controller designs instead of four and the selection of parameter tuning are done without optimization technique. In this paper, there are four enhanced MPID controller designs which are Davison, Penttinen-Koivo, Maciejowski, and Combined method where it is applied to the both original and singularly perturbed system with all of the parameter tuning being obtained based on particle swarm optimization. All of these four designs technique are applied to wastewater treatment plant and Newell and Lee evaporator.

4.1. Davison Method. Multivariable control design based on Davison method simply applies the integral term, which causes decoupling rise at low frequencies

$$K = K_i \frac{1}{s} \underline{e}(s). \tag{34}$$

The controller expression is represented by (34) [16], where K_i and $\underline{e}(s)$ are integral feedback gain and controller error, respectively,

$$K_i = \mu G(s)^{-1}. \tag{35}$$

Since this research is focused on dynamic control, K_i is defined as in (35), where μ is the only controller tuning parameter, which undoubtedly needs to be tuned progressively until the finest solution is discovered. Due to the involvement of the inverse system, the control design is only applicable for square matrix. If the system involves time delay, the time delay needs to be eliminated.

4.2. Penttinen-Koivo Method. This method is somewhat advanced and then the method proposed by Davison. In Penttinen-Koivo method, a proportional term is introduced. It comprises both integral and proportional term. Indirectly, this causes decoupling to take place at low and high frequencies. Davison and Penttinen-Koivo method are only similar in terms of an integral term which is linearly related to the inverse of plant dynamics

$$K = \left(K_p + K_i \frac{1}{s}\right) \underline{e}(s). \tag{36}$$

The controller expression is represented in (36) [16], where K_p, K_i, and $\underline{e}(s)$ are proportional gain, integral feedback gain, and controller error correspondingly

$$\begin{aligned} K_p &= (CB)^{-1} \rho, \\ K_i &= \mu G(s)^{-1}. \end{aligned} \tag{37}$$

Dynamic terms of K_p and K_i are expressed in (37), where ρ and μ are the tuning parameters for both proportional and integral feedback gain.

4.3. Maciejowski Method. Maciejowski method applies all proportional, integral, and derivative gains in its controller design. For maciejowski method, the tuning was done around

the bandwidth frequency, ω_B. Consequently, the interaction is reduced and good decoupling characteristic is provided around the frequency [15]. However, due to the needs of plant frequency analysis experiment, this method is quite difficult to be used throughout the industry [15]

$$K = \left(K_p + K_i \frac{1}{s} + K_d s\right) \underline{e}(s). \tag{38}$$

The controller expression is represented by (38), where K_p, K_i, K_d, and $\underline{e}(s)$ are proportional, integral feedback, derivative gains, and controller error

$$\begin{aligned} K_p &= \rho G(jw_b)^{-1}, \\ K_i &= \mu G(jw_b)^{-1}, \\ K_d &= \delta G(jw_b)^{-1}. \end{aligned} \tag{39}$$

Dynamic terms of K_p, K_i, and K_d are expressed in (39), where ρ, μ, and δ are Maciejowski tuning parameters. Due to a complex gain obtained from the calculation of $G(jw_b)^{-1}$, a real approximation of $G(jw_b)^{-1}$ is necessary which can be done by solving the following optimization problem:

$$\begin{aligned} M(N,\Theta) &= \left[G(jw_b)N - e^{j\Theta}\right]^T \left[G(jw_b)N - e^{j\Theta}\right], \\ \Theta &= \operatorname{diag}(\theta_1, \ldots, \theta_n), \end{aligned} \tag{40}$$

where N is a constant that is used to minimize M.

4.4. Combined Method. In order to overcome the weakness of the Maciejowski method which requires rigorous frequency analysis, a new method was proposed by Wahab et al. [15]. It is the result of the previous controllers where methods by Davison, Penttinen-Koivo, and Maciejowski are combined together:

$$K = \left(\rho Q + \mu Q \frac{1}{s}\right) \underline{e}(s). \tag{41}$$

Equation (41) represents the proposed control design, where ρ, μ, and $\underline{e}(s)$ are the tuning parameters and controller error:

$$Q = [\alpha G(s) + (1-\alpha) CB]^{-1} \tag{42}$$

Q is defined in (42). α is also a tuning parameter. This method keeps some properties in Maciejowski method but excludes the needs of frequency analysis [15].

5. Optimized Singularly Perturbed MPID Parameter Tuning

To ensure a fair comparison, the optimum parameter tuning for each of controller designs is measured by using particle swarm optimization (PSO). PSO optimizes a problem by having a population (swarm) of candidate solutions (birds) which is known as particles that are updated from iteration to iteration [40]. These particles are moved into the search space seeking for a food according to its own flying experience and its companion flying experience. It can be expended to multi-dimensional search. Each particle (solution) is characterized by its position and velocity, and every one of them searches for better positions within the search space by changing its velocity [41]. Each particle preserves the track of its current position within the search space. This value is identified as the particle's local best known position (*pbest*) and leads to the best known position (*gbest*), which is defined as enhanced positions that are found by the other particles. By that, the finest solution is attained

$$\begin{aligned} v(t+1) &= (w * v(t)) + (c_1 * r_1 * (p(t) - x(t))) \\ &\quad + (c_2 * r_2 * (g(t) - x(t))), \\ x(t+1) &= x(t) + v(t+1). \end{aligned} \tag{43}$$

Equation (43) represents the update equations of new velocity and new position, where $v(t+1)$, $x(t+1)$, w, $v(t)$, c_1, c_2, r_1, r_2, $p(t)$, $x(t)$, and $g(t)$ correspond to the velocity at time $t+1$, new particle position, inertia weight, current velocity at time t, cognitive weight, global weight, random variable within the range of $0 \geq r_1 < 1$, random variable within the range of $0 \geq r_2 < 1$, *pbest*, and *gbest*.

The overall performance of PSO can be increased by proper selection of inertia weight, w. Lower value of w provides a good ability for local search and higher value of w provides a good ability for global search [41]:

$$w = w_{\max} - \text{iter} \cdot \frac{w_{\max} - w_{\min}}{\text{iter}_{\max}}. \tag{44}$$

To achieve a respectable performance, w is determined according to (44), where $w_{\max}$ is the maximum value of inertia weight, $w_{\min}$ is the minimum value of inertia weight, iter is the current number of iteration, and $\text{iter}_{\max}$ is the maximum number of iteration. Most of the previous researchers have used $w_{\max} = 0.9$ and $w_{\min} = 0.4$, where significant enhancement of PSO is achieved [42, 43]:

$$\text{ITSE} = \int_0^T t e^2(t)\, dt. \tag{45}$$

Fitness function which is also known as cost function is represented by (45), where $e(t)$ is a system error. The procedure of PID parameter optimization by using PSO is summarized as follows:

(1) Initialization: initialize a population of particles with arbitrary positions and velocities on X dimensions in the problem space. Then, randomly initialize *pbest* and *gbest*.

(2) Fitness: calculate the desired optimization fitness function in X dimensions for every particle.

(3) *pbest*: compare calculated fitness function value for every particle in the population. If current value is smaller than *pbest*, and then update *pbest* as current particle position.

(4) *gbest*: determine the best success particle position among all of the individual best positions and designate as a *gbest*.

(5) New velocity and position: update the velocity and position of the particle based on (43).

(6) Repeat step (2) all over again until a criterion is encountered.

6. Case Studies

In the next subsections, an introduction to the case studies is presented. First, an overview of the wastewater treatment plant (WWTP) is provided and the Newell and Lee evaporator model is explained. These two case studies are considered to demonstrate the performance of the proposed methods.

6.1. Case Study I: Wastewater Treatment Plant (WWTP). Wastewater treatment plant is designed either for carbon removal or for carbon and nitrogen removal. In this project, the carbon removal scenario is considered. The control plant outputs are substrate and dissolved oxygen. Scanty provision of substrate affects the growth of microorganisms that are responsible for treating the wastewater and too many provisions lead to a drop in the microorganisms growth rate. The standard amount of substrate is around 51 mg/L [44]. Meanwhile, insufficient dissolved oxygen will cause the degradation of the pollutants and the plant to become less efficient. Too much dissolved oxygen can cause excessive consumption of energy where it will increase the cost for the treatment. Other than that, it also can cause too much sludge production. The amount of dissolved oxygen concentration needs to be controlled so that it is in the range of 1.5 mg/L–4.0 mg/L [45]. The aim of this case study is to control the concentration of substrate and dissolved oxygen at the desired value by manipulating the manipulate variables of dilution rate and air flow rate, respectively. The state space form of the nonlinear wastewater treatment plant is linearized from [39] as follows:

$$\begin{bmatrix} \dot{X} \\ \dot{S} \\ \dot{\mathrm{DO}} \\ \dot{X}_r \end{bmatrix} = A \begin{bmatrix} X \\ S \\ \mathrm{DO} \\ X_r \end{bmatrix} + B \begin{bmatrix} D \\ W \end{bmatrix} \tag{46a}$$

$$\begin{bmatrix} Y_S \\ Y_{\mathrm{DO}} \end{bmatrix} = C \begin{bmatrix} X \\ S \\ \mathrm{DO} \\ X_r \end{bmatrix} + D \begin{bmatrix} D \\ W \end{bmatrix}, \tag{46b}$$

where the state is composed by X, the biomass, S, the substrate, DO, the dissolved oxygen, and X_r, the recycled biomass. The input variables are D, the dilution rate, and W, an air flow rate. Matrices A, B, C, and D are given by

$$A = \begin{bmatrix} -0.0990 & 0.1234 & 0.2897 & 0.0495 \\ -0.0508 & -0.3219 & -0.4457 & 0 \\ -0.0254 & -0.0949 & -1.9748 & 0 \\ 0.1320 & 0 & 0 & -0.0660 \end{bmatrix}, \tag{47a}$$

$$B = \begin{bmatrix} -87.1159 & 0 \\ 134.0243 & 0 \\ -9.2834 & 0.0699 \\ 0.0001 & 0 \end{bmatrix}, \tag{47b}$$

$$C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \tag{47c}$$

$$D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{47d}$$

These A, B, C, and D matrices are used in the design of singularly perturbed MPID control and tested into the nonlinear wastewater treatment plant.

6.2. Case Study II: Newell and Lee Evaporator. This subsection presents the Newell and Lee evaporator system which is considered as a second case study. The objective is to evaluate the effectiveness and the performance of the proposed singularly perturbed MPID controllers for different system. Here, unstable system is considered. Similar to the first case study, the four different methods of MPID are implemented, which is Davison, Penttinen-Koivo, Maciejowski, and Combined methods. The plant to be controlled is given by the following state space model [46]:

$$\begin{bmatrix} \dot{L}_2 \\ \dot{X}_2 \\ \dot{P}_2 \end{bmatrix} = A \begin{bmatrix} L_2 \\ X_2 \\ P_2 \end{bmatrix} + B \begin{bmatrix} F_2 \\ P_{100} \\ F_{200} \end{bmatrix}, \tag{48a}$$

$$\begin{bmatrix} Y_{L_2} \\ Y_{P_2} \end{bmatrix} = C \begin{bmatrix} L_2 \\ X_2 \\ P_2 \end{bmatrix} + D \begin{bmatrix} F_2 \\ P_{100} \\ F_{200} \end{bmatrix}, \tag{48b}$$

where the state is composed by L_2, the separator level, X_2, the product composition, and P_2, an operating pressure. The input variables are F_2, the product flow rate, P_{100}, the steam pressure, and F_{200}, the cooling water flow rate. The outputs to be controlled are Y_{L_2}, separator level, and Y_{P_2}, operating pressure. Matrices A, B, C, and D are given by

$$A = \begin{bmatrix} 0 & 0.00418 & 0.007512 \\ 0 & -0.10000 & 0 \\ 0 & -0.02091 & -0.05580 \end{bmatrix}, \tag{49a}$$

$$B = \begin{bmatrix} -0.05000 & -0.00192 & 0 \\ -1.25000 & 0 & 0 \\ 0 & 0.00959 & -0.00183 \end{bmatrix}, \tag{49b}$$

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{49c}$$

$$D = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{49d}$$

7. Results and Discussion

This section presents the results and discussion for both case studies. It will include the results of singularly perturbed MPID control based on four proposed methods, which are Davison, Penttinen-Koivo, Maciejowski, and Combined method for both full order and singularly perturbed system. The first section shows the MPID control based on particle swarm optimization. The obtained optimum tuning parameters are presented. The second section shows the simulation results for control of the closed loop system during substrate and dissolved oxygen set point change, while the last section provides the result which shows the stability of the closed loop system.

7.1. Results for the Case Study 1: Wastewater Treatment Plant (WWTP). The eigenvalue of the open loop wastewater treatment plant is as follows:

$$\begin{aligned} e(A) &= \{-0.0076, -0.2009, -0.2579, -1.9953\}, \\ e(A_s) &= \{-0.0076, -0.2009, -0.2579\}, \\ e(A_f) &= \{-1.9953\}. \end{aligned} \tag{50}$$

As a result

$$\begin{aligned} \varepsilon &= \frac{|-0.2579|}{|-1.9953|} \\ &= 0.1293 \ll 1. \end{aligned} \tag{51}$$

Since ε is less than 1, the system is said to possess a two-time scale characteristic. The eigenvalue at -1.9953 is considered as a fast response

$$A_{\text{SPS/Naidu}} = \left[\begin{array}{ccc|c} 0 & 0.1234 & 0.2897 & 0 \\ -0.0508 & -0.3219 & -0.4457 & 0 \\ -0.0254 & -0.0949 & -1.975 & 0 \\ \hline 0 & 0 & 0 & -0.0660 \end{array}\right] \tag{52a}$$

$A_{11} = A_{\text{slow}}$, $A_{12} = Z_{12}$, $A_{21} = Z_{21}$, $A_{22} = A_{\text{fast}}$

$$B_{\text{SPS/Naidu}} = \left[\begin{array}{cc} -87.1159 & 0 \\ 134.0243 & 0 \\ -9.2834 & 0.0699 \\ \hline 0.0001 & 0 \end{array}\right] \tag{52b}$$

$B_1 = B_{\text{slow}}$, $B_2 = B_{\text{fast}}$

$$C_{\text{SPS/Naidu}} = \left[\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \tag{52c}$$

$C_1 = C_{\text{slow}}$, $C_2 = C_{\text{fast}}$

$$D_{\text{SPS/Naidu}} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \longleftarrow D. \tag{52d}$$

By using algorithm discussed in Section 3.1, the original system which refers to (47a) to (47d) is represented in state space form of singularly perturbed system as indicated in (52a) to (52d). This state space form is used to design the controller tuning, while the simulation and performance of the system are based on the original system. From the state space form in (52a) to (52d), it is clearly shown that the eigenvalues of the system are grouped into two distinct and separate sets, which causes the time consumed to obtain the MPID tuning parameters reduce. All eigenvalues of the singularly perturbed system are remained at the left-half plane, which is -1.9955, -0.2798, -0.0214, and -0.0660. It is indicated that the system is boundary input boundary output (BIBO) stable

$$G_{\text{Jian Niu}-S/D} = G_{\text{Jian Niu}-11}(s) = \frac{s + 134.1}{s + 0.0508}, \tag{53a}$$

$$G_{\text{Jian Niu}-C/D} = G_{\text{Jian Niu}-21}(s) = \frac{s + 134.1}{s + 0.0508}, \tag{53b}$$

$$G_{\text{Jian Niu}-S/W} = G_{\text{Jian Niu}-12}(s) = -0.0, \tag{53c}$$

$$G_{\text{Jian Niu}-C/W} = G_{\text{Jian Niu}-22}(s) = -0.05. \tag{53d}$$

Based on the algorithms explained in Section 3.2, the Jian Niu method is successfully able to represent the original system into composite of singularly perturbed system as in (53a) to (53d), where the only existing eigenvalue is located at -0.0508. Since it is in the left-half plane, the system is BIBO stable. To validate the models from both methods, singularly perturbation method obtained by Naidu and Jian Niu, the magnitude and phase plot between the original system and singularly perturbed system are plotted as shown in Figure 1.

Based on Figure 1, singularly perturbed system based on Naidu algorithm provides better dynamic response compared to the singularly perturbed system based on Jian Niu algorithm, where the dynamic response is much identical to the original system. The close approximation between the original system and singularly perturbed system by Naidu in the frequency responses analysis exhibits the authority of model and essentially leads to adequate control performance of the controller design. Hence, a singularly perturbed system by Naidu is used in dynamic matrix inverse of MPID controller designs.

In the considered WWTP, there are two controlled variables and two manipulated variables. Interaction phenomena may occur between these two controls and manipulate variables. Each manipulated variable can affect both the control variables. The process interaction among the variables may cause the closed loop system to become destabilized and the controller tuning is more challenging. In order to minimize the process interaction, the selection of suitable control and manipulated variables pairing is importance. In this case, there are two possible controller pairings.

The relative gain array (RGA) analysis has been used in quantifying the level of interactions in a multivariable system. It is also used to determine the best input output pairing and that pairing should be avoided. To measure the ability of RGA in providing a realistic pairing recommendation, the RGA for

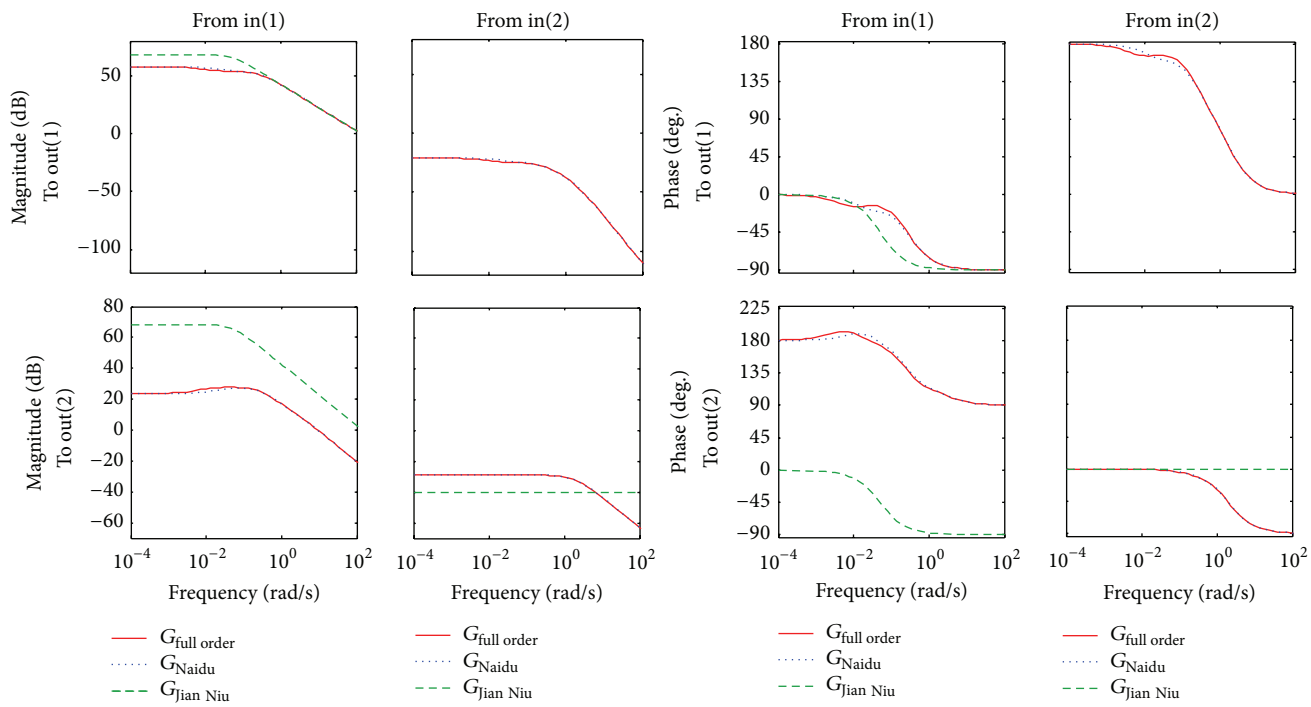


Figure 1: Bode analysis for different methods.

linearized models was calculated. RGA of a nonsquare matrix is defined in (54). The results are displayed in a matrix form, where columns are for each input variable and rows for each output variable. This matrix form can be used in determining which relative gains are associated to which input output variables

$$\text{RGA} = \Lambda = G_{\text{RGA}} \times \left(G_{\text{RGA}}{}^{\dagger}\right)^{T}, \tag{54}$$

where

$$G_{\text{RGA}} = \text{Gain matrix} = \begin{bmatrix} -87.1159 & 0 \\ 134.0243 & 0 \\ -9.2834 & 0.0699 \\ 0.0001 & 0 \end{bmatrix},$$
$$G_{\text{RGA}}{}^{\dagger} = \text{Pseudo inverse of gain matrix} = \begin{bmatrix} -0.0034 & 0.0052 & 0.0000 & 0.0000 \\ -0.4528 & 0.6966 & 14.3062 & 0.0000 \end{bmatrix}. \tag{55}$$

Therefore

$$\text{RGA} = \Lambda = \begin{bmatrix} 0.2970 & 0 \\ 0.7030 & 0 \\ -0.0000 & 1.0000 \\ 0.0000 & 0 \end{bmatrix}. \tag{56}$$

From the RGA obtained, it can be concluded that dissolved oxygen cannot be paired with dilution rate due to the negative relative gain. It corresponds to the worst case, and this is highly undesirable. Biomass, substrate, and recycle biomass on the other hand cannot be paired with air flow rate due to the zero relative gain, which means that air flow rate does not have any effect on biomass, substrate, and recycle biomass. In RGA analysis, the closer the value of RGA element to one is, the configuration is more likely to work, where less interaction exists. Hence, it is concluded that a good pair of dissolved oxygen and air flow rate and substrate and dilution rate are recommended.

Since MPID controller designs are involved with several tuning parameter, PSO was adopted. Due to the PSO characteristic which cannot give a unique solution at every attempt [47], 10 trials of simulation for original and singularly perturbed system of each MPID controller design were conducted. A result with minimum error was selected. Table 1 shows the obtained optimum PID tuning parameter using ITSE fitness function for both systems: original and singularly perturbed system. The results are corresponding to Davison, Penttinen-Koivo, Maciejowski, and Combined method, respectively. Based on the results presented, it clearly shows that singularly perturbed system is able to provide easiness in tuning strategy in terms of computation time where it required less computation time compared to the original system. Table 1 shows that singularly perturbed MPID based on each method is able to reduce more than half of computation time required by original system.

Figures 2 to 5 show the comparison between output response and interaction based on Penttinen-Koivo method for original and singularly perturbed system. Based on the Figures 2 and 4, Penttinen-Koivo based on singularly perturbed system is able to produce better output responses

Table 1: Optimum PID parameter for WWTP based on PSO.

Method	Original system				Singularly perturbed system			
	α	μ	ρ	Time (s)	α	μ	ρ	Time (s)
Davison	—	0.9723		1173.7	—	0.7949		355.5019
Penttinen-Koivo	—	1.7730	0.5741	49.7979	—	1.2010	0.7925	10.4767
Maciejowski	—	9.8842	5.3998	247.4416	—	1.6729	9.4042	69.9072
Combined	0.8843	7.8907	9.5949	370.6488	0.7927	9.4458	7.5924	132.5115

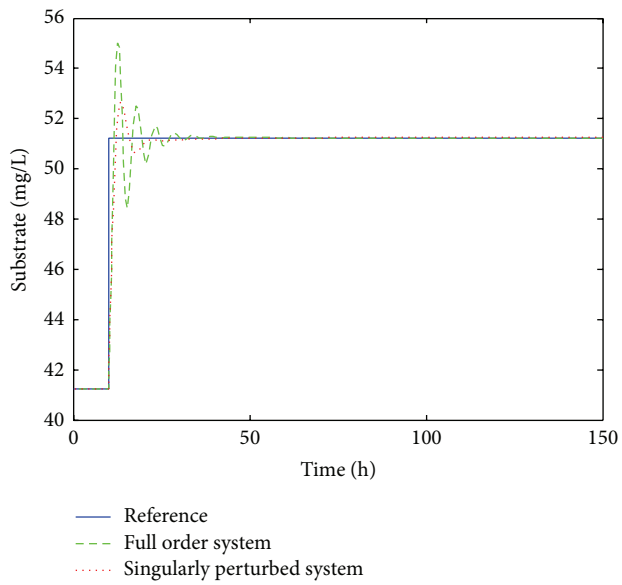


Figure 2: Substrate responses between original and singularly perturbed system during set point change.

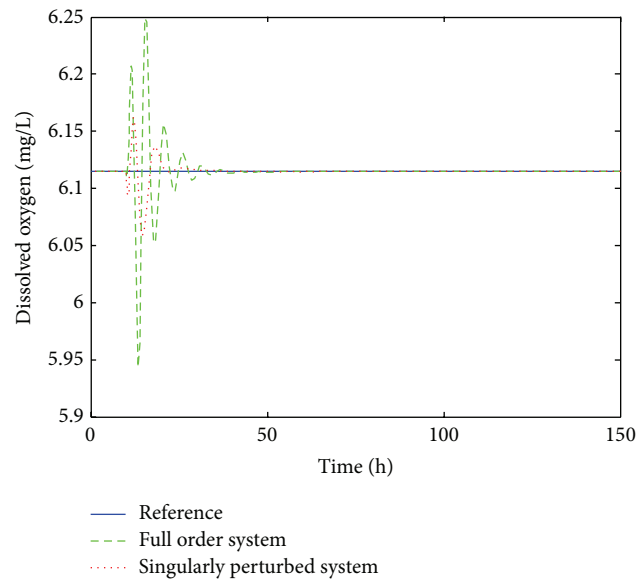


Figure 3: Process interaction between original and singularly perturbed system during substrate change.

with less overshoot and fast settling time. Meanwhile, more oscillation is produced by output responses based on original system. The interaction among the output variables for Penttinen-Koivo based on singularly perturbed system is also reduced as shown in Figures 3 and 5.

Due to the better closed loop performance, reduced process interaction, and less time consuming obtained by the Penttinen-Koivo method based on singularly perturbed system compared to the original system, singularly perturbed system was implemented thoroughly in the case of the controller performance evaluations among others three methods accordingly. To measure the performance quality of four singularly perturbed MPID controller designs, pseudorandom binary sequence (PRBS) was injected as the input signal to the system. PRBS was injected to determine and to test the tracking ability of the proposed singularly perturbed MPID controller designs for each step change. From the responses obtained, all four designs are able to track each step change where the Combined method provides the best response. Figures 6 and 7 show the output responses for both substrate and dissolved oxygen concentration.

To provide a clear view of the set point tracking ability and process interaction, the system was also injected with

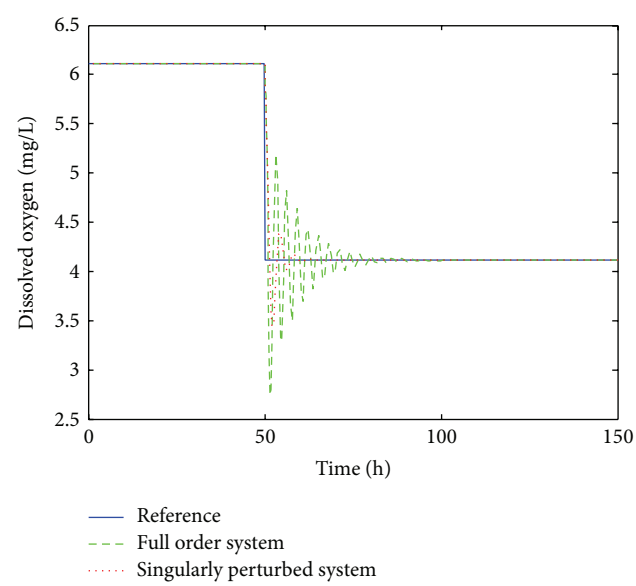


Figure 4: Dissolved oxygen responses between original and singularly perturbed system during set point change.

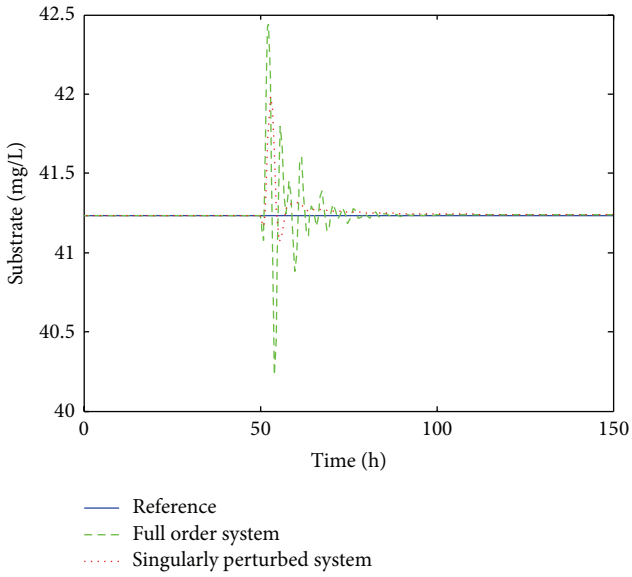


Figure 5: Process interaction between original and singularly perturbed system during dissolved oxygen change.

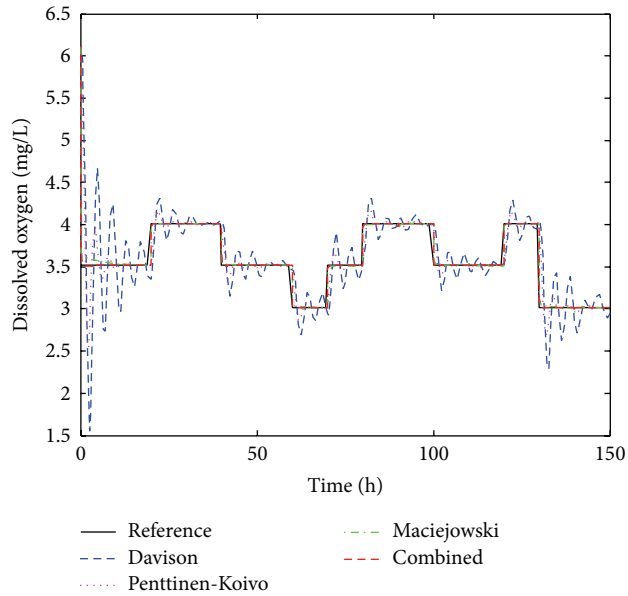


Figure 7: Concentration of dissolved oxygen based on PRBS input.

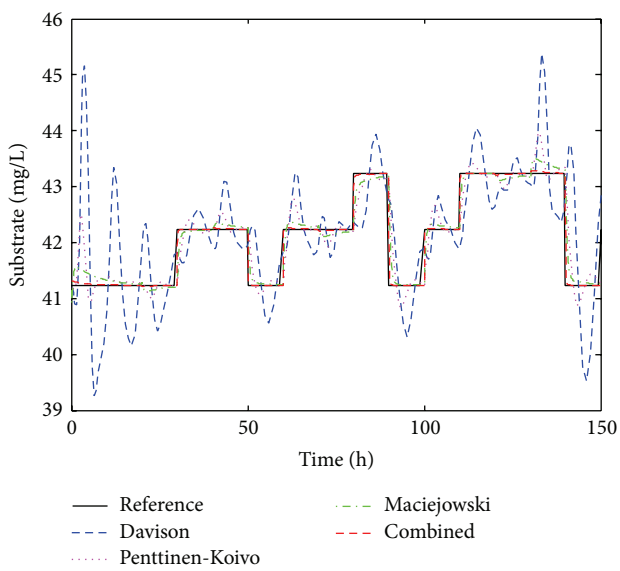


Figure 6: Concentration of substrate based on PRBS input.

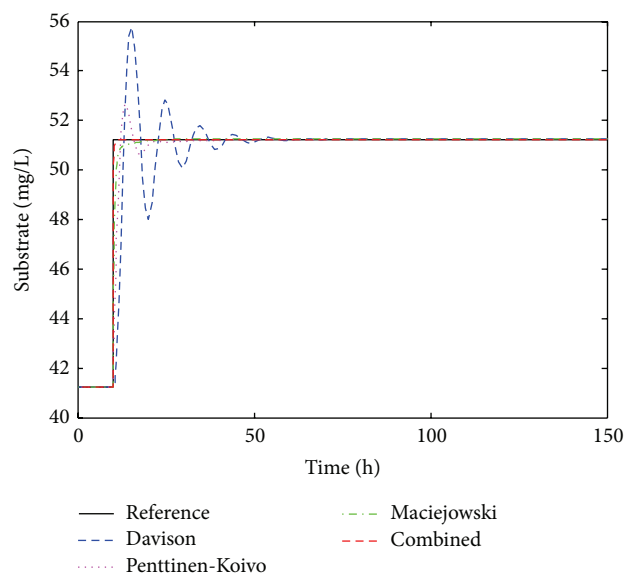


Figure 8: Closed loop responses of substrate during substrate set point change.

step input. The simulation was carried out during substrate and dissolved oxygen set point change. For each change, the step input was injected at $t = 10$ h and $t = 50$ h, respectively. Figures 8 and 9 show the simulation results for substrate and dissolved oxygen for each proposed singularly perturbed MPID controller design. The responses are set with respect to the step change in the substrate input from 41.2348 mg/L to 51.2348 mg/L.

Figure 8 shows that all singularly perturbed MPID controller designs are able to keep the concentration of the substrate close to the desired value. It shows that the control based on Combined method is able to provide the finest control effect among the others' method in terms of settling time and maximum amplitude, where it is able to achieve settling point during 11 h compared to the others which settle during 44 h, 21 h, and 14 h, respectively. Due to the control characteristic which only applies integral gain, control action based on Davison method provides a response with the highest percentage of overshoot (%OS). By using Penttinen-Koivo method, the output response shows better improvement. The presence of both integral and proportional gain is able to minimize the percentage of overshoot (%OS) and offer a better settling time (T_s). However, proportional gain needs to be tuned wisely. High value of proportional gain can cause the system to become unstable, while small value of

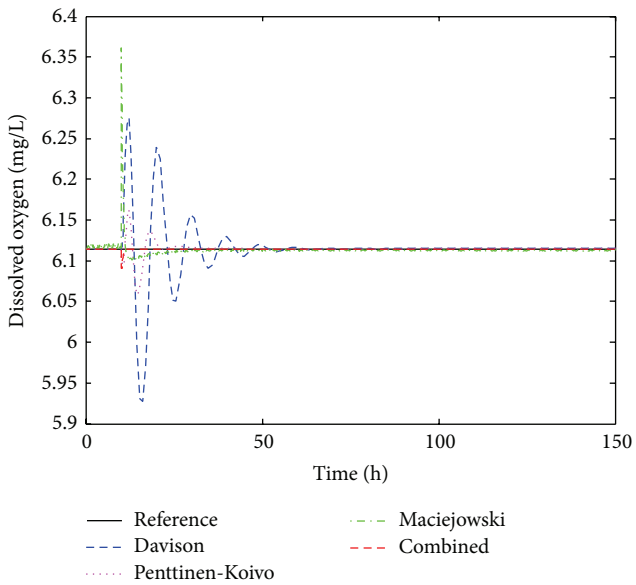


Figure 9: Process interactions during substrate set point change.

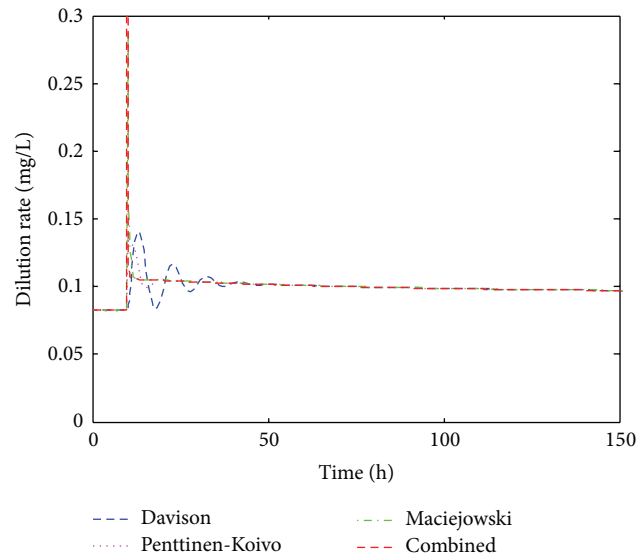


Figure 10: Dilution rate responses during substrate set point change.

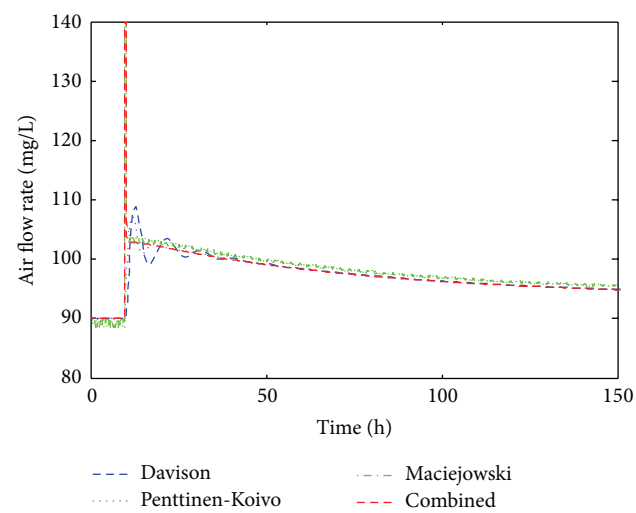


Figure 11: Air flow rate responses during substrate set point change.

proportional gain may reduce the sensitivity of the controller. Compared to the Davison and Penttinen-Koivo method, Maciejowski method gives the best performance with small percentage of overshoot (%OS) and faster settling time (T_s). It is proven that control performance at the selected frequency was able to improve the closed loop response. An important feature in Maciejowski method is the selection of frequency. Frequency must be selected properly to avoid instability. Among all methods, the Combined method exhibited the best tracking to the substrate changes. This method exhibits a faster response than other control designs, but it requires a long time to obtain the tuning parameters.

Since the considered case study involves multivariable system, process interaction may occur. Interactions between the system variables occur because each manipulated variable in the multivariable system certainly will affect the controlled variables. Here, dilution rate and air flow rate will affect the response of both substrate and dissolved oxygen concentration. Evidently, when changing one of the inputs for dilution rate or air flow rate, both outputs will be affected, and this means that there is significant coupling in the system. Figure 9 shows the interaction responses for each proposed singularly perturbed MPID controller design that occurs during substrate change. The response obtained has proven that substrate and dissolved oxygen are coupled since the step changes in the substrate disturb the dissolved oxygen correspondingly. If there is no process interaction, dissolved oxygen should not be affected when the substrate is changed. Fortunately, process interaction was reduced for each controller design where the Combine method provides less interaction, which indicates by the lowest maximum amplitude and less oscillation.

Figures 10 and 11 show the closed loop responses of manipulate variable, which are dilution rate and air flow rate during the substrate set point change, respectively. Large variation exhibits in the response of dilution rate and air flow rate based on Davison method. Penttinen-Koivo and Combined method exhibit moderate variations, while Combined method exhibits smallest variations but with high peak value at one of the points.

Figures 12 and 13 show the simulation results for dissolved oxygen and substrate to the step change in the dissolved oxygen input. From Figure 12, it is clearly shown that the trajectory of the output is improved by each method that has been applied. It is proportionate with the responses during substrate set point change. All the systems step response settles at a final value of 4.1146 mg/L, which is the final value of the unit step input. Figure 12 shows that the responses produced by the Davison and Penttinen-Koivo methods do not asymptotically approach the final value, where overshoot appears in the final value. The maximum values of the outputs

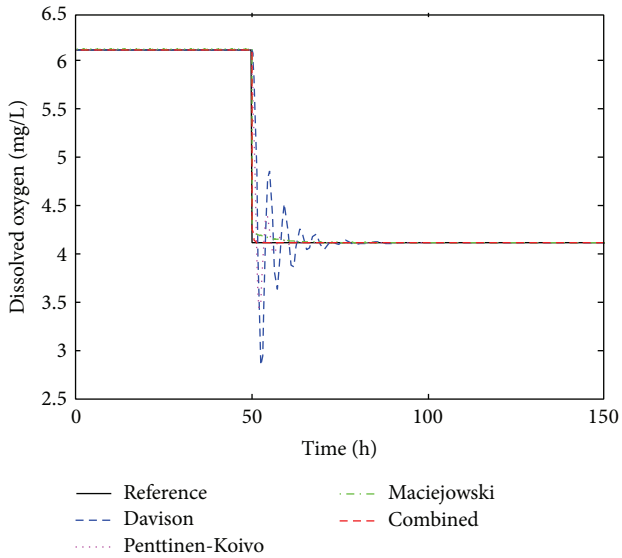


FIGURE 12: Closed loop responses of dissolved oxygen during dissolved oxygen set point change.

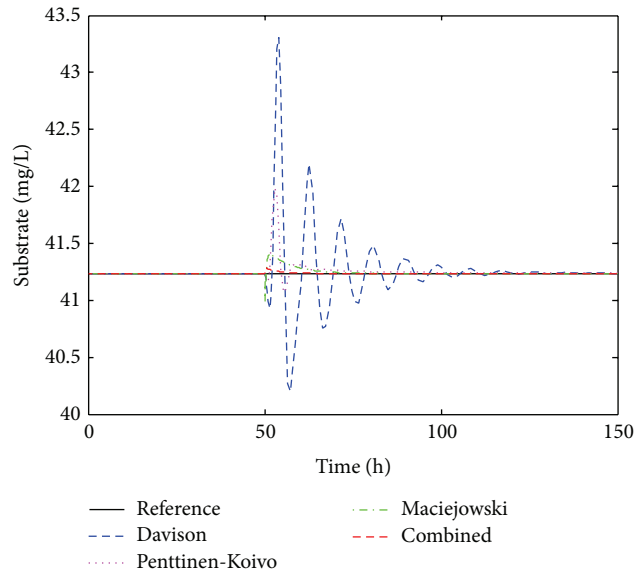


FIGURE 13: Process interactions during dissolved oxygen set point change.

are 4.8649 mg/L and 4.2830 mg/L, respectively, for each of Davison and Penttinen-Koivo methods. The output response yield by Davison method consists of 37.52% of overshoots. However, Penttinen-Koivo method has improved the output response by reducing the overshoot to 8.42%. Meanwhile, the responses acquired from Maciejowski and Combined methods asymptotically approach the final value. These methods provide slightly similar effect in terms of maximum amplitude and settling time. There is no overshoot of the final value and there are no oscillations in the response. The outputs reach the final value at around $t = 0.2$ h and $t = 0.4$ h for both Maciejowski and Combined methods, respectively. Figure 13 shows that interactions also occur during dissolved oxygen set point change. Similar to the responses during substrate set point change, process interaction exists and it is improved by each method proposed by Davison up to the Combined method, respectively.

Figures 14 and 15 show the closed loop responses of manipulate variable, which are dilution rate and air flow rate during the dissolved oxygen set point change, respectively. The characteristic of the closed loop responses for each of singularly perturbed MPID controller design is summarized in Table 2.

The stability of a system can be determined directly from its transfer function or from CLCP. Figure 16 shows the closed loop poles and zeros plot for each singularly perturbed MPID control design. It is mark a pole location by a cross (x) and a zero location by a circle (o). Based on the plot figure, all poles are located on the left-half plane that guarantees a stable system. However, to ensure the reliability of the stability analysis, Routh-Hurwitz analysis was performed. The results show that all methods are able to produce a stable system.

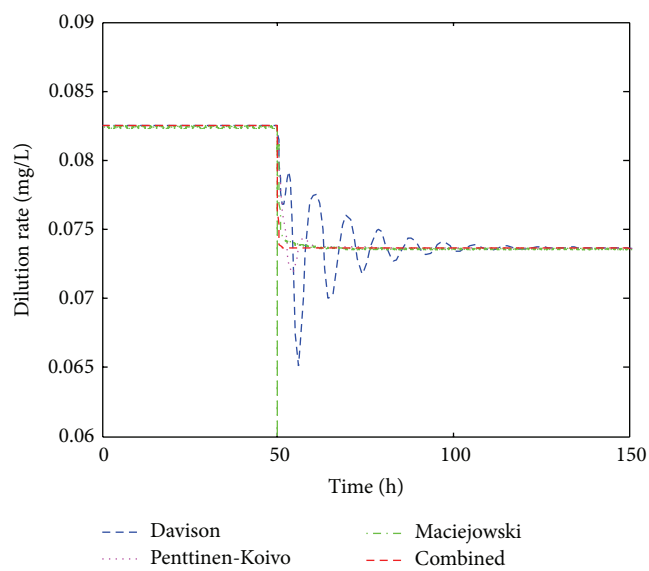


FIGURE 14: Dilution rate responses during dissolved oxygen set point change.

7.2. Results for the Case Study II: Newell and Lee Evaporator. The eigenvalue of the open loop Newell and Lee evaporator is as follows:

$$\begin{aligned} e(A) &= \{0, -0.0558, -0.1000\}, \\ e(A_s) &= \{0\}, \\ e(A_f) &= \{-0.0558, -0.1000\}. \end{aligned} \tag{57}$$

Table 2: Characteristic of closed loop response for WWTP.

Output	Method	Rise time, T_r (h)	Settling time, T_s (h)	Percentage overshoot (%OS)	Steady state error (%)
Substrate, S	Davison	2.0	44	45.65	0.152
	Penttinen-Koivo	1.8	21	8.65	0.092
	Maciejowski	0.9	14	5.65	0.032
	Combined	0.1	11	0	0.012
Dissolved oxygen, DO	Davison	1.0	14	62.49	0.03
	Penttinen-Koivo	1.2	5.5	21.58	0.03
	Maciejowski	0.1	0.2	0	0.17
	Combined	0.1	0.4	0	0.02

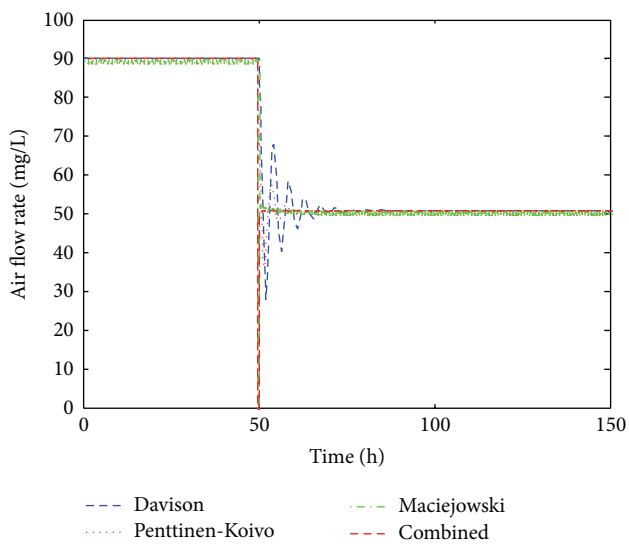


Figure 15: Air flow rate responses during dissolved oxygen set point change.

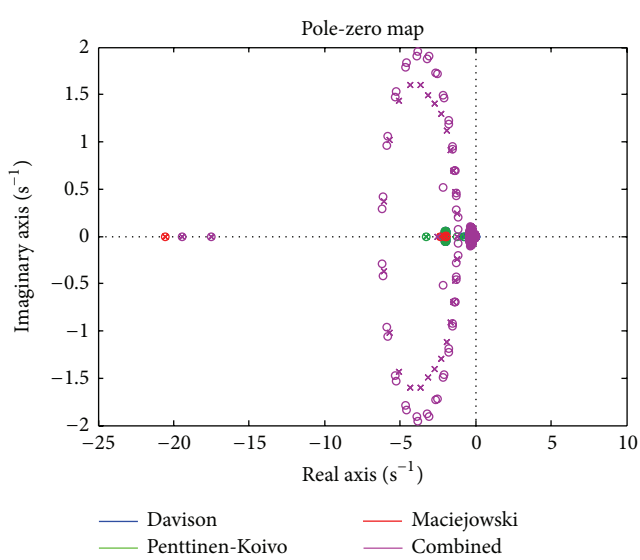


Figure 16: Closed loop pole-zero plot.

As a result

$$\varepsilon = \frac{|0|}{|-1.9953|} = 0 \ll 1. \tag{58}$$

Since ε is less than 1, the system behaved as a two-time scale characteristic. There is one slow variable which is indicated by eigenvalue of 0 and two fast variables which are indicated by the eigenvalue at -0.0558 and -0.1000. Based on the algorithm discussed in Section 3.1, the original system of Newell and Lee evaporator can be represented in singularly perturbed system. The eigenvalues for singularly perturbed system are $0, -0.0558$, and -0.1000, which is similar to the original system

$$A_{11} = A_{\text{slow}} \qquad A_{12} = Z_{12}$$

$$A_{\text{SPS/Naidu}} = \left[\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & -0.1 & 0 \\ 0 & -0.02091 & -0.0558 \end{array}\right] \tag{59a}$$

$$A_{21} = Z_{21} \qquad A_{22} = A_{\text{fast}}$$

$$B_1 = B_{\text{slow}}$$

$$B_{\text{SPS/Naidu}} = \left[\begin{array}{cc} -0.06706 & -0.0002464 \\ \hline -1.25000 & 0 \\ 0 & -0.00183 \end{array}\right] \tag{59b}$$

$$B_2 = B_{\text{fast}}$$

$$C_1 = C_{\text{slow}}$$

$$C_{\text{SPS/Naidu}} = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 0 & 1 \end{array}\right] \tag{59c}$$

$$C_2 = C_{\text{fast}}$$

$$D_{\text{SPS/Naidu}} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \longleftarrow D. \tag{59d}$$

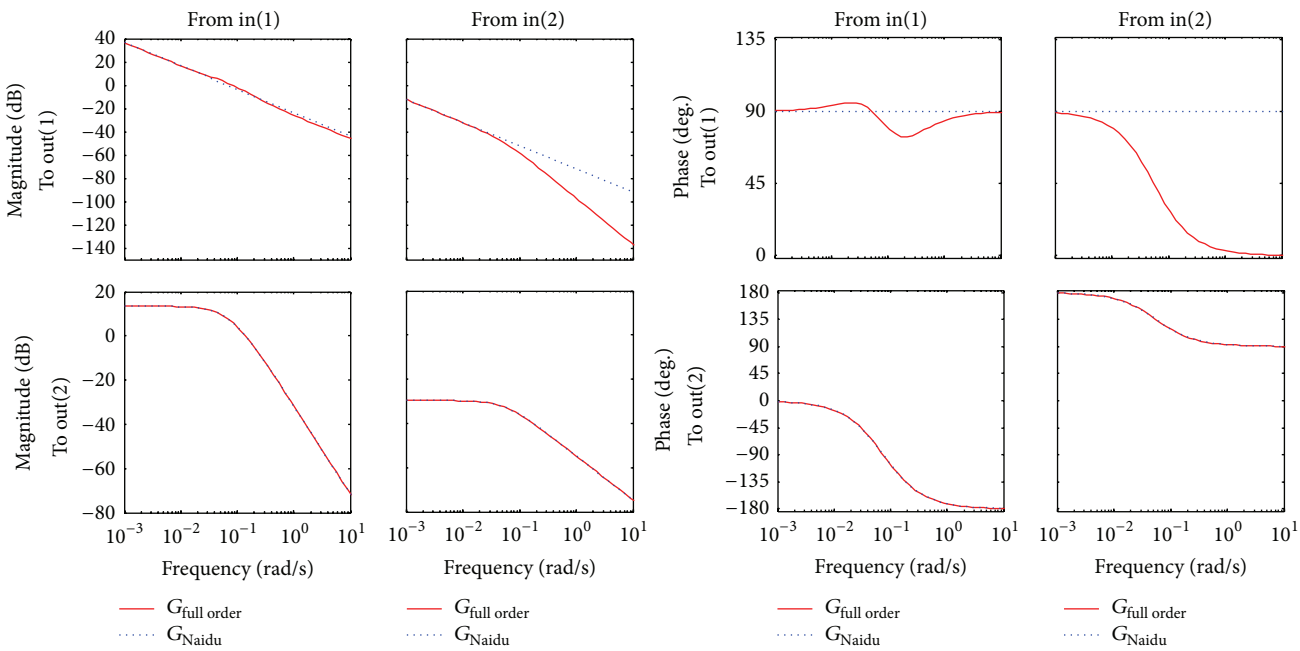


Figure 17: Bode analysis for original and singularly perturbed system.

Equations (59a) to (59d) represent the singularly perturbed system of Newell and Lee evaporator in state space form. To verify the singularly perturbed system with the original system, the Bode diagram is plotted as shown in Figure 17. As seen from Figure 17, singularly perturbed system shows a fairly good tracking with the original system over the low, middle, and high frequencies ranges. The response of the singularly perturbed system is almost identical to the original system. The close approximation between both systems demonstrates the validity of the obtained singularly perturbed system and principally leads to satisfactory control performance.

In this case, there also exists two possible control and manipulate variables paring. By using (54), the RGA for Newell and Lee evaporator was obtained as

$$\text{RGA} = \Lambda = \begin{bmatrix} 0.0028 & 0.0178 \\ 0.9972 & 0 \\ 0 & 0.9822 \end{bmatrix}. \tag{60}$$

Based on the analysis of RGA, it can be concluded that separator level cannot be paired with cooling water flow rate, and operating pressure cannot be paired with product flow rate, respectively. This is due to the zero relative gain. A change of cooling water flow rate will not give any significance to the separator level, and change of product flow rate will not give any significance to the operating pressure. From the RGA analysis, it is highly recommended to pair product flow rate with separator level and cooling water flow rate is paired with operating pressure. Since the value is nonzero and positive, the pairing is possible.

Similar as in case study I, 10 trials of PSO simulation for original and singularly perturbed system of each MPID controller design were conducted. However, due to the unstable open loop response, Maciejowski method cannot be implemented to this multivariable system as it requires information from stable open loop response. Table 3 shows the obtained optimum PID parameter based on PSO for evaporator system. By applied Davison method, both systems are able to provide similar tuning parameter with similar error for 10 number of run. However, it can be seen the advantage of singularly perturbed system which required less computation time compared to the original system. For Penttinen-Koivo and Combined method, the computation time is reduced more than triple times with the adaptation of singularly perturbed system in MPID control.

Figures 18 to 21 show the comparison between output responses based on Penttinen-Koivo method for each Newell and Lee evaporator, original and singularly perturbed system. Figure 18 shows the separator level responses between original and singularly perturbed system during separator level change. The response belonging to the original system has a poor performance as compared to the singularly perturbed system with high oscillation during step up and step down response. The output response achieved by the singularly perturbed system is with less overshoot and fast settling time. Figure 19 shows the interaction response during separator level change. It can be seen that the interaction was reduced by the adaptation of singularly perturbed system in MPID control. Figure 20 shows the operating pressure response between original and singularly perturbed system during operating pressure change. It can be observed that the singularly perturbed system has better performance as compared to the original system with fast settling time, while Figure 21 shows the interaction response during operating pressure change. Due to the good responses exhibited from the

Table 3: Optimum PID parameter for evaporator system based on PSO.

Method	Original system				Singularly perturbed system			
	α	μ	ρ	Time (s)	α	μ	ρ	Time (s)
Davison	—	0.0300	—	746.1308	—	0.0300	—	134.7652
Penttinen-Koivo	—	0.0500	0.0726	147.8473	—	0.0500	2.2488	22.0722
Combined	0.0223	0.1000	1.8896	568.1792	0.0363	0.7703	0.7609	167.2869

Table 4: Step point change.

	Separator level change		Operating pressure change	
Step time	120	250	270	400
Initial value	0	0	0	0
Final value	1.5	−1.5	0.2	−0.2

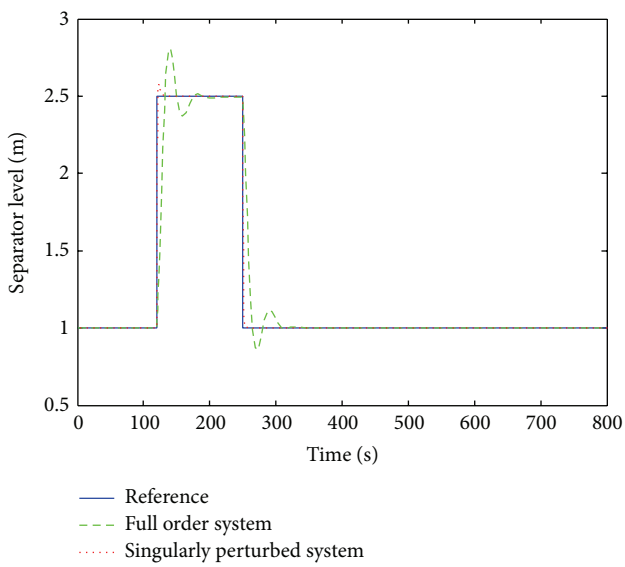


Figure 18: Separator level responses between original and singularly perturbed system.

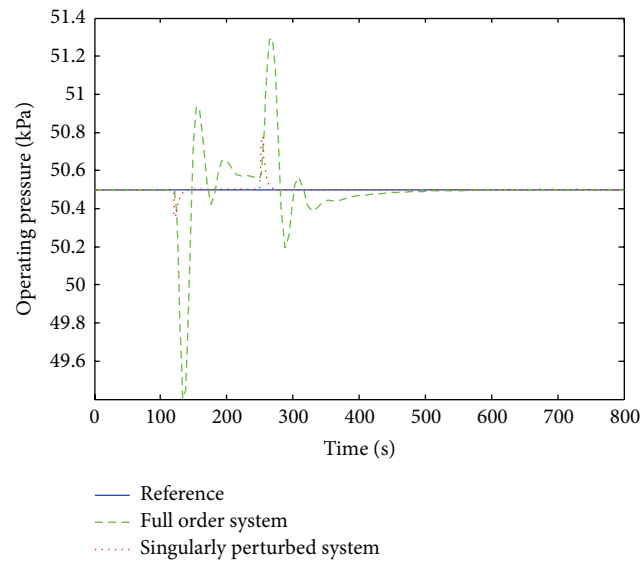


Figure 19: Process interaction between original and singularly perturbed system during separator level change.

control based on singularly perturbed system, this system was implemented thoroughly into the Davison, Penttinen-Koivo, and Combined method. The simulation was performed by setting the set point input values as follows:

value of separator level: 1 m,

value of operating pressure: 50.5 kPa.

The simulation was executed in two-variable change, which is during separator level and operating pressure step point change. The step point change of the separator level and operating pressure is listed in Table 4.

Figures 22 and 23 show the closed loop performance for the separator level and operating pressure to the sequential step point changes in separator level set point. The step point changes were sequentially introduced into the system at $t = 120$ s and $t = 250$ s, respectively. In the simulation study, the comparisons of the closed loop performances were done between Davison, Penttinen-Koivo, Combined, and multivariable controller designed proposed by Fauzi, which is based on multiobjective optimization approach using surrogate modelling. Fauzi's method was compared as they also have been involved with the similar Newell and Lee evaporator system. However, the details concerning to the controller designed are not presented and can be referred in [46]. Figure 22 shows that the output performance based on Davison method provides a response with 11.67% and 8.02% overshoot during the step up and step down input, respectively. These values are higher compared to other controller methods. However, it is still able to track the set point input given. The output performances based on Penttinen-Koivo and Combined methods are quiet similar. Combined method provides faster rise and settling time during the step up input, whereas the Penttinen-Koivo provides faster rise and settling time during the step down input. However, the Combined method provides the best performance with low percentage of overshoot, which specify by the lowest value of maximum amplitude and steady state error. Among the four methods, the method proposed by Fauzi is the poorest. At the early stage, the response is relatively good. Once the step down input is injected, the response shows unstable characteristic where it fails to settle at the given set point input. The response is considered as unstable since the gain error increases as time increases. Figure 23 compares the interactions that occur during the separator level set point change. It is clearly shown that Davison method produces

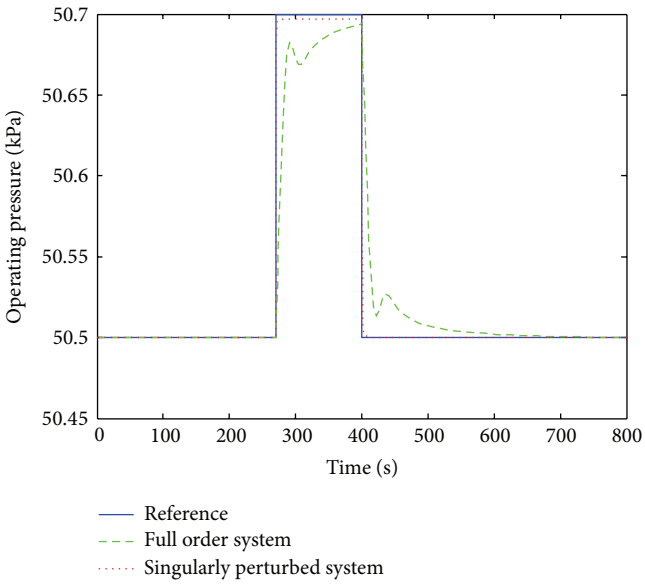


Figure 20: Operating pressure responses between original and singularly perturbed system.

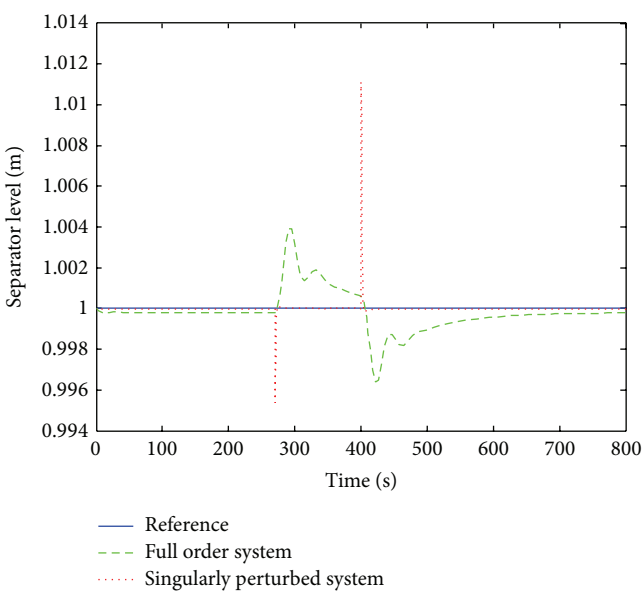


Figure 21: Process interaction between original and singularly perturbed system during operating pressure change.

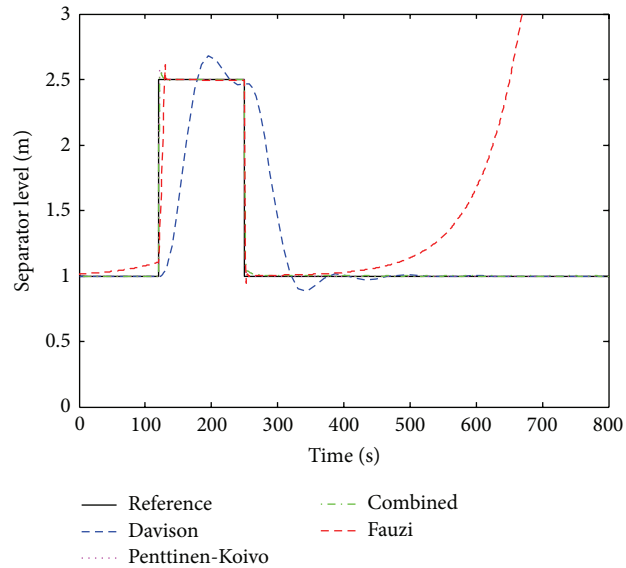


Figure 22: Responses of separator level during separator level set point change.

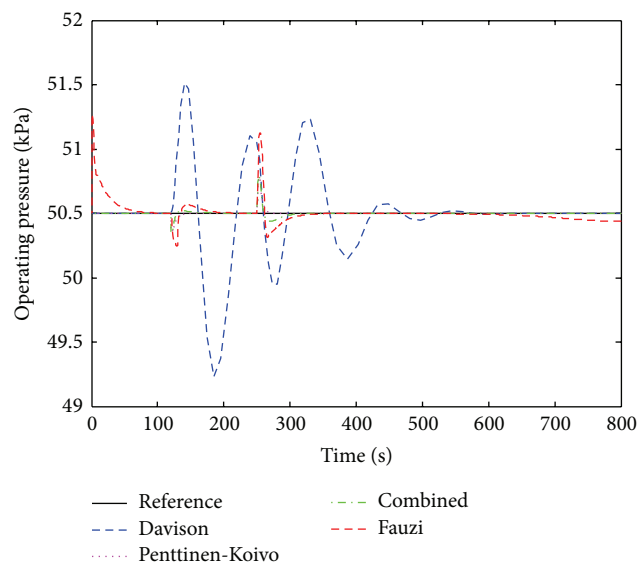


Figure 23: Process interactions during separator level set point change.

large interactions with high maximum amplitude and more oscillation. The sluggishness in the performance is due to the controller algorithm which only involves integral gain. It can be observed that the Combined method is able to reduce the interaction effects well compared to the other methods. The interaction produced by the Combined method is the lowest. Penttinen-Koivo method produces interaction slightly higher than the Combined method, while Fauzi method offers quiet high interaction.

Figures 24 and 25 show the closed loop responses of manipulate variable, which are product flow rate and cooling water flow rate during the separator level set point change respectively. For Davison method, large variations of product flow rate and cooling water flow rate are obtained, while Penttinen-Koivo and Combined method consist of small variations but high peak value. Among the four methods, Mohd Fauzi method exhibits the largest variations.

Figures 26 and 27 show the simulation responses for the operating pressure and separator level to the sequential step point changes in operating pressure set point. Based on Figure 26, Penttinen-Koivo and Combined methods provide a response which is mostly identical to the given set point. During step up input, Penttinen-Koivo method consists of

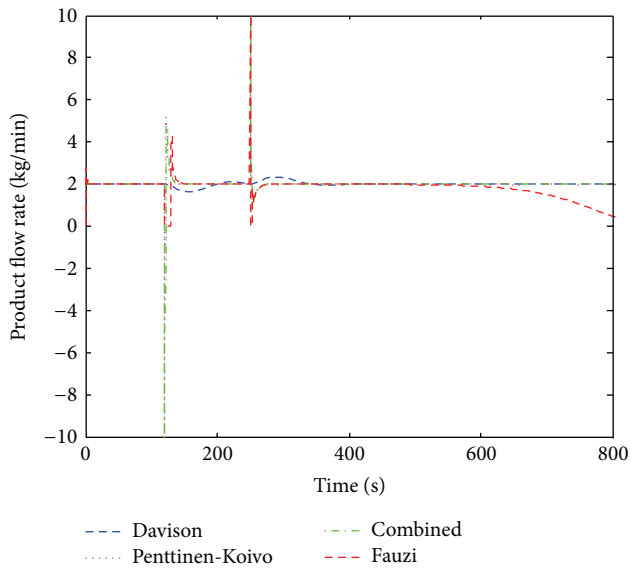


FIGURE 24: Product flow rate responses during separator level set point change.

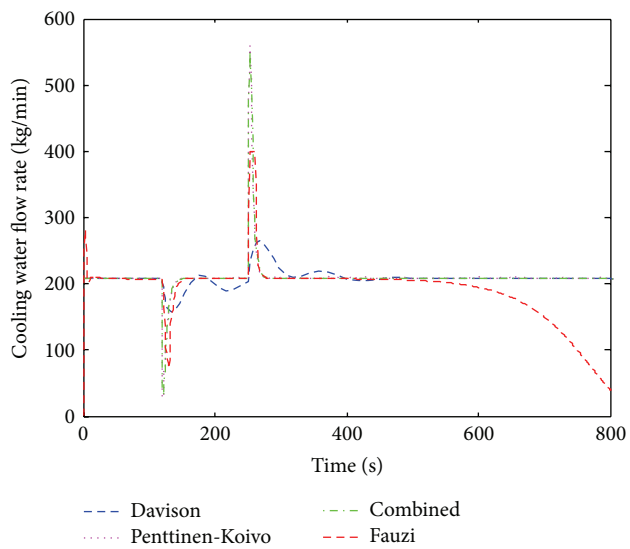


FIGURE 25: Cooling water flow rate responses during separator level set point change.

slightly high steady state errors than the Combined method. The difference is only about 1.40%. Meanwhile, the difference is approximately 0.1% during step down input. Even though the response by Davison method required long computation time for the rise and settling, the response is accomplished to settle at the set point value. But the response is relatively slow and consists of high percentage overshoot and steady state error. Figure 26 also shows that the output performance based on Davison method provides a response with 10% overshoot during the step up and step down input. These values are higher compared to other controller methods. However, it is still able to track the set point. The output performances based on Penttinen-Koivo and Combined methods are almost

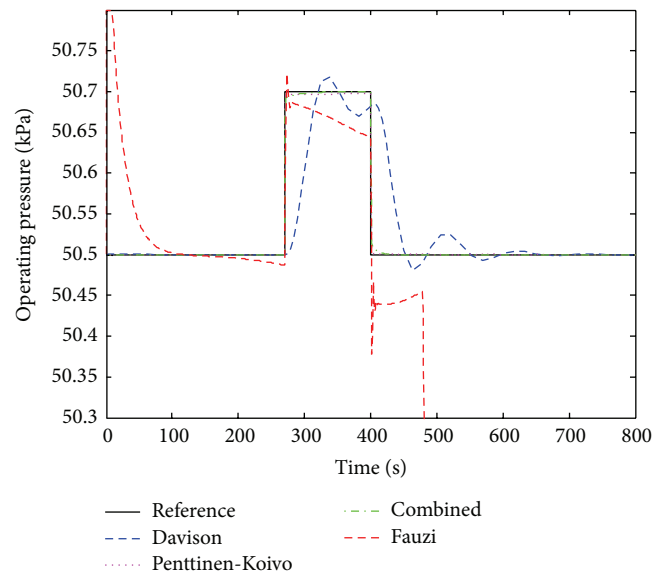


FIGURE 26: Responses of operating pressure during operating pressure set point change.

similar. Penttinen-Koivo provides faster rise and settling time during the step up and step down input. However, the Combined method provides the best performance with the lowest steady state error. Meanwhile, the resulting response by Fauzi obviously shows unstable characteristic. At the beginning, the response already shows that the system is in a state of uncontrollable. After the step down input was injected, the response gradually decreased. At time t = 800 s, the response of operating pressure is at −366.7 kPa. An increase in simulation time will lead the response to be infinity. Figure 27 shows the response of interactions during operating pressure set point change. Among the four methods, Penttinen-Koivo and Combined methods offer the least interaction. It can be seen that the interaction is reduced with the Penttinen-Koivo and Combined methods compared to the Davison and Fauzi methods which consist of high maximum amplitude.

Figures 28 and 29 show the closed loop responses of corresponding manipulated variable, which are product flow rate and cooling water flow rate during operating pressure set point change, respectively. The variation of both manipulate variables is similar during separator level set point change, where control based on Mohd Fauzi method exhibits a response with the largest variations.

The characteristic of closed loop response for evaporator system of all the comparative methods for singularly perturbed MPID control during the separator level and operating pressure set point change is tabulated in Table 5. The good performance of the Combined method is readily apparent. The best performance is given by Combined method, followed by Penttinen-Koivo, Davison, and Fauzi method.

Figure 30 shows the closed loop pole-zero plots for the proposed singularly perturbed MPID controller designs applied to the Newell and Lee evaporator. It can be seen that

Table 5: Characteristic of closed loop response for evaporator system.

Output	Method	Rise time, T_r (s) Step		Settling time, T_s (s) Step		Percentage overshoot (%OS) Step		Steady state error, e_{ss} (%) Step	
		Up	Down	Up	Down	Up	Down	Up	Down
Separator level, L_2	Davison	34.5	43.3	—	117	11.67	8.02	0.2	—
	Penttinen-Koivo	1.1	1.2	5.9	2.19	5.20	0.23	0.03	0.01
	Combined	0.9	1.6	5.4	7.5	3.67	—	0.01	0
	Fauzi	8.1	0.8	*	*	7.40	3.71	*	*
Operating pressure, P_2	Davison	30.3	34.4	—	139.6	10	10	—	0.5
	Penttinen-Koivo	0.7	1.3	1	1.8	—	—	1.45	0.15
	Combined	1	1.9	2	3.1	—	—	0.05	0.05
	Fauzi	2.5	2.4	*	*	15	*	*	*

*Unstable.

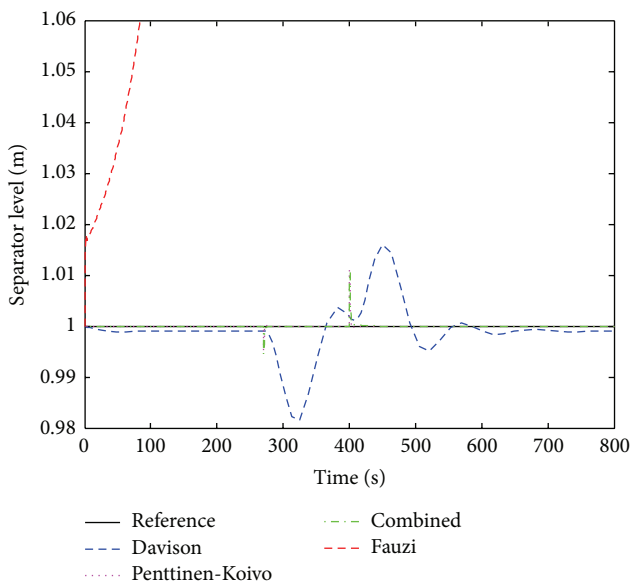


Figure 27: Process interactions during operating pressure set point change.

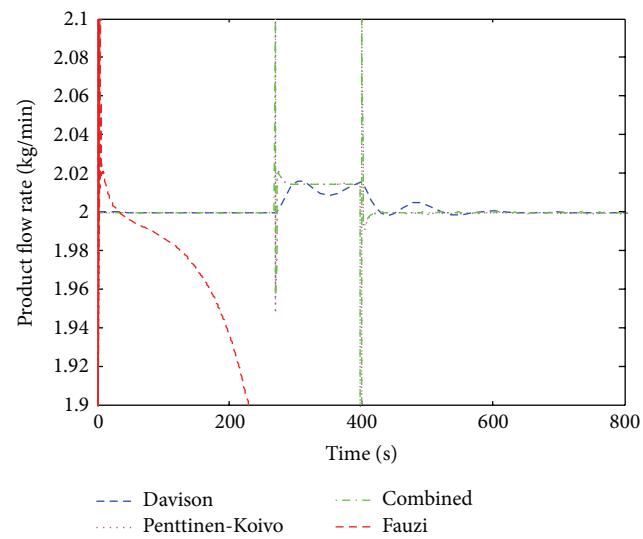


Figure 28: Product flow rate responses during operating pressure set point change.

all eigenvalues are located at the left-half plane of the s-plane. This indicates that the closed loop system is generally stable. A real pole in the left-half plane defines an exponentially decaying component in the homogenous response. The rate of the decay is determined by the eigenvalue location. Eigenvalues far from the origin in the left-half plane correspond to the components that decay rapidly, while eigenvalues near the origin correspond to slowly decaying components. Referring to Table 5, the rise and settling time of the response based on Penttinen-Koivo method are the most faster. It is proportional to the poles location indicated in Figure 30.

8. Conclusion

Designing MPID control tuning based on original and singularly perturbed system for multiinput multioutput (MIMO) processes is presented. Simulation results lead to the inference that, with the appropriate parameter tuning, a satisfactory singularly perturbed MPID control performance can be accomplished to control a nonlinear model of wastewater treatment plant. Ill-defined system like wastewater treatment plant which usually faces difficulties in control system, due to the natural behavior of two-time scale characteristic, can be efficaciously controlled by the implementation of singularly perturbed system into the MPID controller designs. Among the four methods, the Combined method yields somewhat better results with respect to decoupling capabilities, closed loop performances, and process inteaction.

For the second case study, Davison, Penttinen-Koivo, and Combined method were successfully applied to the nonlinear model of Newell and Lee evaporator. The well-tuned parameters of the controller designs were obtained using PSO approach. Simulation results show that the implementation of singularly perturbed system to the dynamic matrix inverse of Davison, Penttinen-Koivo, and Combined method has

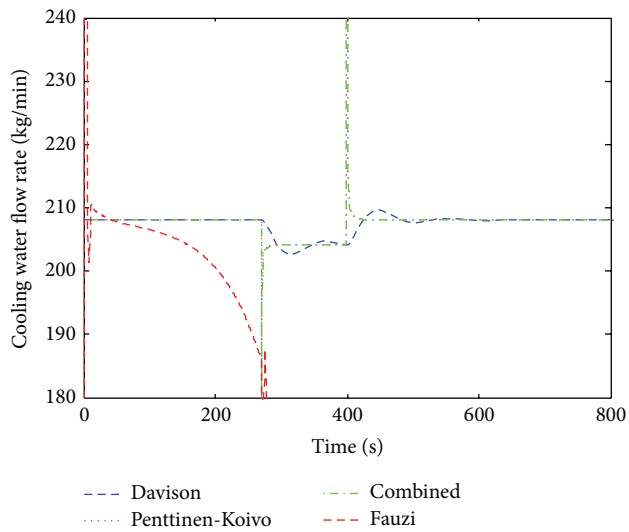


FIGURE 29: Cooling water flow rate responses during operating pressure set point change.

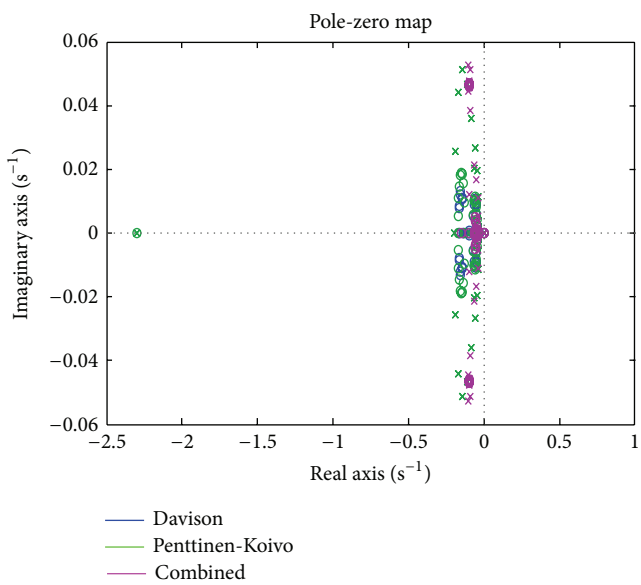


FIGURE 30: Closed loop pole-zero plot of Newell and Lee evaporator.

consistently provided a good performance. Among these three methods, Combined method provides the best control performance. Penttinen-Koivo method offers just a slightly poor control performance than the Combined method. Nevertheless, Maciejowski method is unable to be applied to Newell and Lee evaporator system as the open loop system is unstable which causes the information required by the controller design cannot be retrieved. It is observed that the proposed controller by [46] has weak performance for both separator level and operating pressure output control with high interaction.

Based on the system case studies, we can conclude that the control strategies proposed in these systems are capable of attaining the desired control performance and practically realizable where it is relevant to two-time scale system with a stable open loop system. The attained output responses consist of less percentage overshoot, fast settling time, and low steady state error, and the process interaction between the variables of the system is also reduced.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

The authors wish to thank the Research University Grant (GUP) vote 05H44 Universiti Teknologi Malaysia for the financial support.

References

[1] L. D. S. Coelho and M. W. Pessôa, "A tuning strategy for multivariable PI and PID controllers using differential evolution combined with chaotic Zaslavskii map," *Expert Systems with Applications*, vol. 38, no. 11, pp. 13694–13701, 2011.

[2] K. H. Ang, G. Chong, and Y. Li, "PID control system analysis, design, and technology," *IEEE Transactions on Control Systems Technology*, vol. 13, no. 4, pp. 559–576, 2005.

[3] W. K. Ho, T. H. Lee, and O. P. Gan, "Tuning of multiloop proportional-integral-derivative controllers based on gain and phase margin specifications," *Industrial and Engineering Chemistry Research*, vol. 36, no. 6, pp. 2231–2238, 1997.

[4] V. V. Kumar, V. S. R. Rao, and M. Chidambaram, "Centralized PI controllers for interacting multivariable processes by synthesis method," *ISA Transactions*, vol. 51, no. 3, pp. 400–409, 2012.

[5] K. L. N. Sarma and M. Chidambaram, "Centralized PI/PID controllers for nonsquare systems with RHP zeros," *Indian Institute of Science*, vol. 85, no. 4, pp. 201–214, 2005.

[6] A. Montazar, P. J. Van Overloop, and R. Brouwer, "Centralized controller for the Narmada main canal," *Irrigation and Drainage*, vol. 54, no. 1, pp. 79–89, 2005.

[7] F. Morilla, F. Vázquez, and J. Garrido, "Centralized PID control by decoupling for TITO processes," in *Proceedings of the 13th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA '08)*, pp. 1318–1325, Hamburg, Germany, September 2008.

[8] J. Garrido, F. Morilla, and F. Vázquez, "Centralized PID control by decoupling of a boiler-turbine unit," in *Proceedings of the European Control Conference*, pp. 4007–4012, Budapest, Hungary, August 2009.

[9] M. Willjuice Iruthayarajan and S. Baskar, "Covariance matrix adaptation evolution strategy based design of centralized PID controller," *Expert Systems with Applications*, vol. 37, no. 8, pp. 5775–5781, 2010.

[10] Y. Lee, S. Park, and M. Lee, "PID controller tuning to obtain desired closed loop responses for cascade control systems," *Industrial and Engineering Chemistry Research*, vol. 37, no. 5, pp. 1859–1865, 1998.

[11] H. K. Khalil, "Universal integral controllers for minimum-phase nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 490–494, 2000.

[12] N. A. Mahmoud and H. K. Khalil, "Asymptotic regulation of minimum phase nonlinear systems using output feedback," *IEEE Transactions on Automatic Control*, vol. 41, no. 10, pp. 1402–1412, 1996.

[13] V. D. Yurkevich, "PI and PID controller design for nonlinear systems in the presence of a time delay via singular perturbation technique," in *Proceedings of the 9th International Conference on Actual Problems of Electronic Instrument Engineering (APEIE '08)*, pp. 168–174, IEEE, Novosibirsk, Russia, September 2008.

[14] P. Martin and R. Katebi, "Multivariable PID tuning of dynamic ship positioning control systems," *Journal of Marine Engineering and Technology*, vol. 4, no. 2, pp. 11–24, 2005.

[15] N. A. Wahab, M. R. Katebi, and J. Balderud, "Multivariable PID control design for wastewater systems," in *Proceedings of the Mediterranean Conference on Control and Automation*, pp. 1–6, Athens, Greece, July 2007.

[16] P. Balaguer, N. A. Wahab, M. R. Katebi, and R. Vilanova, "Multivariable PID control tuning: a controller validation approach," in *Proceedings of the 13th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA '08)*, pp. 289–294, Hamburg, Germany, September 2008.

[17] T.-H. S. Li and K.-J. Lin, "Composite fuzzy control of nonlinear singularly perturbed systems," *IEEE Transactions on Fuzzy Systems*, vol. 15, no. 2, pp. 176–187, 2007.

[18] P. V. Kokotovic, R. E. O'Mallley Jr., and P. Sannuti, "Singular perturbations and order reduction in control theory—an overview," *Automatica*, vol. 12, no. 2, pp. 123–132, 1976.

[19] A. K. Packard and S. M. Shahruz, "Estimates of the singular perturbation parameter for stability, controllability, and observability of linear systems," in *Proceedings of the 31st IEEE Conference on Decision and Control*, vol. 4, pp. 3062–3063, Tucson, Ariz, USA, December 1992.

[20] S. Paper, "Singular perturbations and time-scale methods in control theory: survey 1976–1983," *Automatica*, vol. 20, no. 3, pp. 273–293, 1984.

[21] H. Yu, X. Zhang, G. Lu, and Y. Zheng, "On the model-based networked control for singularly perturbed systems with nonlinear uncertainties," in *Proceedings of the IEEE Conference on Decision and Control*, pp. 684–689, Shanghai, China, December 2009.

[22] R. G. Phillips, "A two-stage design of linear feedback controls," *IEEE Transactions on Automatic Control*, vol. 25, no. 6, pp. 1220–1223, 1980.

[23] K.-I. Kang, K.-S. Park, and J.-T. Lim, "Exponential stability of singularly perturbed systems with time delay and uncertainties," *International Journal of Systems Science*, vol. 46, no. 1, pp. 170–178, 2015.

[24] H. K. K. Rabah and W. Aldhaheri, "A real Schur form method for modeling singularly perturbed systems," in *Proceedings of the American Control Conference*, pp. 1719–1721, Atlanta, Ga, USA, June 1988.

[25] D. S. Naidu, *Singular Perturbation Methodology in Control Systems*, vol. 34, Peter Peregrinus, London, UK, 1988.

[26] L. Li and F. Sun, "Stable fuzzy adaptive controller design for nonlinear singularly perturbed systems," in *Proceedings of the IMACS Multiconference on Computational Engineering in Systems Applications*, pp. 1388–1394, IEEE, Beijing, China, October 2006.

[27] A. Saberi and H. Khalil, "Stabilization and regulation of nonlinear singularly perturbed systems—composite control," *IEEE Transactions on Automatic Control*, vol. 30, no. 8, pp. 739–747, 1985.

[28] J.-S. Chiou, "Design of controllers and observer-based controllers for time-delay singularly perturbed systems via composite control," *Journal of Applied Mathematics*, vol. 2013, Article ID 813598, 9 pages, 2013.

[29] H. Bouzaouache and N. B. Braiek, "On guaranteed global exponential stability of polynomial singularly perturbed control systems," in *Proceedings of the IMACS Multiconference on Computational Engineering in System Applications*, vol. 1, pp. 299–305, Beijing, China, October 2006.

[30] Y. J. Kim, B. S. Kim, and M. T. Lim, "Finite-time composite control for a class of singularly perturbed nonlinear systems via successive Galerkin approximation," *IEEE Proceedings—Control Theory and Applications*, vol. 152, no. 5, pp. 507–512, 2005.

[31] Z. Retchkiman and G. Silva, "Stability analysis of singularly perturbed systems via vector Lyapunov methods," in *Proceedings of the 35th IEEE Conference on Decision and Control*, vol. 1, pp. 580–585, Kobe, Japan, December 1996.

[32] S. I. Samsudin, M. F. Rahmat, N. A. Wahab, Zulfatman, S. N. S. Mirin, and M. C. Razali, "Two-time scales matrix decomposition for wastewater treatment plant," in *Proceedings of the IEEE 8th International Colloquium on Signal Processing and Its Applications (CSPA '12)*, pp. 347–351, Melaka, Malaysia, March 2012.

[33] M.-N. Contou-Carrere and P. Daoutidis, "Dynamic precompensation and output feedback control of integrated process networks," in *Proceedings of the 2004 American Control Conference (AAC '04)*, pp. 2909–2914, IEEE, Boston, Mass, USA, June-July 2004.

[34] N. Vora and P. Daoutidis, "Nonlinear model reduction of chemical reaction systems," in *Proceedings of the American Control Conference*, vol. 3, pp. 1583–1587, San Diego, Calif, USA, June 1999.

[35] H. K. Khalil, "Output feedback control of linear two-time-scale systems," *IEEE Transactions on Automatic Control*, vol. 32, no. 9, pp. 784–792, 1987.

[36] K. W. Chang, "Diagonalization method for a vector boundary problem of singular perturbation type," *Journal of Mathematical Analysis and Applications*, vol. 48, no. 3, pp. 652–665, 1974.

[37] J. Niu, J. Zhao, Z. Xu, and J. Qian, "A two-time scale decentralized model predictive controller based on input and output model," *Journal of Automated Methods and Management in Chemistry*, vol. 2009, Article ID 164568, 11 pages, 2009.

[38] K. Zhang and X. An, "Design of multivariable self-tuning PID controllers via quasi-diagonal recurrent wavelet neural network," in *Proceedings of the 2nd International Conference on Intelligent Human-Machine Systems and Cybernetics*, vol. 2, pp. 95–99, Nanjing, China, August 2010.

[39] M. C. Razali, N. A. Wahab, and S. I. Samsudin, "Multivariable PID using singularly perturbed system," *Jurnal Teknologi*, vol. 67, no. 5, pp. 63–69, 2014.

[40] I. C. Trelea, "The particle swarm optimization algorithm: convergence analysis and parameter selection," *Information Processing Letters*, vol. 85, no. 6, pp. 317–325, 2003.

[41] X. Li, F. Yu, and Y. Wang, "PSO algorithm based online self-tuning of PID controller," in *Proceedings of the International Conference on Computational Intelligence and Security*, pp. 128–132, Harbin, China, December 2007.

[42] M. I. Solihin, M. A. S. Kamal, and A. Legowo, "Optimal PID controller tuning of automatic gantry crane using PSO algorithm," in *Proceedings of the 5th International Symposium on Mechatronics and its Applications (ISMA '08)*, pp. 1–5, Amman, Jordan, May 2008.

[43] W. U. Dongsheng, Y. Qing, and W. Dazhi, "A novel PSO-PID controller application to bar rolling process," in *Proceedings of the 30th Chinese Control Conference (CCC '11)*, pp. 2036–2039, Yantai, China, July 2011.

[44] F. Nejjari, B. Dahhou, A. Benhammou, and G. Roux, "Nonlinear multivariable adaptive control of an activated sludge wastewater treatment process," *International Journal of Adaptive Control and Signal Processing*, vol. 13, no. 5, pp. 347–365, 1999.

[45] D.-S. Joo and H. Park, "Control of the dissolved oxygen concentration in the activated sludge process," *Environmental Engineering Research*, vol. 3, no. 2, pp. 115–121, 1998.

[46] M. F. B. N. Shah, *Multi-objective optimization of MIMO control system using surrogate modeling [M.S. thesis]*, Universiti Teknologi Malaysia, 2012.

[47] M. Tajjudin, R. Adnan, N. Ishak, M. H. F. Rahiman, and H. Ismail, "Model reference input for an optimal PID tuning using PSO," in *Proceedings of the IEEE International Conference on Control System, Computing and Engineering (ICCSCE '11)*, pp. 162–167, IEEE, Penang, Malaysia, November 2011.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 684370, 9 pages
http://dx.doi.org/10.1155/2015/684370

Research Article

Feature Selection Tracking Algorithm Based on Sparse Representation

Hui-dong Lou,[1] Wei-guang Li,[1] Yue-en Hou,[2] Qing-he Yao,[3] Guo-qiang Ye,[1] and Hao Wan[1]

[1] *School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou, Guangdong, China*
[2] *School of Computer Science, Jiaying University, Meizhou, Guangdong, China*
[3] *School of Engineering, Sun Yat-sen University, Guangzhou, Guangdong, China*

Correspondence should be addressed to Hui-dong Lou; loudong@mail.gdufs.edu.cn

Received 11 September 2015; Accepted 28 September 2015

Academic Editor: Xinguang Zhang

In order to enhance the robustness of visual tracking algorithm in complex environment, a novel visual tracking algorithm based on multifeature selection and sparse representation is proposed. In the framework of particles filter, particles with low target similarity are first filtered out by a fast algorithm; then, based on the principle of sparsely reconstructing the sample label, the features with high differentiation against the background are involved in the computation so as to reduce the disturbance of occlusions and noises. Finally, candidate targets are linearly reconstructed via sparse representation and the sparse equation is solved by using APG method to obtain the state of the target. Four comparative experiments demonstrate that the proposed algorithm in this paper has effectively improved the robustness of the target tracking algorithm.

1. Introduction

Visual target tracking algorithm, due to its wide application in fields such as robotics visual control, human-machine interaction, intelligent assistance driving, and video surveillance, has attracted increasing attention from the researchers. However, in complex environment, due to changes in illumination and expression of the target object, object occlusion, and noises disturbance, design of an accurate and real-time visual tracker remains a challenging problem [1].

In recent years, a kind of appearance modeling technology called sparse representation has been widely used for information compression and pattern recognition. Agarwal and Roth [2] have achieved good results in object recognition via sparse representation. Researches [3–5] demonstrate that high recognition rate has been obtained in face recognition via sparse representation. Mei and Ling [6, 7] first introduced sparse representation theory into visual tracking and developed a sparse representation tracking algorithm based on particles filter. The algorithm uses the template dictionary to linearly reconstruct the candidate targets and imposes sparsity constraints on the reconstruction coefficients. Apart from the target template, the algorithm uses the trivial template to construct the template dictionary, showing good robustness to occlusion; however, it requires a large amount of calculation because the algorithm adopts the LESSO method to solve the sparse equation. Bai and Li [8] developed a structural sparse representation model, which divides the sample into blocks and uses the Block Orthogonal Matching Pursuit (BOMP) algorithm to solve the sparse equation. It effectively improves the calculation speed, but the algorithm is not robust to illumination change. Hou et al. [9] provided a tracking algorithm of block sparse representation, in which each block is given the corresponding weights to improve the robustness to occlusion, but there may be drift due to large noises. A target tracking algorithm based on the structural sparse representation is proposed in [10], which constructs a vector pool of sparse representation coefficients by blocking the target sample and identifies the target state through the similarity information in the vector pool. But this algorithm fails to effectively use the residual error information; its robustness needs to be further improved. Based on this,

Hou et al. [11] developed a target tracking algorithm with sparse representation based on ranking. The algorithm gives full consideration to the sparse representation coefficients and residual error information while locating the target, which improves the robustness of the tracking algorithm. In the framework of structural sparse representation, the sparse coefficients of the candidate targets is classified in [12] by training the Naive Bayes Classifier, strengthening the algorithm's ability to differentiate between the target and its background. But the algorithm fails to extract the features with differentiation; its robustness needs to be further improved. Wang et al. [13] propose a least soft-threshold squares tracking algorithm based on sparse representation, which models the reconstructed residual error terms with the Gaussian-Laplacian distribution and finds the optimal solution of the objective equation by using iterative and soft-threshold methods. Bao et al. [14] impose constraints on the target template coefficients and the trivial template with L_1-norm and L_2-norm, respectively. Additionally, it adopts the Accelerated Proximal Gradient (APG) method to solve the objective equation with sparsity constraints. As a result, the accuracy of the tracker and the speed of calculation are significantly improved. Zhuang et al. [15] propose a multitask concept that is similar to [14]; the algorithm also uses the Accelerated Proximal Gradient (APG) method to solve the objective equation by iteration, but it requires a fair amount of calculation. In the framework of sparse representation, Lan et al. [16] propose an objective equation based on adaptive multifeature selection and solve the equation by using the Accelerated Proximal Gradient (APG) method. The multifeature selection method has effectively improved the robustness of the tracking algorithm, but it requires a large amount of calculation.

In order to improve the robustness of visual tracking algorithm, this paper proposes a novel tracking algorithm based on multifeature fusion and sparse representation. According to the image intensity, the proposed algorithm uses the simple algorithm to filter out the candidate targets largely dissimilar to the target. Then, the discriminative features, which come from multifeature, are selected by using a method of sparsely reconstructing sample label. Finally, it uses L_2-norm to linearly reconstruct the candidate targets and obtains the state of the target.

2. Overview of the Tracking Algorithm

The theorem of proposed tracking algorithm based on multifeature fusion and sparse representation in this paper is shown in Figure 1. Compared with other tracking algorithms, it has the following contributions:

(1) This paper proposes a fast algorithm that can quickly filter out the particles with low similarity to the target template, improving the calculation speed of the algorithm.

(2) Based on the sample feature constructed from multifeature, the proposed algorithm uses the method of sparsely reconstructing positive and negative sample label to extract sample features, which can help these extracted features more discriminatively.

(3) The conventional tracking algorithms based on sparse representation always use L_1-norm to impose the constraints on the linear reconstruction function, but they require a large amount of calculation, whereas the proposed algorithm in this paper uses APG method to solve the sparse function with nonnegative constraints, greatly improving the calculation speed of the algorithm.

3. Fast Particle Filter Algorithm

In the framework of particle filtering, most particles bear little resemblance to the target. So they can be filtered out through simple algorithm. Based on this idea, a fast algorithm is proposed to filter out the particles with low similarity to the target.

Let $Z = [Z_1, Z_2, \ldots, Z_n]$ denote the sample set of the candidate targets, and $\overline{Z} \in R^m$ denote the target template, where $Z_i \in R^m$ is one of the candidate target samples, m is the dimension of the sample, and n represents the number of particles. In order to reduce the amount of calculation, only the gray image information of the sample is involved in the calculation. By normalizing the candidate target Z_i and target template $\overline{Z}$, we can get $z_i' = z_i/(\sum_{j=1}^m z_{i,j})$ and $\overline{z}' = \overline{z}/(\sum_{j=1}^m \overline{z}_j)$, where $z_{i,j}$ is the jth feature of z_i, $\overline{z}_j$ is the jth feature of $\overline{z}$, and u_i is the similarity measure of z_i and $\overline{z}$. If z_i bears much resemblance to $\overline{z}$, then there will be no remarkable difference between the values of u_i. So we can filter out the particles according to the fluctuation of values of u_i:

$$\overline{u_i} = u_{i,j} - \text{mean}\left(u_i\right), \tag{1}$$

where $u_{i,j}$ is the jth element of u_i and mean (u_i) is the mean of all elements of u_i. The value of $\overline{u_i}$ reflects the fluctuation of values of u_i. If the target is disturbed by occlusions or noises, the real target can possibly be filtered out when the value of $\overline{u_i}$ is too small. In order to solve this problem, only the elements with 50% variance value in $\overline{u_i}$ will be involved in the calculation. Sort the elements in $\overline{u_i}$ in descent order according to their values and eliminate the elements with values in the top half; then we can get the vector $\widetilde{u_i}$. Suppose $y_i = \sum_{j=1}^{m/2} \widetilde{u_{i,j}}$, where $\widetilde{u_{i,j}}$ is the jth element of $\widetilde{u_i}$; then we can filter out the particles according to the value of y_i. Through experiment, we select the y_i with smaller value ($n_0 = n/3$) from all the particles to participate in the later operation. Through the algorithm, we can quickly eliminate two-thirds of particles with low similarity to the object, which effectively improves the calculation speed of the tracking algorithm.

4. Feature Selection

In order to improve the robustness of the tracking algorithm, we use the image intensity and LBP feature $s_L \in R^{m_2}$ to construct the sample features. Denote the gray features of the sample by $s_G \in R^{m_1}$ and the LBP feature of the sample

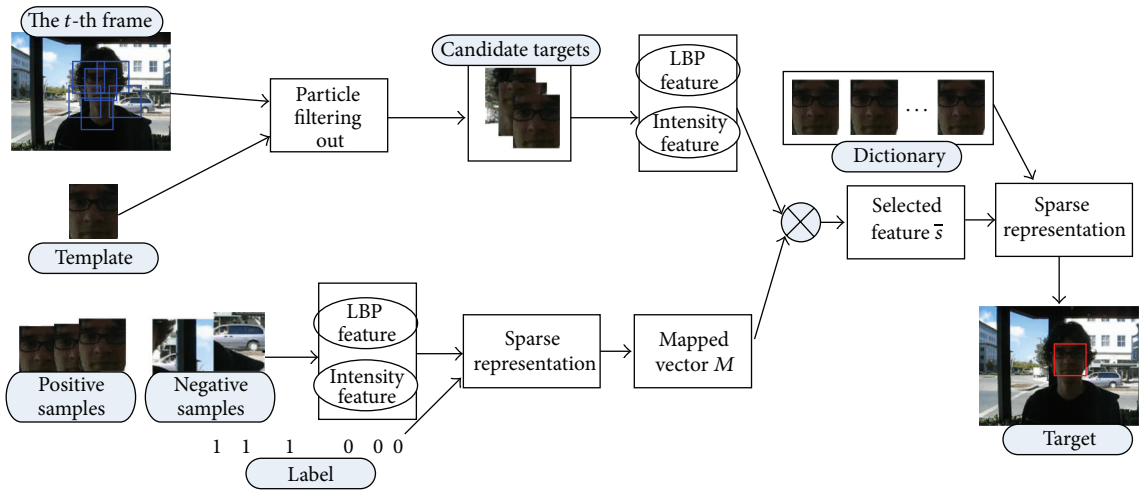


FIGURE 1: Proposed tracking algorithm.

by $s_L \in R^{m_2}$. When s_G and s_L are combined, we can get $s = [s_G, s_L]^T$, where $s \in R^k$ represents the sample feature, $k = m_1 + m_2$. Because the dimension of s is high, it contains a fair amount of redundant information. In order that the algorithm can differentiate between target and background better, we need to select the discriminative features from s.

During the tracking process, according to the target sample in the first frame and the tracking result, we can get the positive and negative sample set of the target $T = [t_p, t_n]$, where the positive sample of the target is $T_p \in R^{k\times p}$ and the negative sample is $T_n \in R^{k\times q}$, where p represents the number of the positive samples and q indicates the number of negative samples. Use the sparse reconstruction theory to select the sample features; then we can get

$$\min_A \left(\frac{1}{2} \left\| T^T A - l \right\|_2^2 + \lambda \|A\|_2 \right), \tag{2}$$

where $\|\cdot\|$ is the norm operator and l represents the label vectors of the positive and negative samples. The positive and negative samples are set to 1 and -1, respectively. λ is the sparse adjustment coefficient and $A \in R^k$ represents the reconstruction vectors, $A = [A_1, A_2, \ldots, A_k]$. It should be noted that, in (2), we use the 2-norm to impose constraints on the sparsity of A. The benefit is that it can effectively reduce the amount of calculation in solving the sparse function. Set threshold as τ, and when $A_i > \tau$, the corresponding feature has a strong ability to differentiate between the target and its background.

Construct the mapped vector of feature selection $M = [M_1, M_2, \ldots, M_k]^T$, $M \in R^k$. Consider

$$\begin{aligned} M_i &= 0, \quad Ai \le \tau, \\ M_i &= 1, \quad Ai > \tau. \end{aligned} \tag{3}$$

Here, s is the feature vector before the feature selection. After the feature selection, the feature vector $\bar{s}$ is defined as

$$\bar{s} = s \otimes M, \tag{4}$$

where $\otimes$ denotes the multiplication of the corresponding elements within the vector.

5. Solution to the Sparse Equation with APG Approach

After feature selection, we get the feature vector set of the candidate targets $S = [\overline{s_1}, \overline{s_2}, \ldots, \overline{s_{n_0}}]$, $\bar{s} \in R^{k_0}$, and k_0 denotes the dimension of each sample after feature selection. Let $D = [D_0, I, -I]$ denote the template dictionary, where D_0 is the target sample set, $D_0 \in R^{k_0\times f}$, and $I \in R^{k_0\times k_0}$ represents the unit diagonal matrix used to reduce the disturbance of occlusions and noises.

Using template dictionary D to make sparse linear reconstruction for the candidate targets, we can get

$$\begin{aligned} \arg\min_C \quad & \frac{1}{2} \|\bar{s} - DC\|_2^2 + \lambda \|C\|_1 \\ \text{s.t.} \quad & C_T \ge 0, \end{aligned} \tag{5}$$

where $C = [C_T, C_I]^T$ represents the sparse coefficients, C_T is the sparse coefficients corresponding to the target template, and C_I is the sparse coefficients corresponding to I and $-I$. Impose the nonnegative constraints on the elements in C_T to improve the robustness of the tracker [7].

After adding a penalty term, (5) with nonnegative constraints can be updated by

$$\arg\min_C \frac{1}{2} \|\bar{s} - DC\|_2^2 + \lambda \|C\|_1 + J(C_T), \tag{6}$$

where $J(\cdot)$ is the penalty term defined by

$$J(a) = 1, \quad a \geq 0; \qquad J(a) = +\infty, \quad a < 0. \tag{7}$$

The solution to (6) is equivalent to optimizing the convex function. This paper uses the Accelerated Proximal Gradient (APG) approach to find the optimal solution to (6). Set

$$F(C) = \frac{1}{2}\|\bar{s} - DC\|_2^2 + \lambda\|C\|_1, \qquad G(C) = J(C_T), \tag{8}$$

where $F(C)$ is a differentiable convex function and $G(C)$ is a discontinuous convex function. The specific algorithm is shown in Algorithm 1.

Algorithm 1. Minimization algorithm for (6) by using APG method is as follows:

(1) Initialize, $\alpha_0 = \alpha_{-1} = 0, t_0 = t_{-1} = 1$

(2) For $i = 1:4$

(3) $\beta_{i+1} = \alpha_i + ((t_{i-1} - 1)/t_i)(\alpha_i - \alpha_{i-1})$;

(4) $\alpha_{i+1} = \arg\min_C(\gamma/2)\|C - \beta_{i+1} + \nabla F(\beta_{i+1})/\gamma\|_2^2 + G(C)$;

(5) $t_{i+1} = (1 + \sqrt{1 + 4t_i^2})/2$;

(6) end for

In Algorithm 1, step (4) needs to find the minimum value of the function, so it can be updated by

$$\arg\min_C \frac{\gamma}{2}\|C - h_{i+1}\|_2^2 + J(C_T), \tag{9}$$

where $h_{i+1} = \beta_{i+1} - \nabla F(\beta_{i+1})/\gamma$. Equation (9) can be seen as to solve the minimum values of the two functions as follow:

$$\arg\min_{C_T} \frac{\gamma}{2}\left\|C_T - h_{i+1}^T\right\|_2^2 + J(C_T), \tag{10}$$

$$\arg\min_{C_I} \frac{\gamma}{2}\left\|C_I - h_{i+1}^I\right\|_2^2, \tag{11}$$

where h_{i+1}^T and h_{i+1}^I are the elements corresponding to T and $[I, -I]$ in h_{i+1}. It can be seen that the optimal solution to (10) is $C_T = \max(0, h_{i+1}^T)$ and the optimal solution to (11) is $C_I = h_{i+1}^I$.

6. Object Tracking

The proposed algorithm in this paper is implemented in the framework of particles filter. In the first frame of the video, the initial state of the target is picked by mouse or captured through target recognition. Let $\{o_1, o_2, \ldots, o_t\}$ denote the observation values from the first frame to the tth frame of the video. x_t^i is the state of the ith particle in the tth frame. The target state in tth frame is

$$x_t = \arg\max_{x_t^i}\left(p\left(x_t^i \mid o_{1:t}\right)\right). \tag{12}$$

Here, $p(x_t^i \mid o_{1:t})$ can be obtained by solving the following equations:

$$p(x_t \mid o_{1:t-1}) = \int p(x_t \mid x_{t-1})\, p(x_{t-1} \mid o_{1:t-1})\, dx_{t-1}, \tag{13}$$

$$p(x_t \mid o_{1:t}) = \frac{p(o_t \mid x_t)\, p(x_t \mid o_{1:t-1})}{p(o_t \mid o_{1:t-1})}. \tag{14}$$

In (13), $p(x_t \mid x_{t-1})$ is the state transfer function. The state of the sample is defined by six-dimensional affine vector $[\tau_1, \tau_2, \tau_3, \tau_4, \tau_5, \tau_6]$, which represents the x coordinate, y coordinate, length-width ratio, rotation angle, torsion angle, and scale, respectively. Suppose that the six parameters are mutually independent; then the state transfer function $p(x_t \mid x_{t-1})$ can be represented as

$$p(x_t \mid x_{t-1}) = N(x_t \mid x_{t-1}, \Sigma), \tag{15}$$

where Σ is the diagonal matrix constituted by $[\tau_1, \tau_2, \tau_3, \tau_4, \tau_5, \tau_6]$.

The relation between the similarity measure $p(o_t \mid x_t)$ in (14) and the residual error e based on sparse representation can be represented as

$$p(o_t \mid x_t) \propto \frac{1}{e}. \tag{16}$$

The proposed algorithm is shown in Algorithm 2.

Algorithm 2. The proposed algorithm in this paper is as follows.

Input. Video, the target state in the first frame

(1) Obtain the target dictionary of the initial target and positive and negative samples

(2) for $i = 1: f$ (f is the number of the frames of the video)

(3) Update particles with normal distribution

(4) Extract the intensity of the particles s_G and filter out particles

(5) Extract LBP feature of the sample s_L and get s

(6) Use sparse reconstruction theory to select the sample features, and then get the feature vector $\bar{s}$

(7) Construct a sparse equation and solve it with APG method

(8) if f can be divided by 5

(9) Update the feature vector $\bar{s}$, update template dictionary, and update the positive and negative samples

(10) end if

(11) end for

Ouput. The state of target in each frame.

Figure 2: Tracking results on sequence "*Panda*."

7. Experiments and Discussion

The proposed algorithm in this paper is implemented in Matlab 2009b. In order to make the tracking results of the proposed algorithm more convincing, the experiment selects other three representative tracking algorithms for comparison, including L_1 tracking algorithm [7], IVT tracking algorithm [17], and ASLSAM tracking algorithm. The four algorithms will be tested by four internationally used tracking test videos including "Panda," "Woman Square," "Trellis," and "ThreePassShop2cor". It is a challenging task to track these targets because all the videos pose challenging factors such as partial occlusion and variations in illumination, pose, and scale.

The first test video is Panda with a cartoon panda as the tracked target. For this video, the difficulty in tracking the target lies in the partial occlusion and the target rotation. Figure 2 shows the test results of the four algorithms in the 4th, 55th, 108th, 176th, 201th, and 241th frame. The test results of the proposed algorithm, ASLSAM algorithm, L_1 algorithm, and IVT algorithm are marked by the blue, red, gray, and purple boxes, respectively. As shown in Figure 2, our proposed algorithm can handle the target very well before the 201th frame but drift from the target in the 201th and 241th frame due to the occlusion. However, it has never missed the target throughout the tracking process. While the other three algorithms miss the target when there are occlusion and target rotation.

The second test video is Trellis with the face as the target. There is a great difficulty in tracking the target because of the drastic illumination and variation in pose in the video. As shown in Figure 3, our proposed algorithm can track the target faithfully throughout the tracking process without being affected by the pose and illumination change. But L_1 algorithm and IVT algorithm miss the target in the 276th frame and ASLSAM algorithm misses the target in the 532th frame.

The third test video is Woman Square with a pedestrian as the target. The main challenge of tracking the target comes from the partial occlusion. As shown in Figure 4, only our proposed algorithm has successfully tracked the target without being affected by the partial occlusion while all the other three algorithms miss the target in the 143th frame as a result of the partial occlusion.

The fourth test video is ThreePassShop2Cor. The difficulty of tracking the target lies in the occlusion, the disturbance of the similar object, and scale variation. The tracking results of the four algorithms are shown in Figure 5. Because of the disturbance of occlusion and similar objects, our proposed tracker drifts from the target in the 125th frame but relocates the target later. While the ASLSAM and IVT algorithms locate the similar object instead of the tracked target in the 125th frame and the L_1 algorithm misses the target in the 303th frame.

In order to compare the tracking results better, Table 1 lists the maximum, mean, and standard variance values of the tracking error of the four algorithms in four image sequences. The tracking error refers to the Euclidean distance between the center point of the target derived from the tracking algorithm and the center point of the actual target. It can be seen that the tracking results on all the four videos demonstrate that our proposed algorithm achieves more favorable performance than the other three.

Apart from handling the tracking error, in order to reflect the relation between the tracking results and the actual

Table 1: Maximum, mean, and standard variance values of the tracking error.

Video	Proposed algorithm			ASL algorithm			L_1 algorithm			IVT algorithm		
	Max	Mean	STD	Max	Mean	STD	Max	Mean	STD	Max	Mean	STD
TPS	24.2	3.9	4.6	115.3	65.6	38.0	40.9	18.2	14.1	114.9	65.8	38.4
WS	12.6	3.4	2.2	259.8	15.2	90.3	225.3	129.4	75.2	343.8	156.8	98.8
Panda	23.4	8.6	5.1	385.8	136.3	158.9	410.1	129.9	112.2	190.4	68.3	68.8
Trellis	13.6	3.6	2.9	191.2	21.6	45.4	222.1	59.2	54.4	151.6	72.3	51.0
Average	18.5	4.9	3.7	238.0	59.7	83.2	224.6	84.2	64.0	200.2	90.8	64.25

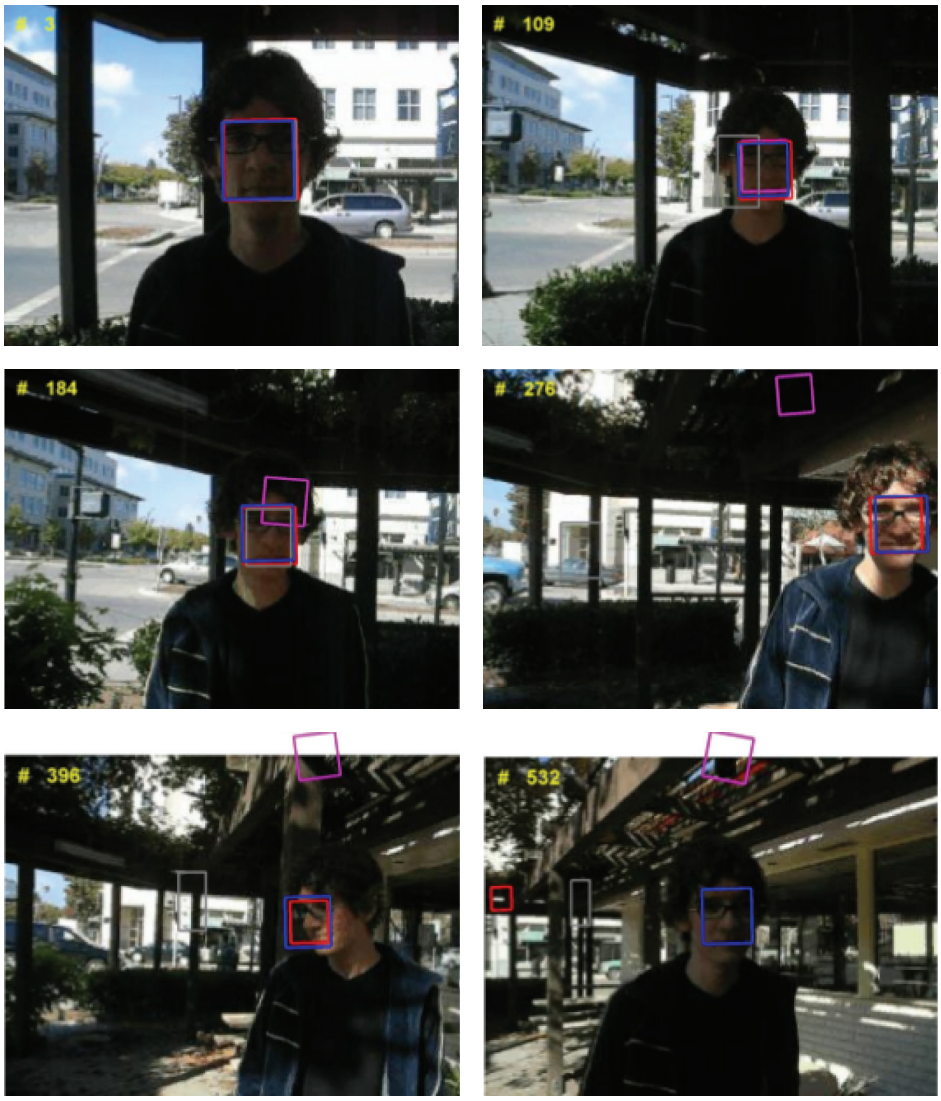


Figure 3: Tracking results on sequence "*Trellis*."

appearance of the target, Table 2 lists the success rate of the four tracking algorithms based on the PASCAL VOC standard [18]. It can be seen that the success rate of our proposed algorithm is remarkably higher than that of the other three.

8. Conclusion

A new sparse representation-based tracking algorithm in the framework of particles filter is proposed in this paper. The new method first uses a fast algorithm to filter out

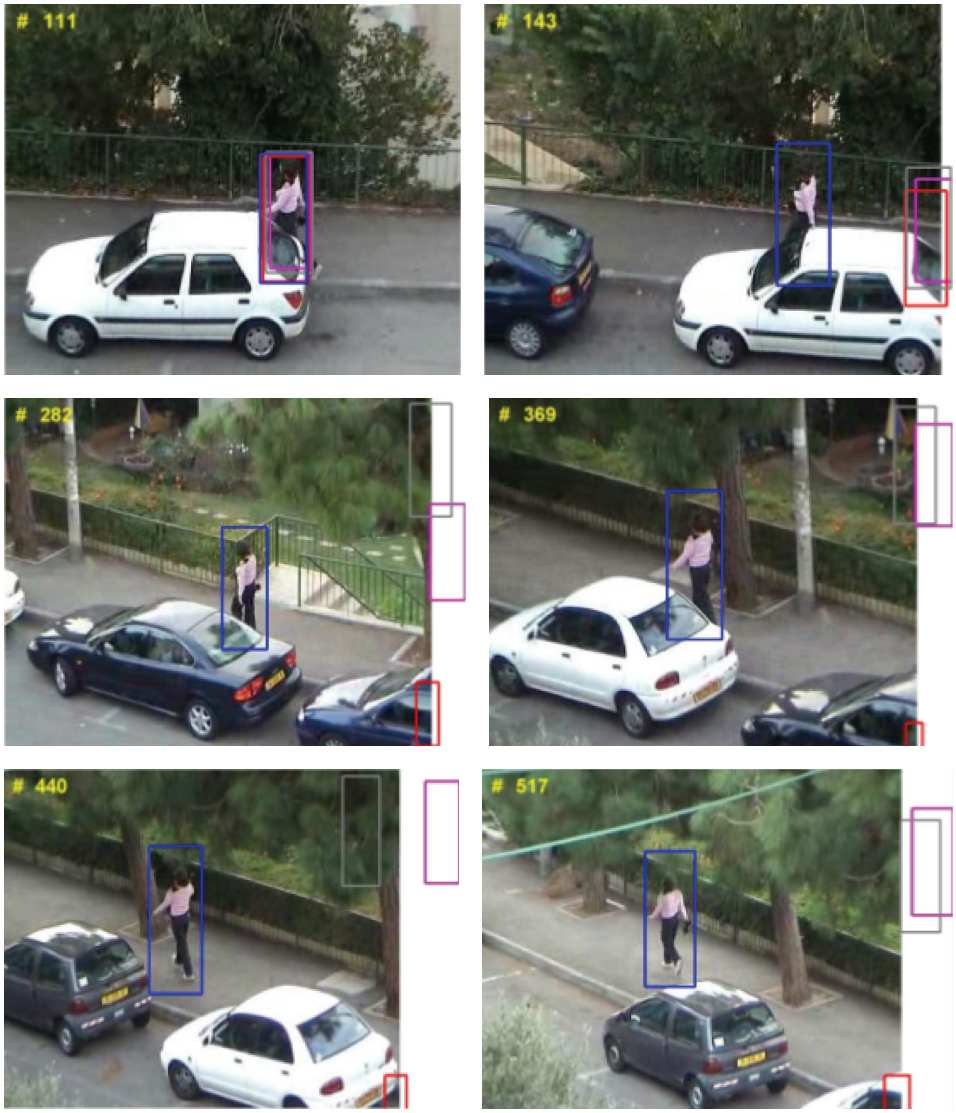


Figure 4: Tracking results on sequence "*Woman Square*."

Table 2: Success rate of tracking algorithms.

Video	Proposed algorithm	ASLSAM algorithm	L_1 algorithm	IVT algorithm
TPS	0.99	0.21	0.49	0.24
WS	0.76	0.21	0.21	0.20
Panda	0.85	0.52	0.15	0.42
Trellis	0.94	0.74	0.24	0.26
Average	0.86	0.42	0.27	0.28

the particles largely dissimilar to the tracked target, which reduces the amount of calculation. Then, with the intensity feature and LBP feature combined, the algorithm extracts the features with discriminative ability via L_2 sparse representation, improving its robustness to the occlusion and disturbance. Furthermore, it adopts the APG method to solve the sparse equation with nonnegative coefficient constraints, improving the computational speed and robustness of the proposed algorithm. Finally, the experiments demonstrate that the proposed tracker achieves more favorable tracking results.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publishing of this paper.

Acknowledgment

This work is supported by the Foundation of Innovation Demonstration Project for NC mechanical product in Guangdong, China (no. 2013B011301026).

Figure 5: Tracking results on sequence *"ThreePassShop2Cor."*

References

[1] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: a survey," *ACM Computing Surveys*, vol. 38, no. 4, pp. 1158–1166, 2006.

[2] S. Agarwal and D. Roth, "Learning a sparse representation for object detection," *Computer Science*, vol. 2353, pp. 97–101, 2006.

[3] J. Wright, Y. Yang Allen, A. Ganesh et al., "Robust face recognition via sparse representation," in *Proceedings of the 8th IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 210–227, Amsterdam, The Netherlands, September 2008.

[4] L. Zhang, M. Yang, and X. Feng, "Sparse representation or collaborative representation: which helps face recognition?" in *Proceedings of the IEEE International Conference on Computer Vision (ICCV '11)*, pp. 471–478, IEEE, Barcelona, Spain, November 2011.

[5] M. Yang, L. Zhang, D. Zhang, and S. Wang, "Relaxed collaborative representation for pattern classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 2224–2231, IEEE, Providence, RI, USA, June 2012.

[6] X. Mei and H. B. Ling, "Robust visual tracking and vehicle classification via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 11, pp. 2259–2272, 2011.

[7] X. Mei and H. B. Ling, "Robust visual tracking using $l1$ minimization," in *Proceedings of the 12th IEEE Computer Vision*, pp. 1436–1443, IEEE, Kyoto, Japan, September-October 2009.

[8] T. X. Bai and Y. F. Li, "Robust visual tracking with structured sparse representation appearance model," *Pattern Recognition*, vol. 45, no. 6, pp. 2390–2404, 2012.

[9] Y.-E. Hou, W.-G. Li, A.-Q. Rong, and G.-Q. Ye, "Tracking algorithm of block sparse representation with background information," *Journal of South China University of Technology*, vol. 41, no. 8, pp. 21–27, 2013.

[10] X. Jia, H. C. Lu, and M.-H. Yang, "Visual tracking via adaptive structural local sparse appearance model," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 1822–1829, IEEE, Providence, RI, USA, June 2012.

[11] Y.-E. Hou, W.-G. Li, S. Sekou et al., "Target tracking algorithm with structured sparse representation based on ranks," *Journal of South China University of Technology (Natural Science Edition)*, vol. 41, no. 11, pp. 23–29, 2013.

[12] Y. Hou, W. Li, A. Rong, H. Lou, and S. Quan, "Robust visual ℓ_2-regularized least squares tracker with Bayes classifier and coding error," *Journal of Electronic Imaging*, vol. 22, no. 4, Article ID 043036, 2013.

[13] D. Wang, H. Lu, and M.-H. Yang, "Least soft-threshold squares tracking," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '13)*, pp. 2371–2378, Portland, Ore, USA, June 2013.

[14] C. Bao, Y. Wu, H. Ling, and H. Ji, "Real time robust L1 tracker using accelerated proximal gradient approach," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 1830–1837, IEEE, Providence, RI, USA, June 2012.

[15] B. Zhuang, H. Lu, Z. Xiao, and D. Wang, "Visual tracking via discriminative sparse similarity map," *IEEE Transactions on Image Processing*, vol. 23, no. 4, pp. 1872–1881, 2014.

[16] X. Lan, A. J. Ma, and P. C. Yuen, "Multi-cue visual tracking using robust feature-level fusion based on joint sparse representation," in *Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '14)*, pp. 1194–1201, Columbus, Ohio, USA, June 2014.

[17] D. A. Ross, J. Lim, R.-S. Lin, and M.-H. Yang, "Incremental learning for robust visual tracking," *International Journal of Computer Vision*, vol. 77, no. 1–3, pp. 125–141, 2008.

[18] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (VOC) challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303–338, 2010.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 737949, 11 pages
http://dx.doi.org/10.1155/2015/737949

Research Article

Exponential Stabilization of a Class of Time-Varying Delay Systems with Nonlinear Perturbations

Yazhou Tian,[1,2] Yuanli Cai,[1] Yuangong Sun,[3] and Tongxing Li[2]

[1] *School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China*
[2] *Qingdao Technological University, Feixian, Shandong 273400, China*
[3] *School of Mathematical Sciences, University of Jinan, Jinan, Shandong 250022, China*

Correspondence should be addressed to Yuanli Cai; ylicai@mail.xjtu.edu.cn and Tongxing Li; litongx2007@163.com

Received 3 April 2015; Accepted 25 June 2015

Academic Editor: Xinguang Zhang

This paper addresses the problem of exponential stabilization of a class of time-varying delay systems with nonlinear perturbations. These perturbations are related not only with current state $x(t)$ and the delayed state $x(t-h(t))$ but also with $\beta(t)$, where $\beta(t)$ is a continuous function defined on $[0,+\infty)$. With the delay interval divided into two equidistant subintervals, a novel Lyapunov functional is introduced, and several new exponential stabilization criteria are derived in terms of linear matrix inequalities (LMIs) by employing reciprocally convex approach. Two examples are given to illustrate the effectiveness of the main results.

1. Introduction

Time delay is commonly encountered in various physical and engineering systems such as aircraft, biological systems, and networked control systems. Since the existence of time delays causes poor performance, oscillation, or even instability, it is very important to investigate stability analysis for systems with time delays before designing control systems. On the other hand, the systems almost present some uncertainties because it is not easy to obtain an exact mathematical model due to environmental noise, uncertain or slowly varying parameters, and so forth. Therefore, considerable amounts of efforts have been done to the stability and stabilization of time-delay systems and time-delay systems with nonlinear perturbations; see, for example, [1–25] and the references cited therein.

Recently, Zhang et al. [12] considered interval time-varying delay systems and obtained some delay-dependent conditions by employing Finsler's lemma. Combining the descriptor model transformation and the integral inequality method, Han [3] investigated the robust stability of linear systems with time-varying delay and nonlinear perturbations and obtained several improved stability conditions. On the basis of free weighting matrices technique, robust stabilization criteria for neutral systems with nonlinear perturbations were reported in [9]. Wang et al. [6] introduced a new parameter in the Lyapunov functional for the time-varying delay systems with nonlinear perturbations and obtained less conservative results, whereas the range of the time delays considered in the paper was assumed from zero to an upper bound. Note that the stability investigated in the above-mentioned papers was primarily focused on asymptotic stability. Using delay decomposition method and Finsler's lemma, Liu et al. [24] studied the exponential stability of neutral systems with interval time-varying delays and nonlinear perturbations. So far, there are few articles concerning the problem of exponential stabilization of time-varying delay systems with nonlinear perturbations. Thuan et al. [16] provided a detailed analysis for the problem of designing state feedback controllers to exponential stabilization of time-delay systems with nonlinear perturbations by using the integral inequality method and constructing a Lyapunov functional containing the triple integral terms. However, there still exists a gap for reducing both the conservatism and the number of decision variables.

In this paper, we study the exponential stabilization of a class of time-delay systems with nonlinear perturbations. The main contributions of this paper can be summarized as follows: (i) a novel Lyapunov functional containing the center point of time-delay interval is constructed; (ii) compared with the systems studied in [3, 8, 16], the nonlinear perturbations of (1) are related not only with the current state $x(t)$ and the delayed state $x(t-h(t))$ but also with $\beta(t)$, where $\beta(t)$ is a continuous function satisfying $\int_0^{+\infty}\beta^2(s)e^{2\alpha s}ds < +\infty$ and α is a positive constant; new sufficient conditions are obtained that ensure the stability of a closed-loop system, which extend and improve the main results of [16]. Furthermore, the stabilization conditions are shown to be less conservative than those reported in Zhang et al. [12] when there are no nonlinear perturbations in the system. Finally, two numerical examples are presented to demonstrate the effectiveness and advantages of the main results.

Notation. Throughout the paper, R^n denotes the n-dimensional Euclidean space with vector norm $\|\cdot\|$, and $R^{n\times m}$ is the set of all $n\times m$-dimensional real matrices. I denotes the identity matrix of appropriate dimensions, and the superscript "T" stands for matrix transposition. The notation $P > 0$ (≥ 0) means that P is symmetric and positive (semipositive) definite. $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ denote the minimum and maximum eigenvalues of A, respectively. In addition, in symmetric block matrices or long matrix expressions, we use an asterisk $(*)$ to represent a term that is induced by symmetry.

2. Problem Description and Preliminaries

Consider the following system with a nonlinear perturbation:

$$\begin{aligned}\dot{x}(t) &= Ax(t)+Dx(t-h(t))\\ &\quad + f(t,x(t),x(t-h(t)))+Bu(t), \quad t\geq 0,\\ x(t) &= \phi(t), \quad t\in[-h_2,0],\end{aligned} \tag{1}$$

where $x(t)\in R^n$ is the state vector, $u(t)\in R^m$ is the control input vector, $A,D\in R^{n\times n}$, and $B\in R^{n\times m}$, $\phi(t)\in C([-h_2,0],R^n)$ with $\|\phi\| = \sup_{t\in[-h_2,0]}\{\|\phi(t)\|,\|\dot{\phi}(t)\|\}$, where $C([-h_2,0],R^n)$ is the Banach space of continuous functions. The delay $h(t)$ is time-varying and satisfies

$$0\leq h_1\leq h(t)\leq h_2, \quad \dot{h}(t)\leq\mu, \tag{2}$$

where μ is a positive constant and h_1 and h_2 are constants representing the lower and upper bounds of the delay, respectively. $f(t,x(t),x(t-h(t)))$ is a nonlinear perturbation satisfying

$$\begin{aligned}&\|f(t,x(t),x(t-h(t)))\|\\ &\quad\leq a\|x(t)\|+b\|x(t-h(t))\|+\beta(t),\end{aligned} \tag{3}$$

where a and b are positive scalars and $\beta(t)$ satisfies $\int_0^{+\infty}\beta^2(s)e^{2\alpha s}ds < +\infty$.

Remark 1. In [3, 8, 16], the authors assumed that the nonlinear terms satisfy

$$\begin{aligned}&f^T(t,x(t))f(t,x(t))\leq a^2x^T(t)F^TFx(t),\\ &g^T(t,x(t-h(t)))g(t,x(t-h(t)))\\ &\quad\leq b^2x^T(t-h(t))G^TGx(t-h(t)),\end{aligned} \tag{4}$$

where F and G are constant matrices and a and b are positive scalars. It is obvious that the assumptions on the nonlinear terms given in (2) and (3) are more general.

The following definitions and lemmas will be used for providing the main results in the sequel.

Definition 2. Given a scalar $\alpha > 0$: system (1) with $u(t) = 0$ is α-stable if there exists a positive number $\gamma > 0$ such that every solution $x(t,\phi)$ of the system satisfies

$$\|x(t,\phi)\|\leq\gamma e^{-\alpha t}, \quad t\geq 0. \tag{5}$$

Definition 3. Given a scalar $\alpha > 0$: system (1) is α-stabilizable if there exists a state feedback control $u(t) = Kx(t)$ such that the closed-loop system

$$\begin{aligned}\dot{x}(t) &= (A+BK)x(t)+Dx(t-h(t))\\ &\quad + f(t,x(t),x(t-h(t))), \quad t\geq 0,\\ x(t) &= \phi(t), \quad t\in[-h_2,0]\end{aligned} \tag{6}$$

is α-stable.

Lemma 4 (see [25]). *For any $x,y\in R^n$ and a positive symmetric definite matrix $N\in R^{n\times n}$,*

$$\pm 2y^Tx\leq x^TN^{-1}x+y^TNy. \tag{7}$$

Lemma 5 (lower bound lemma for reciprocal convexity; see [20]). *Let $f_1,f_2,\ldots,f_N : R^m\to R$ have positive values in an open subset D of R^m. Then, the reciprocally convex combination of f_i over D satisfies*

$$\min_{\{\alpha_i|\alpha_i>0,\sum_i\alpha_i=1\}}\sum_i\frac{1}{\alpha_i}f_i(t) = \sum_i f_i(t)+\max_{g_{i,j}(t)}\sum_{i\neq j}g_{i,j}(t) \tag{8}$$

subject to

$$\left\{g_{i,j}:R^m\longrightarrow R,\ g_{j,i}(t)\triangleq g_{i,j}(t),\ \begin{bmatrix} f_i(t) & g_{i,j}(t)\\ g_{i,j}(t) & f_j(t)\end{bmatrix}\geq 0\right\}. \tag{9}$$

3. Main Results

We use the following notation for the convenience:

$$\begin{aligned}
\bar{h} &= \frac{h_1+h_2}{2},\\
\delta &= \frac{h_2-h_1}{2},\\
\lambda &= \lambda_{\min}\left(P^{-1}\right),\\
\Gamma &= \lambda_{\max}\left(P^{-1}\right)+h_1\lambda_{\max}\left(P^{-1}Q_1P^{-1}\right)+\left(h_2-h_1\right)\lambda_{\max}\left(P^{-1}Q_2P^{-1}\right)+\left(h_2-h_1\right)\lambda_{\max}\begin{bmatrix} P^{-1}R_{11}P^{-1} & P^{-1}R_{12}P^{-1}\\ * & P^{-1}R_{22}P^{-1}\end{bmatrix}\\
&\quad+\frac{\left(3h_1+h_2\right)\left(h_2-h_1\right)^2}{16}\lambda_{\max}\left(P^{-1}S_1P^{-1}\right)+\frac{\left(h_1+3h_2\right)\left(h_2-h_1\right)^2}{16}\lambda_{\max}\left(P^{-1}S_2P^{-1}\right)+\frac{h_1^3}{2}\lambda_{\max}\left(P^{-1}S_3P^{-1}\right),\\
\Lambda &= \Gamma\left\|\phi\right\|^2+6\int_0^{+\infty}\beta^2(s)\,e^{2\alpha s}ds,\\
\xi^T(t) &= \left[x^T(t)\quad x^T(t-h(t))\quad x^T\left(t-h_1\right)\quad x^T\left(t-\bar{h}\right)\quad x^T\left(t-h_2\right)\quad \dot{x}^T(t)\right].
\end{aligned} \tag{10}$$

The following theorem presents an exponential stabilization condition for (1).

Theorem 6. *Let $\alpha > 0$ and assume that conditions (2) and (3) are satisfied. If there exist matrices $P > 0$, $Q_1 > 0$, $Q_2 > 0$, $S_i > 0$ $(i = 1, 2, 3)$, $\left[\begin{smallmatrix} R_{11} & R_{12}\\ * & R_{22}\end{smallmatrix}\right] > 0$, M_1, M_2, and Y such that the following LMIs hold*

$$\Sigma_1 = \begin{bmatrix}
\Sigma_{11} & DP & \Sigma_{13} & 0 & 0 & PA^T+Y^TB^T & a^2P & 0\\
* & \Sigma_{22} & \Sigma_{23} & \Sigma_{24} & 0 & PD^T & 0 & b^2P\\
* & * & \Sigma_{33} & \Sigma_{34} & 0 & 0 & 0 & 0\\
* & * & * & \Sigma_{44} & \Sigma_{45} & 0 & 0 & 0\\
* & * & * & * & \Sigma_{55} & 0 & 0 & 0\\
* & * & * & * & * & \Sigma_{66} & 0 & 0\\
* & * & * & * & * & * & -\frac{1}{6}a^2I & 0\\
* & * & * & * & * & * & 0 & -\frac{1}{6}b^2I
\end{bmatrix} < 0, \tag{11}$$

$$\Sigma_2 = \begin{bmatrix}
\Sigma_{11} & DP & \Sigma_{13} & 0 & 0 & PA^T+Y^TB^T & a^2P & 0\\
* & \overline{\Sigma}_{22} & 0 & \overline{\Sigma}_{24} & \overline{\Sigma}_{25} & PD^T & 0 & b^2P\\
* & * & \Sigma_{33} & \overline{\Sigma}_{34} & 0 & 0 & 0 & 0\\
* & * & * & \Sigma_{44} & \overline{\Sigma}_{45} & 0 & 0 & 0\\
* & * & * & * & \Sigma_{55} & 0 & 0 & 0\\
* & * & * & * & * & \Sigma_{66} & 0 & 0\\
* & * & * & * & * & * & -\frac{1}{6}a^2I & 0\\
* & * & * & * & * & * & 0 & -\frac{1}{6}b^2I
\end{bmatrix} < 0, \tag{12}$$

$$\begin{bmatrix} S_i & M_i\\ * & S_i\end{bmatrix} \geq 0 \quad (i = 1, 2), \tag{13}$$

where

$$\begin{aligned}
\Sigma_{11} &= AP+PA^T+BY+Y^TB^T+2\alpha P+I+Q_1\\
&\quad -e^{-2\alpha h_1}S_3,\\
\Sigma_{22} &= -\left(1-\mu\right)e^{-2\alpha h_2}Q_2+e^{-2\alpha\bar{h}}M_1+e^{-2\alpha\bar{h}}M_1^T\\
&\quad -2e^{-2\alpha\bar{h}}S_1,\\
\Sigma_{33} &= e^{-2\alpha h_1}Q_2+e^{-2\alpha h_1}R_{11}-e^{-2\alpha h_1}Q_1-e^{-2\alpha\bar{h}}S_1\\
&\quad -e^{-2\alpha h_1}S_3,\\
\Sigma_{44} &= e^{-2\alpha h_1}R_{22}-e^{-2\alpha\bar{h}}R_{11}-e^{-2\alpha\bar{h}}S_1-e^{-2\alpha h_2}S_2,\\
\Sigma_{55} &= -e^{-2\alpha\bar{h}}R_{22}-e^{-2\alpha h_2}S_2,\\
\Sigma_{66} &= h_1^2S_3+\delta^2S_2+\delta^2S_1+I-2P,\\
\Sigma_{13} &= e^{-2\alpha h_1}S_3,\\
\Sigma_{23} &= e^{-2\alpha\bar{h}}\left(S_1-M_1\right),\\
\Sigma_{24} &= e^{-2\alpha\bar{h}}\left(S_1-M_1^T\right),\\
\Sigma_{34} &= e^{-2\alpha h_1}R_{12}+e^{-2\alpha\bar{h}}M_1^T,\\
\Sigma_{45} &= -e^{-2\alpha\bar{h}}R_{12}+e^{-2\alpha h_2}S_2,\\
\overline{\Sigma}_{22} &= -\left(1-\mu\right)e^{-2\alpha h_2}Q_2+e^{-2\alpha h_2}M_2+e^{-2\alpha h_2}M_2^T\\
&\quad -2e^{-2\alpha h_2}S_2,
\end{aligned}$$

$$\begin{aligned}
\overline{\Sigma}_{24} &= e^{-2\alpha h_2}\left(S_2 - M_2\right),\\
\overline{\Sigma}_{25} &= e^{-2\alpha h_2}\left(S_2 - M_2^T\right),\\
\overline{\Sigma}_{34} &= e^{-2\alpha h_1}R_{12} + e^{-2\alpha \overline{h}}S_1,\\
\overline{\Sigma}_{45} &= -e^{-2\alpha \overline{h}}R_{12} + e^{-2\alpha h_2}M_2^T,
\end{aligned} \tag{14}$$

then system (1) is robustly α-stabilizable, the state feedback control $u(t) = YP^{-1}x(t)$, and the solution $x(t,\phi)$ of the closed-loop system satisfies

$$\left\|x(t,\phi)\right\| \le \sqrt{\frac{\Lambda}{\lambda}}e^{-\alpha t}, \quad t \ge 0. \tag{15}$$

Proof. Let us denote

$$\begin{aligned}
\overline{Q}_i &= P^{-1}Q_iP^{-1},\\
\overline{M}_i &= P^{-1}M_iP^{-1}, \quad i = 1,2;\\
\overline{S}_i &= P^{-1}S_iP^{-1}, \quad i = 1,2,3;\\
\overline{R}_{11} &= P^{-1}R_{11}P^{-1},\\
\overline{R}_{12} &= P^{-1}R_{12}P^{-1},\\
\overline{R}_{22} &= P^{-1}R_{22}P^{-1}.
\end{aligned} \tag{16}$$

Define a Lyapunov functional by

$$V\left(x_t\right) = V_1\left(x_t\right) + V_2\left(x_t\right) + V_3\left(x_t\right), \tag{17}$$

where

$$\begin{aligned}
V_1\left(x_t\right) &= x^T(t)P^{-1}x(t),\\
V_2\left(x_t\right) &= \int_{t-h_1}^{t} e^{2\alpha(s-t)}x^T(s)\overline{Q}_1x(s)\,ds\\
&\quad + \int_{t-h(t)}^{t-h_1} e^{2\alpha(s-t)}x^T(s)\overline{Q}_2x(s)\,ds\\
&\quad + \int_{t-\overline{h}}^{t-h_1} e^{2\alpha(s-t)}\begin{bmatrix} x(s)\\ x(s-\delta)\end{bmatrix}^T\\
&\quad \cdot \begin{bmatrix} \overline{R}_{11} & \overline{R}_{12}\\ * & \overline{R}_{22}\end{bmatrix}\begin{bmatrix} x(s)\\ x(s-\delta)\end{bmatrix}ds,\\
V_3\left(x_t\right) &= \delta\int_{-\overline{h}}^{-h_1}\int_{t+s}^{t} e^{2\alpha(\tau-t)}\dot{x}^T(\tau)\overline{S}_1\dot{x}(\tau)\,d\tau\,ds\\
&\quad + \delta\int_{-h_2}^{-\overline{h}}\int_{t+s}^{t} e^{2\alpha(\tau-t)}\dot{x}^T(\tau)\overline{S}_2\dot{x}(\tau)\,d\tau\,ds\\
&\quad + h_1\int_{-h_1}^{0}\int_{t+s}^{t} e^{2\alpha(\tau-t)}\dot{x}^T(\tau)\overline{S}_3\dot{x}(\tau)\,d\tau\,ds.
\end{aligned} \tag{18}$$

Calculating the time derivative of $V(x_t)$ along the trajectories of (6), we conclude that

$$\begin{aligned}
&\dot{V}_1\left(x_t\right)\\
&\quad = x^T(t)\left[P^{-1}(A+BK) + (A+BK)^TP^{-1}\right]x(t)\\
&\qquad + 2x^T(t)P^{-1}Dx(t-h(t))\\
&\qquad + 2x^T(t)P^{-1}f(t,x(t),x(t-h(t))),
\end{aligned} \tag{19}$$

$$\begin{aligned}
&\dot{V}_2\left(x_t\right)\\
&\quad \le -2\alpha V_2\left(x_t\right) + x^T(t)\overline{Q}_1x(t)\\
&\qquad - (1-\mu)e^{-2\alpha h_2}x^T(t-h(t))\overline{Q}_2x(t-h(t))\\
&\qquad + e^{-2\alpha h_1}x^T(t-h_1)\left[\overline{Q}_2 - \overline{Q}_1\right]x(t-h_1)\\
&\qquad + e^{-2\alpha h_1}\begin{bmatrix} x(t-h_1)\\ x(t-\overline{h})\end{bmatrix}^T\begin{bmatrix} \overline{R}_{11} & \overline{R}_{12}\\ * & \overline{R}_{22}\end{bmatrix}\begin{bmatrix} x(t-h_1)\\ x(t-\overline{h})\end{bmatrix}\\
&\qquad - e^{-2\alpha \overline{h}}\begin{bmatrix} x(t-\overline{h})\\ x(t-h_2)\end{bmatrix}^T\begin{bmatrix} \overline{R}_{11} & \overline{R}_{12}\\ * & \overline{R}_{22}\end{bmatrix}\begin{bmatrix} x(t-\overline{h})\\ x(t-h_2)\end{bmatrix},
\end{aligned} \tag{20}$$

$$\begin{aligned}
&\dot{V}_3\left(x_t\right)\\
&\quad \le -2\alpha V_3\left(x_t\right) + \dot{x}^T(t)\left(h_1^2\overline{S}_3 + \delta^2\overline{S}_2 + \delta^2\overline{S}_1\right)\dot{x}(t)\\
&\qquad - \delta e^{-2\alpha h_2}\int_{t-h_2}^{t-\overline{h}}\dot{x}^T(s)\overline{S}_2\dot{x}(s)\,ds\\
&\qquad - \delta e^{-2\alpha \overline{h}}\int_{t-\overline{h}}^{t-h_1}\dot{x}^T(s)\overline{S}_1\dot{x}(s)\,ds\\
&\qquad - h_1e^{-2\alpha h_1}\int_{t-h_1}^{t}\dot{x}^T(s)\overline{S}_3\dot{x}(s)\,ds.
\end{aligned} \tag{21}$$

By virtue of Lemma 4, we have

$$\begin{aligned}
&2x^T(t)P^{-1}f(t,x(t),x(t-h(t)))\\
&\quad \le x^T(t)P^{-1}P^{-1}x(t)\\
&\quad + f^T(t,x(t),x(t-h(t)))f(t,x(t),x(t-h(t)))\\
&\quad = x^T(t)P^{-1}P^{-1}x(t) + \left\|f(t,x(t),x(t-h(t)))\right\|^2.
\end{aligned} \tag{22}$$

Using (3) and the inequality $(H+J+L)^2 \le 3H^2 + 3J^2 + 3L^2$, where H, J, and L are constants, we obtain

$$\begin{aligned}
&\left\|f(t,x(t),x(t-h(t)))\right\|^2\\
&\quad \le \left(a\left\|x(t)\right\| + b\left\|x(t-h(t))\right\| + \beta(t)\right)^2\\
&\quad \le 3a^2\left\|x(t)\right\|^2 + 3b^2\left\|x(t-h(t))\right\|^2 + 3\beta^2(t).
\end{aligned} \tag{23}$$

It follows from (22) and (23) that

$$\begin{aligned} &2x^T(t)P^{-1}f(t,x(t),x(t-h(t))) \\ &\quad\le x^T(t)\left(P^{-1}P^{-1}+3a^2I\right)x(t) \\ &\qquad+3b^2x^T(t-h(t))x(t-h(t))+3\beta^2(t). \end{aligned} \tag{24}$$

Combining (19) and (24), we get

$$\begin{aligned} \dot{V}_1(x_t) \le{}& -2\alpha V_1(x_t)+x^T(t)\left[P^{-1}(A+BK)\right. \\ &\left.+(A+BK)^TP^{-1}+2\alpha P^{-1}+P^{-1}P^{-1}+3a^2I\right]x(t) \\ &+2x^T(t)P^{-1}Dx(t-h(t))+3b^2x^T(t-h(t))x(t \\ &-h(t))+3\beta^2(t). \end{aligned} \tag{25}$$

Now, we estimate the upper bounds of the last three integral terms in inequality (21) as follows.

(i) Assume first that $h_1 \le h(t) \le \overline{h}$. From Jensen's inequality [19] and Lemma 5,

$$\begin{aligned} &-\delta\int_{t-\overline{h}}^{t-h_1}\dot{x}^T(s)\overline{S}_1\dot{x}(s)\,ds \\ &\quad=-\delta\int_{t-\overline{h}}^{t-h(t)}\dot{x}^T(s)\overline{S}_1\dot{x}(s)\,ds \\ &\qquad-\delta\int_{t-h(t)}^{t-h_1}\dot{x}^T(s)\overline{S}_1\dot{x}(s)\,ds \\ &\quad\le-\frac{\delta}{\overline{h}-h(t)}\left(\int_{t-\overline{h}}^{t-h(t)}\dot{x}(s)\,ds\right)^T \\ &\qquad\cdot\overline{S}_1\left(\int_{t-\overline{h}}^{t-h(t)}\dot{x}(s)\,ds\right) \\ &\qquad-\frac{\delta}{h(t)-h_1}\left(\int_{t-h(t)}^{t-h_1}\dot{x}(s)\,ds\right)^T \\ &\qquad\cdot\overline{S}_1\left(\int_{t-h(t)}^{t-h_1}\dot{x}(s)\,ds\right) \\ &\quad\le-\begin{bmatrix} x(t-h(t))-x\left(t-\overline{h}\right) \\ x(t-h_1)-x(t-h(t)) \end{bmatrix}^T \\ &\qquad\cdot\begin{bmatrix} \overline{S}_1 & \overline{M}_1 \\ * & \overline{S}_1 \end{bmatrix}\begin{bmatrix} x(t-h(t))-x\left(t-\overline{h}\right) \\ x(t-h_1)-x(t-h(t)) \end{bmatrix}, \end{aligned} \tag{26}$$

and so

$$\begin{aligned} &-\delta e^{-2\alpha\overline{h}}\int_{t-\overline{h}}^{t-h_1}\dot{x}^T(s)\overline{S}_1\dot{x}(s)\,ds \\ &\quad\le-\begin{bmatrix} x(t-h(t))-x\left(t-\overline{h}\right) \\ x(t-h_1)-x(t-h(t)) \end{bmatrix}^T \\ &\qquad\cdot\begin{bmatrix} e^{-2\alpha\overline{h}}\overline{S}_1 & e^{-2\alpha\overline{h}}\overline{M}_1 \\ * & e^{-2\alpha\overline{h}}\overline{S}_1 \end{bmatrix}\begin{bmatrix} x(t-h(t))-x\left(t-\overline{h}\right) \\ x(t-h_1)-x(t-h(t)) \end{bmatrix}. \end{aligned} \tag{27}$$

On the other hand, if $h(t) = \overline{h}$ or $h(t) = h_1$, then $\int_{t-\overline{h}}^{t-h(t)}\dot{x}(s)ds = 0$ or $\int_{t-h(t)}^{t-h_1}\dot{x}(s)ds = 0$, respectively. Hence, inequality (27) holds.

Using Jensen's inequality [19], it is not difficult to arrive at the inequalities

$$\begin{aligned} &-h_1e^{-2\alpha h_1}\int_{t-h_1}^{t}\dot{x}^T(s)\overline{S}_3\dot{x}(s)\,ds \\ &\quad\le-\left[x(t)-x(t-h_1)\right]^T \\ &\qquad\cdot e^{-2\alpha h_1}\overline{S}_3\left[x(t)-x(t-h_1)\right], \end{aligned} \tag{28}$$

$$\begin{aligned} &-\delta e^{-2\alpha h_2}\int_{t-h_2}^{t-\overline{h}}\dot{x}^T(s)\overline{S}_2\dot{x}(s)\,ds \\ &\quad\le-\left[x\left(t-\overline{h}\right)-x(t-h_2)\right]^T \\ &\qquad\cdot e^{-2\alpha h_2}\overline{S}_2\left[x\left(t-\overline{h}\right)-x(t-h_2)\right]. \end{aligned} \tag{29}$$

It follows now from (6) that

$$\begin{aligned} &2\dot{x}^T(t)P^{-1}\left[(A+BK)x(t)+Dx(t-h(t))\right. \\ &\qquad\left.+f(t,x(t),x(t-h(t)))-\dot{x}(t)\right]=0. \end{aligned} \tag{30}$$

On the other hand,

$$\begin{aligned} &2\dot{x}^T(t)P^{-1}f(t,x(t),x(t-h(t))) \\ &\quad\le\dot{x}^T(t)P^{-1}P^{-1}\dot{x}(t) \\ &\qquad+f^T(t,x(t),x(t-h(t)))f(t,x(t),x(t-h(t))) \\ &\quad=\dot{x}^T(t)P^{-1}P^{-1}\dot{x}(t)+\left\|f(t,x(t),x(t-h(t)))\right\|^2 \\ &\quad\le\dot{x}^T(t)P^{-1}P^{-1}\dot{x}(t)+3a^2x^T(t)x(t) \\ &\qquad+3b^2x^T(t-h(t))x(t-h(t))+3\beta^2(t). \end{aligned} \tag{31}$$

Therefore, formulas (19)–(31) imply that

$$\dot{V}(x_t)+2\alpha V(x_t)\le\xi^T(t)\Phi_1\xi(t)+6\beta^2(t), \tag{32}$$

where

$$\Phi_1 = \begin{bmatrix} \Phi_{11} & P^{-1}D & \Phi_{13} & 0 & 0 & (A+BK)^TP^{-1} \\ * & \Phi_{22} & \Phi_{23} & \Phi_{24} & 0 & D^TP^{-1} \\ * & * & \Phi_{33} & \Phi_{34} & 0 & 0 \\ * & * & * & \Phi_{44} & \Phi_{45} & 0 \\ * & * & * & * & \Phi_{55} & 0 \\ * & * & * & * & * & \Phi_{66} \end{bmatrix},$$

$$
\begin{aligned}
&\Phi_{11} \\
&\quad = P^{-1}(A+BK)+(A+BK)^T P^{-1}+2\alpha P^{-1} \\
&\qquad + P^{-1}P^{-1}+\overline{Q}_1-e^{-2\alpha h_1}\overline{S}_3+6a^2 I, \\
&\Phi_{22} \\
&\quad = -(1-\mu)e^{-2\alpha h_2}\overline{Q}_2+e^{-2\alpha\overline{h}}\overline{M}_1+e^{-2\alpha\overline{h}}\overline{M}_1^T \\
&\qquad - 2e^{-2\alpha\overline{h}}\overline{S}_1+6b^2 I, \\
&\Phi_{33} \\
&\quad = e^{-2\alpha h_1}\overline{Q}_2+e^{-2\alpha h_1}\overline{R}_{11}-e^{-2\alpha h_1}\overline{Q}_1-e^{-2\alpha\overline{h}}\overline{S}_1 \\
&\qquad - e^{-2\alpha h_1}\overline{S}_3, \\
&\Phi_{44} = e^{-2\alpha h_1}\overline{R}_{22}-e^{-2\alpha\overline{h}}\overline{R}_{11}-e^{-2\alpha\overline{h}}\overline{S}_1-e^{-2\alpha h_2}\overline{S}_2, \\
&\Phi_{55} = -e^{-2\alpha\overline{h}}\overline{R}_{22}-e^{-2\alpha h_2}\overline{S}_2, \\
&\Phi_{66} = h_1^2\overline{S}_3+\delta^2\overline{S}_2+\delta^2\overline{S}_1+P^{-1}P^{-1}-2P^{-1}, \\
&\Phi_{13} = e^{-2\alpha h_1}\overline{S}_3, \\
&\Phi_{23} = e^{-2\alpha\overline{h}}\left(\overline{S}_1-\overline{M}_1\right), \\
&\Phi_{24} = e^{-2\alpha\overline{h}}\left(\overline{S}_1-\overline{M}_1^T\right), \\
&\Phi_{34} = e^{-2\alpha h_1}\overline{R}_{12}+e^{-2\alpha\overline{h}}\overline{M}_1^T, \\
&\Phi_{45} = -e^{-2\alpha\overline{h}}\overline{R}_{12}+e^{-2\alpha h_2}\overline{S}_2.
\end{aligned}
\tag{33}
$$

If we pre- and postmultiply Φ_1 by diag$\{P, P, P, P, P, P\}$ and let

$$K = YP^{-1}, \tag{34}$$

then the condition $\Phi_1 < 0$ is equivalent to condition (11) by using Schur Complement Lemma.

(ii) Assume now that $\overline{h} \le h(t) \le h_2$. Applications of Jensen's inequality [19] and Lemma 5 yield

$$
\begin{aligned}
&-\delta\int_{t-h_2}^{t-\overline{h}}\dot{x}^T(s)\overline{S}_2\dot{x}(s)\,ds \\
&\quad = -\delta\left[\int_{t-h_2}^{t-h(t)}\dot{x}^T(s)\overline{S}_2\dot{x}(s)\,ds\right. \\
&\qquad \left.+\int_{t-h(t)}^{t-\overline{h}}\dot{x}^T(s)\overline{S}_2\dot{x}(s)\,ds\right] \\
&\quad \le -\frac{\delta}{h_2-h(t)}\left(\int_{t-h_2}^{t-h(t)}\dot{x}(s)\,ds\right)^T \\
&\qquad \cdot\overline{S}_2\left(\int_{t-h_2}^{t-h(t)}\dot{x}(s)\,ds\right) \\
&\qquad -\frac{\delta}{h(t)-\overline{h}}\left(\int_{t-h(t)}^{t-\overline{h}}\dot{x}(s)\,ds\right)^T\overline{S}_2\left(\int_{t-h(t)}^{t-\overline{h}}\dot{x}(s)\,ds\right) \\
&\quad \le -\begin{bmatrix} x(t-h(t))-x(t-h_2) \\ x(t-\overline{h})-x(t-h(t)) \end{bmatrix}^T \\
&\qquad \cdot\begin{bmatrix} \overline{S}_2 & \overline{M}_2 \\ * & \overline{S}_2 \end{bmatrix}\begin{bmatrix} x(t-h(t))-x(t-h_2) \\ x(t-\overline{h})-x(t-h(t)) \end{bmatrix},
\end{aligned}
\tag{35}
$$

and hence

$$
\begin{aligned}
&-\delta e^{-2\alpha h_2}\int_{t-h_2}^{t-\overline{h}}\dot{x}^T(s)\overline{S}_2\dot{x}(s)\,ds \\
&\quad \le -\begin{bmatrix} x(t-h(t))-x(t-h_2) \\ x(t-\overline{h})-x(t-h(t)) \end{bmatrix}^T \\
&\quad \cdot\begin{bmatrix} e^{-2\alpha h_2}\overline{S}_2 & e^{-2\alpha h_2}\overline{M}_2 \\ * & e^{-2\alpha h_2}\overline{S}_2 \end{bmatrix}\begin{bmatrix} x(t-h(t))-x(t-h_2) \\ x(t-\overline{h})-x(t-h(t)) \end{bmatrix}.
\end{aligned}
\tag{36}
$$

Note that when $h(t) = \overline{h}$ or $h(t) = h_2$, one can obtain $\int_{t-\overline{h}}^{t-h(t)}\dot{x}(s)ds = 0$ or $\int_{t-h(t)}^{t-h_2}\dot{x}(s)ds = 0$, respectively. Therefore, inequality (36) is satisfied.

Using Jensen's inequality [19], we deduce that

$$
\begin{aligned}
&-\delta e^{-2\alpha\overline{h}}\int_{t-\overline{h}}^{t-h_1}\dot{x}^T(s)\overline{S}_1\dot{x}(s)\,ds \\
&\quad \le -\left[x(t-h_1)-x(t-\overline{h})\right]^T \\
&\quad \cdot e^{-2\alpha\overline{h}}\overline{S}_1\left[x(t-h_1)-x(t-\overline{h})\right].
\end{aligned}
\tag{37}
$$

Combining (19)–(25), (28), and (30)–(37), we obtain

$$\dot{V}(x_t)+2\alpha V(x_t) \le \xi^T(t)\Phi_2\xi(t)+6\beta^2(t), \tag{38}$$

where

$$
\begin{aligned}
&\Phi_2 \\
&\quad = \begin{bmatrix}
\Phi_{11} & P^{-1}D & \Phi_{13} & 0 & 0 & (A+BK)^T P^{-1} \\
* & \overline{\Phi}_{22} & 0 & \overline{\Phi}_{24} & \overline{\Phi}_{25} & D^T P^{-1} \\
* & * & \Phi_{33} & \overline{\Phi}_{34} & 0 & 0 \\
* & * & * & \Phi_{44} & \overline{\Phi}_{45} & 0 \\
* & * & * & * & \Phi_{55} & 0 \\
* & * & * & * & * & \Phi_{66}
\end{bmatrix}, \\
&\overline{\Phi}_{22} \\
&\quad = -(1-\mu)e^{-2\alpha h_2}\overline{Q}_2+e^{-2\alpha h_2}\overline{M}_2+e^{-2\alpha h_2}\overline{M}_2^T \\
&\qquad - 2e^{-2\alpha h_2}\overline{S}_2+6b^2 I,
\end{aligned}
$$

$$\begin{aligned}
\overline{\Phi}_{24} &= e^{-2\alpha h_2}\left(\overline{S}_2 - \overline{M}_2\right),\\
\overline{\Phi}_{25} &= e^{-2\alpha h_2}\left(\overline{S}_2 - \overline{M}_2^T\right),\\
\overline{\Phi}_{34} &= e^{-2\alpha h_1}\overline{R}_{12} + e^{-2\alpha\overline{h}}\overline{S}_1,\\
\overline{\Phi}_{45} &= -e^{-2\alpha\overline{h}}\overline{R}_{12} + e^{-2\alpha h_2}\overline{M}_2^T.
\end{aligned} \tag{39}$$

Pre- and postmultiplying Φ_2 by $\mathrm{diag}\{P, P, P, P, P, P\}$ and letting $K = YP^{-1}$, the condition $\Phi_2 < 0$ is equivalent to condition (12) by using Schur Complement Lemma.

From the above discussion, if conditions (11)–(13) are satisfied, then

$$\dot{V}(x_t) + 2\alpha V(x_t) \le 6\beta^2(t), \quad t \ge 0. \tag{40}$$

By virtue of (40) and the definition of $V(x(t))$,

$$\begin{aligned}
\lambda \|x(t,\phi)\|^2 &\le V(x(t))\\
&\le e^{-2\alpha t}\left(V(x(0)) + 6\int_0^t \beta^2(s)\, e^{\int_0^s 2\alpha d\tau} ds\right)\\
&\le e^{-2\alpha t}\left(\Gamma\|\phi\|^2 + 6\int_0^{+\infty} \beta^2(s)\, e^{2\alpha s} ds\right) = \Lambda e^{-2\alpha t}.
\end{aligned} \tag{41}$$

Hence, we have

$$\|x(t,\phi)\| \le \sqrt{\frac{\Lambda}{\lambda}} e^{-\alpha t}, \quad t \ge 0, \tag{42}$$

which implies that the closed-loop system is α-stable. The proof is completed. □

If there is no perturbation in system (1), that is, $f(t, x(t), x(t-h(t))) = 0$, then system (1) reduces to

$$\begin{aligned}
\dot{x}(t) &= Ax(t) + Dx(t-h(t)) + Bu(t), \quad t \ge 0,\\
x(t) &= \phi(t), \quad t \in [-h_2, 0].
\end{aligned} \tag{43}$$

Application of Theorem 6 yields the following result.

Corollary 7. *Assume that $\alpha > 0$ and condition (2) is satisfied. If there exist matrices $P > 0$, $Q_1 > 0$, $Q_2 > 0$, $S_i > 0$ $(i = 1, 2, 3)$, $\begin{bmatrix} R_{11} & R_{12} \\ * & R_{22} \end{bmatrix} > 0$, M_1, M_2, and Y such that*

$$\Xi_1 = \begin{bmatrix}
\Xi_{11} & DP & \Xi_{13} & 0 & 0 & PA^T + Y^TB^T\\
* & \Xi_{22} & \Xi_{23} & \Xi_{24} & 0 & PD^T\\
* & * & \Xi_{33} & \Xi_{34} & 0 & 0\\
* & * & * & \Xi_{44} & \Xi_{45} & 0\\
* & * & * & * & \Xi_{55} & 0\\
* & * & * & * & * & \Xi_{66}
\end{bmatrix} < 0,$$

$$\Xi_2 = \begin{bmatrix}
\Xi_{11} & DP & \Xi_{13} & 0 & 0 & PA^T + Y^TB^T\\
* & \overline{\Xi}_{22} & 0 & \overline{\Xi}_{24} & \overline{\Xi}_{25} & PD^T\\
* & * & \Xi_{33} & \overline{\Xi}_{34} & 0 & 0\\
* & * & * & \Xi_{44} & \overline{\Xi}_{45} & 0\\
* & * & * & * & \Xi_{55} & 0\\
* & * & * & * & * & \Xi_{66}
\end{bmatrix} < 0,$$

$$\begin{bmatrix} S_i & M_i \\ * & S_i \end{bmatrix} \ge 0 \quad (i = 1, 2), \tag{44}$$

where

$$\begin{aligned}
\Xi_{11} &= AP + PA^T + BY + Y^TB^T + 2\alpha P + Q_1\\
&\quad - e^{-2\alpha h_1}S_3,\\
\Xi_{22} &= -(1-\mu)e^{-2\alpha h_2}Q_2 + e^{-2\alpha\overline{h}}M_1 + e^{-2\alpha\overline{h}}M_1^T\\
&\quad - 2e^{-2\alpha\overline{h}}S_1,\\
\Xi_{33} &= e^{-2\alpha h_1}Q_2 + e^{-2\alpha h_1}R_{11} - e^{-2\alpha h_1}Q_1 - e^{-2\alpha\overline{h}}S_1\\
&\quad - e^{-2\alpha h_1}S_3,\\
\Xi_{44} &= e^{-2\alpha h_1}R_{22} - e^{-2\alpha\overline{h}}R_{11} - e^{-2\alpha\overline{h}}S_1 - e^{-2\alpha h_2}S_2,\\
\Xi_{55} &= -e^{-2\alpha\overline{h}}R_{22} - e^{-2\alpha h_2}S_2,\\
\Xi_{66} &= h_1^2S_3 + \delta^2S_2 + \delta^2S_1 - 2P,\\
\Xi_{13} &= e^{-2\alpha h_1}S_3,\\
\Xi_{23} &= e^{-2\alpha\overline{h}}(S_1 - M_1),\\
\Xi_{24} &= e^{-2\alpha\overline{h}}\left(S_1 - M_1^T\right),\\
\Xi_{34} &= e^{-2\alpha h_1}R_{12} + e^{-2\alpha\overline{h}}M_1^T,\\
\Xi_{45} &= -e^{-2\alpha\overline{h}}R_{12} + e^{-2\alpha h_2}S_2,\\
\overline{\Xi}_{22} &= -(1-\mu)e^{-2\alpha h_2}Q_2 + e^{-2\alpha h_2}M_2 + e^{-2\alpha h_2}M_2^T\\
&\quad - 2e^{-2\alpha h_2}S_2,\\
\overline{\Xi}_{24} &= e^{-2\alpha h_2}(S_2 - M_2),\\
\overline{\Xi}_{25} &= e^{-2\alpha h_2}\left(S_2 - M_2^T\right),\\
\overline{\Xi}_{34} &= e^{-2\alpha h_1}R_{12} + e^{-2\alpha\overline{h}}S_1,\\
\overline{\Xi}_{45} &= -e^{-2\alpha\overline{h}}R_{12} + e^{-2\alpha h_2}M_2^T,
\end{aligned} \tag{45}$$

then system (43) is robustly α-stabilizable, the state feedback control $u(t) = YP^{-1}x(t)$, and the solution $x(t,\phi)$ of the closed-loop system satisfies

$$\|x(t,\phi)\| \le \sqrt{\frac{\Lambda}{\lambda}} e^{-\alpha t} = \sqrt{\frac{\Gamma}{\lambda}} e^{-\alpha t}\|\phi\|, \quad t \ge 0. \tag{46}$$

4. Numerical Examples

In this section, two numerical examples are given to illustrate the effectiveness of the results obtained in this paper.

Example 1. Consider the system with a nonlinear perturbation

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Dx(t-h(t)) \\ &\quad + f(t, x(t), x(t-h(t))) + Bu(t), \\ &\qquad\qquad\qquad\qquad t \geq 0, \\ x(t) &= \begin{bmatrix}1 & 1\end{bmatrix}^T, \quad t \in [-0.4, 0],\end{aligned} \tag{47}$$

where

$$\begin{aligned}A &= \begin{bmatrix}1 & 0 \\ 0 & -1\end{bmatrix}, \\ B &= \begin{bmatrix}0 \\ 1\end{bmatrix}, \\ D &= \begin{bmatrix}-2 & -0.1 \\ 0 & 1.1\end{bmatrix},\end{aligned} \tag{48}$$

$h(t) = 0.1 + 0.3\sin^2(5t/3)$, and the nonlinear perturbation satisfies

$$\begin{aligned}&\|f(t, x(t), x(t-h(t)))\| \\ &\quad \leq 0.1\|x(t)\| + 0.1\|x(t-h(t))\| + e^{-0.66t}.\end{aligned} \tag{49}$$

Note that $h_1 = 0.1$, $h_2 = 0.4$, $\mu = 0.5$, and $\beta(t) = e^{-0.66t}$. It is not difficult to check that A and $A + D$ are Hurwitz unstable.

Let $\alpha = 0.16$. Then $\int_0^{+\infty} \beta^2(s)e^{2\alpha s}ds = \int_0^{+\infty} e^{-s}ds = 1 < +\infty$. Using the LMI Toolbox in MATLAB, LMIs (11)–(13) in Theorem 6 are satisfied with

$$\begin{aligned}P &= \begin{bmatrix}4.4365 & -0.1306 \\ -0.1306 & 4.9389\end{bmatrix}, \\ Q_1 &= \begin{bmatrix}0.1510 & 0.4068 \\ 0.4068 & 3.4481\end{bmatrix}, \\ Q_2 &= \begin{bmatrix}0.0115 & 0.0066 \\ 0.0066 & 0.2019\end{bmatrix}, \\ S_1 &= \begin{bmatrix}62.2011 & -5.1166 \\ -5.1166 & 52.7981\end{bmatrix}, \\ S_2 &= \begin{bmatrix}36.2972 & -1.8408 \\ -1.8408 & 31.9872\end{bmatrix}, \\ S_3 &= \begin{bmatrix}71.6879 & -4.4063 \\ -4.4063 & 40.8855\end{bmatrix}, \\ R_{11} &= \begin{bmatrix}12.6433 & -1.6878 \\ -1.6878 & 21.3645\end{bmatrix}, \\ R_{12} &= \begin{bmatrix}-12.6769 & 2.2144 \\ 2.2144 & -18.8952\end{bmatrix}, \\ R_{22} &= \begin{bmatrix}13.2723 & -1.3613 \\ -1.3613 & 21.8765\end{bmatrix}, \\ M_1 &= \begin{bmatrix}11.3018 & -6.7305 \\ -6.7305 & 20.9430\end{bmatrix}, \\ M_2 &= \begin{bmatrix}-14.5051 & -2.3510 \\ -2.3510 & 0.3503\end{bmatrix}, \\ Y &= \begin{bmatrix}0.2410 & -6.8850\end{bmatrix}.\end{aligned} \tag{50}$$

Furthermore, the solution $x(t, \phi)$ of the system satisfies

$$\|x(t, \phi)\| \leq 6.1258e^{-0.16t}, \quad t \geq 0, \tag{51}$$

and the stabilizing feedback control

$$u(t) = \begin{bmatrix}0.0133 & -1.3937\end{bmatrix} x(t), \quad t \geq 0. \tag{52}$$

Observe that the results reported in [3, 8, 16] cannot be applied to (47) since the nonlinear perturbation is related with the term $\beta(t) = e^{-0.66t}$.

Example 2. Consider a linear system with an interval time-varying delay

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Dx(t-h(t)) + Bu(t), \quad t \geq 0, \\ x(t) &= \phi(t), \quad \forall t \in [-h_2, 0],\end{aligned} \tag{53}$$

where

$$\begin{aligned}A &= \begin{bmatrix}0 & 0 \\ 0 & 1\end{bmatrix}, \\ B &= \begin{bmatrix}0 \\ 1\end{bmatrix}, \\ D &= \begin{bmatrix}-1 & -1 \\ 0 & -0.9\end{bmatrix},\end{aligned} \tag{54}$$

and $h(t) = 0.5 + 1.28\sin^2(25t/64)$. Note that $h_1 = 0.5$, $h_2 = 1.78$, and $\mu = 0.5$. It is easy to check that A and $A + D$ are Hurwitz unstable. Given $\alpha = 0.01$, using the LMI Toolbox in MATLAB, LMIs (44) in Corollary 7 are satisfied with

$$\begin{aligned}P &= \begin{bmatrix}84.8721 & -32.8554 \\ -32.8554 & 18.7095\end{bmatrix}, \\ Q_1 &= \begin{bmatrix}0.0864 & -0.0424 \\ -0.0424 & 11.7837\end{bmatrix}, \\ Q_2 &= \begin{bmatrix}0.2099 & 0.2035 \\ 0.2035 & 11.4665\end{bmatrix},\end{aligned}$$

$$S_1 = \begin{bmatrix} 109.3732 & -42.3701 \\ -42.3701 & 16.5117 \end{bmatrix},$$

$$S_2 = \begin{bmatrix} 102.7380 & -39.7555 \\ -39.7555 & 15.5155 \end{bmatrix},$$

$$S_3 = \begin{bmatrix} 124.7698 & -48.3825 \\ -48.3825 & 18.8240 \end{bmatrix},$$

$$R_{11} = \begin{bmatrix} 32.4350 & -12.8325 \\ -12.8325 & 5.1838 \end{bmatrix},$$

$$R_{12} = \begin{bmatrix} -2.3196 & 0.9331 \\ 0.9331 & -0.3969 \end{bmatrix},$$

$$R_{22} = \begin{bmatrix} 54.1560 & -21.0105 \\ -21.0105 & 8.2532 \end{bmatrix},$$

$$M_1 = \begin{bmatrix} -22.9733 & 9.4881 \\ 9.4881 & -3.9352 \end{bmatrix},$$

$$M_2 = \begin{bmatrix} -53.6877 & 20.8764 \\ 20.8764 & -8.2092 \end{bmatrix},$$

$$Y = \begin{bmatrix} 23.8997 & -27.0789 \end{bmatrix}. \tag{55}$$

Moreover, the solution $x(t, \phi)$ of the system satisfies

$$\|x(t, \phi)\| \leq 9.3399 e^{-0.01t} \|\phi\|, \quad t \geq 0, \tag{56}$$

and the stabilizing feedback control

$$u(t) = \begin{bmatrix} -0.8704 & -2.9758 \end{bmatrix} x(t), \quad t \geq 0. \tag{57}$$

Figure 1 shows the trajectories of $x_1(t)$ and $x_2(t)$ of the open-loop system with the initial condition $\phi(t) = [30 \quad 10]^T$, $t \in [-1.78, 0]$. Figure 2 shows the trajectories of $x_1(t)$ and $x_2(t)$ of the closed-loop system with the state feedback $u(t) = [-0.8704 \quad -2.9758]x(t)$ and the initial condition $\phi(t) = [30 \quad 10]^T$, $t \in [-1.78, 0]$.

Letting the lower and upper bounds of the time delay be the same as in Zhang et al. [12], our results also ensure exponential stability with an α-convergence rate as given in Table 1. Note that Zhang et al. [12] discussed asymptotic stability, whereas the controller derived in this paper provides exponential stability for the closed-loop system. Furthermore, the maximum bound for α is better than Thuan et al. [16] by letting h_1 and h_2 be the same as in [16]. For selected h_1 and $\mu = 0.5$, using Corollary 7, one can easily observe that the maximum allowable delay bounds for h_2 are better than those reported in the papers by Zhang et al. [12] and Thuan et al. [16].

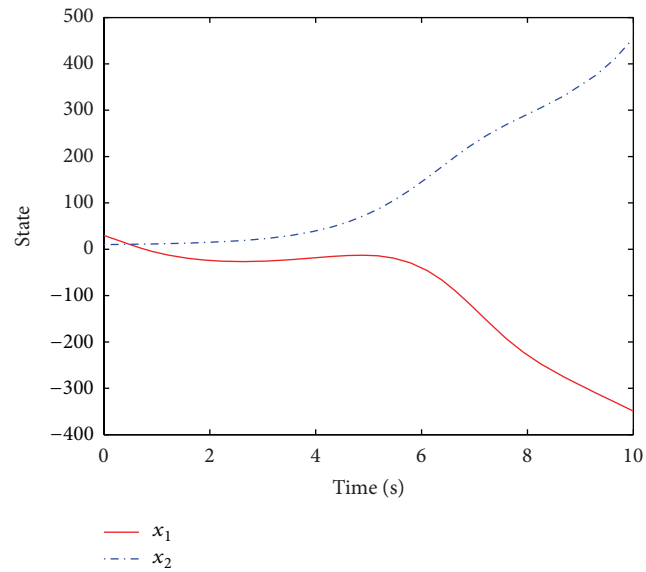


Figure 1: Open-loop trajectories of $x_1(t)$ and $x_2(t)$.

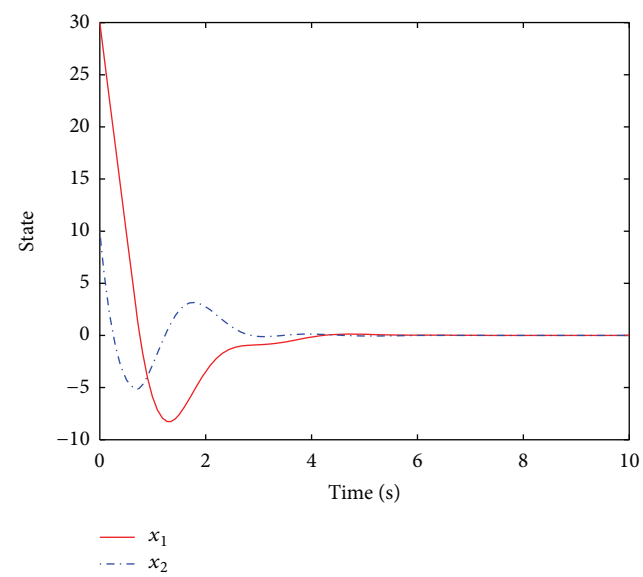


Figure 2: Closed-loop trajectories of $x_1(t)$ and $x_2(t)$.

5. Conclusions

In this paper, exponential stabilization of a class of time-varying delay systems with nonlinear perturbations has been investigated. By using the delay decomposition approach and constructing a novel Lyapunov functional, some new delay-dependent stabilization criteria are obtained in order to ensure closed-loop stability of the system with any prescribed α-convergence rate. Numerical examples are given to illustrate that the results obtained are much less conservative than some existing results in the literature. Exponential stabilization of impulsive switched delay systems with nonlinear perturbations will be further investigated in the future.

Table 1: Admissible upper bound h_2 and K for given h_1 with $\mu = 0.5$.

Method	h_1	h_2	K
Zhang et al. [12]	0.5	0.967	$\begin{bmatrix} -0.5290 & -2.7166 \end{bmatrix}$
Thuan et al. [16] ($\alpha = 0.314$)	0.5	0.967	$\begin{bmatrix} -0.0967 & -1.8338 \end{bmatrix}$
Corollary 7 ($\alpha = 0.353$)	0.5	0.967	$\begin{bmatrix} -0.1662 & -2.0605 \end{bmatrix}$
Thuan et al. [16] ($\alpha = 0.01$)	0.5	1.66	$\begin{bmatrix} -0.8765 & -2.9873 \end{bmatrix}$
Corollary 7 ($\alpha = 0.01$)	0.5	1.78	$\begin{bmatrix} -0.8704 & -2.9758 \end{bmatrix}$
Zhang et al. [12]	1.0	1.114	$\begin{bmatrix} -0.3096 & -2.1894 \end{bmatrix}$
Thuan et al. [16] ($\alpha = 0.225$)	1.0	1.114	$\begin{bmatrix} -0.6344 & -2.9059 \end{bmatrix}$
Corollary 7 ($\alpha = 0.236$)	1.0	1.114	$\begin{bmatrix} -0.6714 & -2.9317 \end{bmatrix}$
Thuan et al. [16] ($\alpha = 0.01$)	1.0	1.672	$\begin{bmatrix} -0.9023 & -2.9983 \end{bmatrix}$
Corollary 7 ($\alpha = 0.01$)	1.0	1.765	$\begin{bmatrix} -0.8701 & -2.9712 \end{bmatrix}$

Conflict of Interests

The authors declare that they have no conflict of interests.

Acknowledgments

The authors express their sincere gratitude to the editors and three anonymous referees for their constructive comments and suggestions that helped to improve the presentation of the results and accentuate important details. This work was supported by the Natural Science Foundation of Shandong Province under Grant no. JQ201119 and the National Natural Science Foundation of China under Grant nos. 61174217, 61374074, and 61473133.

References

[1] J. K. Hale and S. M. V. Lunel, *Introduction to Functional Differential Equations*, Springer, New York, NY, USA, 1993.

[2] M. Malek-Zavarei and M. Jamshidi, *Time-Delay Systems: Analysis, Optimization and Applications*, North-Holland, Amsterdam, The Netherklands, 1987.

[3] Q.-L. Han, "Robust stability for a class of linear systems with time-varying delay and nonlinear perturbations," *Computers & Mathematics with Applications*, vol. 47, no. 8-9, pp. 1201–1209, 2004.

[4] F. Qiu, B. T. Cui, and Y. Ji, "Further results on robust stability of neutral system with mixed time-varying delays and nonlinear perturbations," *Nonlinear Analysis. Real World Applications*, vol. 11, no. 2, pp. 895–906, 2010.

[5] R. Rakkiyappan, P. Balasubramaniam, and R. Krishnasamy, "Delay dependent stability analysis of neutral systems with mixed time-varying delays and nonlinear perturbations," *Journal of Computational and Applied Mathematics*, vol. 235, no. 8, pp. 2147–2156, 2011.

[6] W. Wang, S. K. Nguang, S. Zhong, and F. Liu, "Novel delay-dependent stability criterion for time-varying delay systems with parameter uncertainties and nonlinear perturbations," *Information Sciences*, vol. 281, pp. 321–333, 2014.

[7] L. D. Guo, H. Gu, J. Xing, and X. Q. He, "Asymptotic and exponential stability of uncertain system with interval delay," *Applied Mathematics and Computation*, vol. 218, no. 19, pp. 9997–10006, 2012.

[8] W. Zhang, X.-S. Cai, and Z.-Z. Han, "Robust stability criteria for systems with interval time-varying delay and nonlinear perturbations," *Journal of Computational and Applied Mathematics*, vol. 234, no. 1, pp. 174–180, 2010.

[9] M. Li and G. D. Hu, "Delay-dependent robust stabilization for a class of neutral systems with nonlinear perturbations," *Journal of Control Theory and Applications*, vol. 5, no. 4, pp. 409–414, 2007.

[10] T. Li, L. Guo, and Y. M. Zhang, "Delay-range-dependent robust stability and stabilization for uncertain systems with time-varying delay," *International Journal of Robust and Nonlinear Control*, vol. 18, no. 13, pp. 1372–1387, 2008.

[11] P. G. Wang and X. Liu, "ϕ_0-stability of hybrid impulsive dynamic systems on time scales," *Journal of Mathematical Analysis and Applications*, vol. 334, no. 2, pp. 1220–1231, 2007.

[12] J. Zhang, Y. Xia, P. Shi, and M. S. Mahmoud, "New results on stability and stabilisation of systems with interval time-varying delay," *IET Control Theory & Applications*, vol. 5, no. 3, pp. 429–436, 2011.

[13] P.-L. Liu, "State feedback stabilization of time-varying delay uncertain systems: a delay decomposition approach," *Linear Algebra and its Applications*, vol. 438, no. 5, pp. 2188–2209, 2013.

[14] L. V. Hien and V. N. Phat, "Exponential stability and stabilization of a class of uncertain linear time-delay systems," *Journal of the Franklin Institute*, vol. 346, no. 6, pp. 611–625, 2009.

[15] J. Cheng, H. Zhu, S. M. Zhong, and G. H. Li, "Novel delay-dependent robust stability criteria for neutral systems with mixed time-varying delays and nonlinear perturbations," *Applied Mathematics and Computation*, vol. 219, no. 14, pp. 7741–7753, 2013.

[16] M. V. Thuan, V. N. Phat, T. Fernando, and H. Trinh, "Exponential stabilization of time-varying delay systems with nonlinear perturbations," *IMA Journal of Mathematical Control and Information*, vol. 31, no. 4, pp. 441–464, 2014.

[17] Y. G. Sun, L. Wang, and G. M. Xie, "Exponential stability of switched systems with interval time-varying delay," *IET Control Theory & Applications*, vol. 3, no. 8, pp. 1033–1040, 2009.

[18] Y. G. Sun and J. Qi, "Note on exponential stability of certain nonlinear differential systems with time-varying delays," *Applied Mathematics Letters*, vol. 25, no. 12, pp. 2240–2245, 2012.

[19] K. Gu, "An integral inequality in the stability problem of time delay systems," in *Proceedings of the 39th IEEE Conference on Decision Control*, pp. 2805–2810, 2000.

[20] P. Park, J. W. Ko, and C. Jeong, "Reciprocally convex approach to stability of systems with time-varying delays," *Automatica*, vol. 47, no. 1, pp. 235–238, 2011.

[21] Y. J. Zhang, D. Yue, and E. G. Tian, "New stability criteria of neural networks with interval time-varying delay: a piecewise delay method," *Applied Mathematics and Computation*, vol. 208, no. 1, pp. 249–259, 2009.

[22] A. Seuret and F. Gouaisbaut, "Wirtinger-based integral inequality: application to time-delay systems," *Automatica*, vol. 49, no. 9, pp. 2860–2866, 2013.

[23] O. M. Kwon, J. H. Park, and S. M. Lee, "On robust stability criterion for dynamic systems with time-varying delays and nonlinear perturbations," *Applied Mathematics and Computation*, vol. 203, no. 2, pp. 937–942, 2008.

[24] Y. Liu, S. M. Lee, O. M. Kwon, and J. H. Park, "Delay-dependent exponential stability criteria for neutral systems with interval time-varying delays and nonlinear perturbations," *Journal of the Franklin Institute*, vol. 350, no. 10, pp. 3313–3327, 2013.

[25] K. Zhou and P. P. Khargonekar, "Robust stabilization of linear systems with norm-bounded time-varying uncertainty," *Systems & Control Letters*, vol. 10, no. 1, pp. 17–20, 1988.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 736720, 10 pages
http://dx.doi.org/10.1155/2015/736720

Research Article

Multivariable Fuzzy Control Based Mobile Robot Odor Source Localization via Semitensor Product

Ping Jiang,[1,2] Yuzhen Wang,[2] and Aidong Ge[2]

[1] *School of Electronic and Engineering, University of Jinan, Jinan, Shandong 250022, China*
[2] *School of Control Science and Engineering, Shandong University, Jinan, Shandong 250061, China*

Correspondence should be addressed to Ping Jiang; cse_jiangp@ujn.edu.cn

Received 8 April 2015; Accepted 18 June 2015

Academic Editor: Xinguang Zhang

In order to take full advantage of the multisensor information, a MIMO fuzzy control system based on semitensor product (STP) is set up for mobile robot odor source localization (OSL). Multisensor information, such as vision, olfaction, laser, wind speed, and direction, is the input of the fuzzy control system and the relative searching strategies, such as random searching (RS), nearest distance-based vision searching (NDVS), and odor source declaration (OSD), are the outputs. Fuzzy control rules with algebraic equations are given according to the multisensor information via STP. Any output can be updated in the proposed fuzzy control system and has no influence on the other searching strategies. The proposed MIMO fuzzy control scheme based on STP can reach the theoretical system of the mobile robot OSL. Experimental results show the efficiency of the proposed method.

1. Introduction

In natural world, many organisms such as drosophila, moth, and lobster use olfaction or/and vision cues to find the same species, avoid predators, exchange information, and search for food [1–3]. Inspired by those biological activities, in the early 1990s researchers started to build single or multiple mobile robots with onboard odor sensors or/and winds sensor to accomplish the odor source localization (OSL) task. Existing methods can be categorized along two lines. One is olfaction-based method, which mainly uses olfaction or/and wind information to search for gas sources without visual information. The other is vision-based method, which takes the visual information as an assistant of olfaction to accomplish the OSL task. Most work has been focused on the first field and it has become a mature and popular filed. However, the vision-based method is immature and needs to do deep study due to the late beginning. Russell [4], Meng and Li [5], Lilienthal et al. [6], Naeem et al. [7], Kowadlo and Russell [8], and Ishida et al. [9] have given relative reviews about mobile robot OSL from a different angle or application. The interested reader is referred to [4–9] for a comprehensive review of olfaction-based mobile robot OSL. Compared with organisms, robots can be deployed quickly, maintained at low cost, and work for a long time without fatigue. Moreover, they can enter the dangerous or harmful areas. Mobile robot OSL is a multidisciplinary research field with wide potential applications, such as judging toxic/harmful gas leakage location, checking contraband (e.g., heroin), locating unexploded mines and bombs, and fighting against terrorist attacks.

It is well known that human beings normally first look around to search for the most potential region or object and then identify whether the region or object is an odor source by olfaction. Vision contains abundant information, so visual sensor could be a good assistant of olfaction for mobile robot OSL. Meanwhile, large amount of leakage accidents indicate that some devices are more likely to leak, such as valves, bottles, and pipelines. In this paper, such devices are called potential gas sources and the areas which contain such devices are called plausible areas. It would improve the searching efficiency if these potential gas sources are recognized accurately and the plausible areas are determined rapidly in the searching process.

In recent years, a few researchers attempted to integrate vision and olfaction to localize the odor source. Kowadlo et al. [10] took crackles as the vision feature assisting olfaction to search for the odor source. Ishida et al. [11] proposed a color-based algorithm to deal with the vision information in

the searching process. These methods were verified in the experiments, which indicate that vision as an assistant of olfaction for mobile robot OSL is efficient. Inspired by these researches, Jiang et al. [12] proposed a support vector machine based algorithm to localize an odor source and the author also presented a top-down visual attention mechanism-based algorithm [13] for mobile robot OSL. And then least square estimation was used to fuse the vision and olfaction information to accomplish the OSL task in stable airflow environment [14]. Meanwhile, Jiang and Zhang [15] attempted to integrate the vision and olfaction using subsumption architecture to accomplish the OSL task. However, how to fuse the uncertainty, ambiguity, vagueness, incompleteness, and granularity of the multisensor information from the mobile robot, especially vision and olfaction information, needs further study from the deep analysis to those few reports. It is noteworthy that multisensor data fusion is developed in recent years and new fusion algorithms and models are constantly emerging such as Dempster-Shafer evidence theory, probability theory, fuzzy theory, possibility theory, rough set theory, and the improved algorithms of these methods [16, 17]. Meanwhile, these methods have been successfully used in many fields, such as image processing, fault diagnosis, and target tracking. Inspired by these successful cases, we attempt to set up a multivariable fuzzy control system based on semitensor product for mobile robot OSL by fusing multisensor information and obtain some interesting results.

Fuzzy control as an intelligent control strategy needs no precise mathematical model for the objective system. They have found a great variety of applications ranging from control engineering, qualitative modeling, pattern recognition, signal processing, machine intelligence, and so on [18, 19]. In particular, fuzzy logic control (FLC), as one of the earliest applications of fuzzy sets and systems, has become one of the most successful applications. In fact, FLC has been proved to be a successful control approach to many complex nonlinear systems or even nonanalytic ones. The fuzzy control algorithm consists of a set of heuristic control rules, and fuzzy sets and fuzzy logic are used, respectively, to represent linguistic terms and to evaluate the rules. Since then, fuzzy logic control has attracted great attention from both academic and industrial communities and a lot of excellent books and tutorial articles on the topic have been published. However, it is difficult to infer the proper control input for a multivariable system since the dimension of its relation matrix is very large. The high dimensionality of the relation matrix might lead to not only computational difficulties but also memory overload. To solve this problem, a fuzzy control algorithm by which the multivariable fuzzy system is decomposed into a set of one-dimensional systems [18, 19]. The decomposition of control rules is preferable since it alleviates the complexity of the problem.

Recently, the semitensor product (STP) of matrices was proposed in [20]. And up to now, it has been widely applied in many fields, such as boolean network [21, 22] and coloring problems [23]. The logic expression can be expressed into an algebraic form by constructing its structure matrix. In [22], the observed data was expressed into a two-valued algebraic form. For the mobile robot odor source localization different sizes of the multisensor information play the different roles in the searching process. Therefore, the multisensor information for the mobile robot odor source localization cannot be divided into two-valued true and false cases simply. This multisensor information is expressed as multivalued algebraic form according to the actual demand. It is noted that the fuzzy logic also can be considered as an extended mix-valued logic in which the truth-values are the ones of memberships of all the elements in a fuzzy set, and the complex reasoning process can be converted into a problem of solving a set of algebraic equations via STP, which greatly simplifies the process of logical reasoning.

In this paper, we attempt to set up a multi-input multi-output (MIMO) fuzzy control framework based on STP for the mobile robot OSL. The multisensor information obtained by the mobile robot is the inputs and the relative searching strategies are the outputs. Several interesting results are obtained. The main contributions of this paper are as follows:

(1) A MIMO fuzzy control system is set up for the mobile robot OSL.

(2) Fuzzy control rules with algebraic equations are given according to the multisensor information.

(3) Any output can be updated in this framework and has no influence to the others.

(4) The proposed method based on MIMO fuzzy control scheme via STP for mobile robot OSL can reach the theory of this field.

The rest of this paper is organized as follows. Section 2 provides some necessary preliminaries on the semitensor product of matrices and the expression of logical function and logical variables. Section 3 presents the proposed algorithms for mobile robot OSL. Section 4 shows experimental results and analysis and the conclusion is given in Section 5.

2. Matrices with Logical Variables

First, some notations are introduced, which will be used in this paper:

(i) δ_k^i: the ith column of the identity matrix I_k.

(ii) $\Delta_k := \{\delta_k^i \mid i = 1, 2, \ldots, k\}$; especially, $\Delta := \Delta_2$.

(iii) $\mathscr{D} := \{1, 0\}$; to use matrix expression, "1" and "0" can be expressed with the following vectors, respectively: $1 \sim \delta_2^1$, $0 \sim \delta_2^2$.

(iv) $\mathscr{D}_k := \{1, (k-2)/(k-1), (k-3)/(k-1), \ldots, 0\}$, $k \geq 2$.

(v) A matrix $L \in \mathbb{R}^{m\times n}$ is called a logical matrix if the columns of L, denoted by $\mathrm{Col}(L)$, are of the form δ_n^k; that is, $\mathrm{Col}(L) \subset \Delta_n$.

(vi) Let $\mathscr{L}_{n\times r}$ denote the set of $n \times r$ logical matrices; if $L \in \mathscr{L}_{n\times r}$, by definition, it can be expressed as $L = [\delta_n^{i_1}\ \delta_n^{i_2}\ \cdots\ \delta_n^{i_r}]$; for the sake of compactness, it is briefly denoted as $L = \delta_n[i_1\ i_2\ \cdots\ i_r]$.

(vii) Each k-valued logical value with a vector can be denoted as $(k-i)/(k-1) \sim \delta_k^i$, $i = 1, 2, \ldots, k$; then, $\mathscr{D}_k \sim \Delta_k$.

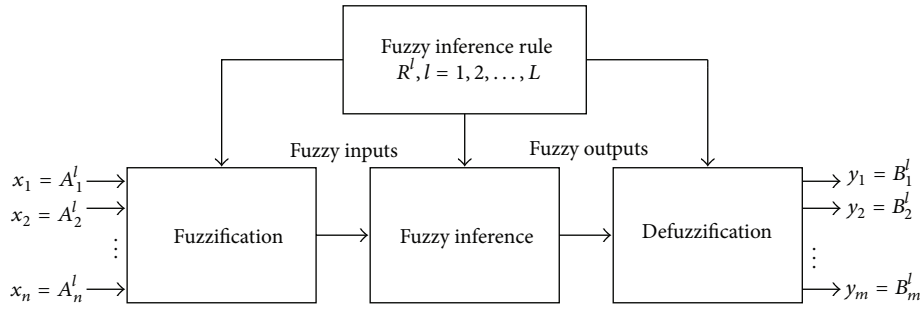


Figure 1: The fuzzy control scheme.

In the following, we recall some definitions and basic properties about the STP [20].

Definition 1. Let $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{p\times q}$. Let $s = \mathrm{lcm}(n, p)$ denote the least common multiple of n and p. Then, the semitensor product of A and B is defined as

$$A \ltimes B = \left(A \otimes I_{s/n}\right)\left(B \otimes I_{s/p}\right), \tag{1}$$

where "$\otimes$" is the Kronecker product.

Remark 2. It is noted that when $n = p$, the STP of A and B becomes the conventional matrix product. Hence, the STP is a generalization of the conventional matrix product. Because of this, we can omit the sign "$\ltimes$" without confusion.

Definition 3. A swap matrix $W_{[m,n]}$ is an $mn \times mn$ matrix. Its rows and columns are labeled by double index (i, j), the columns are arranged by the ordered multi-index $\mathrm{Id}(i, j; m, n)$, and the rows are arranged by the ordered multi-index $\mathrm{Id}(J, I; n, m)$. Then the elements at position $[(I, J), (i, j)]$ are

$$w_{(I,J)(i,j)} = \delta_{i,j}^{I,J} = \begin{cases} 1, & I = i,\ J = j, \\ 0, & \text{others}. \end{cases} \tag{2}$$

Remark 4. Let $X \in \mathbb{R}^n$ and $Y \in \mathbb{R}^m$ be column vectors; then $W_{[n,m]}XY = YX$. Let $X_i \in \mathbb{R}^{n_i}$, $i = 1, 2, \ldots, k$, be column vectors; then $[I_{n_1 n_2 \cdots n_{t-1}} \otimes W_{[n_t, n_{t+1}]}]X_1 \cdots X_{t-1}X_tX_{t+1} \cdots X_k = X_1 \cdots X_{t-1}X_{t+1}X_t \cdots X_k$.

Let $x_i \in \mathscr{D}_{k_i}$, $i = 1, \ldots, n$ and $y_j \in \mathscr{D}_{s_j}$, $j = 1, \ldots, m$. Assume that a logic mapping,

$$F : \mathscr{D}_{k_1} \times \cdots \times \mathscr{D}_{k_n} \longrightarrow \mathscr{D}_{s_1} \times \cdots \times \mathscr{D}_{s_m}, \tag{3}$$

can be expressed as

$$\begin{aligned} y_1 &= f_1(x_1, x_2, \ldots, x_n), \\ y_2 &= f_2(x_1, x_2, \ldots, x_n), \\ &\vdots \\ y_m &= f_m(x_1, x_2, \ldots, x_n), \end{aligned} \tag{4}$$

where $f_j : \mathscr{D}_{k_1} \times \cdots \times \mathscr{D}_{k_n} \to \mathscr{D}_{s_j}$, $j = 1, \ldots, m$.

Lemma 5. *Any logical function $y = F(x_1, x_2, \ldots, x_n)$ can be uniquely expressed into the multilinear form of*

$$y = F(x_1, x_2, \ldots, x_n) = M_f \ltimes_i^n x_i, \tag{5}$$

where $M_f \in \mathscr{L}_{s\times k}$ is called the structural matrix of F, $y \in \Delta_s$, $s = s_1 s_2 \cdots s_m$, and $k = k_1 k_2 \cdots k_n$.

Lemma 6. *Consider (5). For the sake of compactness, we denote $M_f W_{[k_i, \prod_{p=1}^{i-1} k_p]} = M$. For any $1 \le i \le n$, we split M into k_i equal-size blocks as $[Blk_1(M), \ldots, Blk_{k_i}(M)]$. If all the blocks are the same, then x_i is a redundant variable. Thus, y can be replaced by*

$$y = M_f' x_1 \cdots x_{i-1} x_{i+1} \cdots x_n, \tag{6}$$

where $M_f' = Blk_1(M) = M_f W_{[k_i, \prod_{p=1}^{i-1} k_p]} \delta_{k_i}^1$.

3. Multivariable FLC Based on STP for Mobile Robot OSL

Consider the linguistic control rules of the multivariable fuzzy system:

$$\begin{aligned} &R^l\text{: IF } x_1 \text{ is } A_1^l, \ldots, \text{ and } x_n \text{ is } A_n^l, \\ &\text{THEN } y_1 \text{ is } B_1^l, \ldots, \text{ and } y_m \text{ is } B_m^l, \end{aligned} \tag{7}$$

where x_i and y_j are linguistic variables representing the process state and the control variable, respectively. R^l denotes the lth fuzzy inference rule, where $l \in \{1, \ldots, L\}$, and L is the number of fuzzy rules. $A_i, i = 1, \ldots, n$, and B_j, $j = 1, \ldots, m$, are the normalized fuzzy set of linguistic values on universes of discourses X_i and Y_j, respectively. The control system is shown in Figure 1.

3.1. Controller Design of MFS with Complete Fuzzy Control Rules. The fuzzy control rules are in accordance with consistency and correctness. For the n inputs and m outputs fuzzy

controller (7), let the number of the linguistic values of x_i and y_j be k_i and s_j, respectively; that is,

$$\begin{aligned} &x_i \in \mathscr{D}_{k_i}, \\ &A_i = \left\{A_i^1, \ldots, A_i^{k_i}\right\}, \\ &\qquad\qquad\qquad\qquad i = 1, \ldots, n, \\ &y_j \in \mathscr{D}_{s_j}, \\ &B_j = \left\{B_j^1, \ldots, B_j^{s_j}\right\}, \\ &\qquad\qquad\qquad\qquad j = 1, \ldots, m. \end{aligned} \tag{8}$$

We identify $A_1^{i_1} \sim \delta_{k_1}^{i_1}; \ldots; A_n^{i_n} \sim \delta_{k_n}^{i_n}$, $B_1^{j_1} \sim \delta_{s_1}^{j_1}; \ldots; B_m^{j_m} \sim \delta_{s_m}^{j_m}$, $i_1 = 1, \ldots, k_1; \ldots; i_n = 1, \ldots, k_n$; $j_1 = 1, \ldots, s_1; \ldots; j_m = 1, \ldots, s_m$.

Then, (7) can be written as

$$\begin{aligned} R^l\text{: IF } x_1 = \delta_{k_1}^{i_1}, \ldots, \text{ and } x_n = \delta_{k_n}^{i_n}, \\ \text{THEN } y_1 = \delta_{s_1}^{j_1}, \ldots, \text{ and } y_m = \delta_{s_m}^{j_m}. \end{aligned} \tag{9}$$

Using the vector form of logical variables, we express the fuzzy controller as

$$\begin{aligned} y_1 &= M_1 x, \\ y_2 &= M_2 x, \\ &\vdots \\ y_m &= M_m x, \end{aligned} \tag{10}$$

$$y = M_f x, \tag{11}$$

where $y := \ltimes_{j=1}^m y_j$, $x := \ltimes_{i=1}^n x_i$, $M_j \in \mathscr{L}_{s_j \times k}$, $j = 1, \ldots, m$, and $\mathrm{Col}_i(M_f) = \mathrm{Col}_i(M_1) \ltimes \cdots \ltimes \mathrm{Col}_i(M_m)$, where $\mathrm{Col}_i(M_f)$ denotes the ith column of matrix M_f. For rules l and $y_j = M_j x$, since $x = \ltimes_{i=1}^n x_i = \delta_{k_1}^{i_1} \ltimes \cdots \ltimes \delta_{k_n}^{i_n} = \delta_k^i$, $y_j = \delta_{s_j}^{j_j}$, we have $\mathrm{Col}_i(M_j) = \delta_{s_j}^{j_j}$. If the fuzzy rules are complete, all the columns of M_j, $j = 1, \ldots, m$ can be obtained. Then, we have the following result.

Theorem 7. *The structural matrices M_j, $j = 1, \ldots, m$ and M_f of the fuzzy controller can be uniquely determined, if and only if the fuzzy rules of the fuzzy controller are complete.*

Proof. Consider the following.

Sufficiency. For the fuzzy rules (9), let $x = x_1 \ltimes \cdots \ltimes x_n$. Assume the fuzzy rules of the fuzzy controller are complete; that is, there are k fuzzy rules. For the lth, $l = 1, \ldots, k$, fuzzy rule, we have $x = \delta_k^i$ and $y_1 = \delta_{s_1}^{j_1}$. Then the ith column of M_1 can be obtained as

$$\mathrm{Col}_i\left(M_1\right) = \delta_{s_1}^{j_1}. \tag{12}$$

Repeating this procedure, one can obtain all the columns of M_1 if the fuzzy rules are complete. Similarly, all $M_2, \ldots, M_m$ and M_f can be determined.

Necessity. If the structural matrices M_j and M_f of the fuzzy controller are uniquely determined, then all the columns of M_j and M_f are uniquely determined. Because one column of M_f can generate one fuzzy rule, one can obtain k fuzzy rules from k columns of M_j or M_f; that is, the fuzzy rules are complete. □

Remark 8. If the rules are not complete, some columns of M_j and M_f can be determined. In this case, the model is not unique. In addition, uncertain columns of M_j and M_f can be chosen arbitrarily.

3.2. Controller Design of MFS with Incomplete Fuzzy Control Rules. The fuzzy control rules are also in accordance with consistency and correctness. We first define a kind of incidence matrix to express the dynamic connection of the inputs and the outputs for a fuzzy controller.

Definition 9. Consider a fuzzy controller with m controls and n input variables. An $m \times n$ matrix, $\mathscr{J} = (r_{j,i}) \in \mathscr{R}^{m \times n}$, is called its incidence matrix, if

$$r_{j,i} = \begin{cases} 1, & y_j \text{ depends on } x_i, \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

Consider fuzzy controllers (9) and (10); the indegree $d(y_j)$ is the number of the inputs and it influences y_j directly. From the incidence matrix of the fuzzy controller, we have

$$d\left(y_j\right) = \sum_{k=1}^{n} r_{jk}, \quad j = 1, \ldots, m. \tag{14}$$

A set of controls (10) is said to be a feasible one to (9), if (9) satisfies (10). A feasible set of controls (10) with the indegree $d^*(y_j)$, $j = 1, \ldots, m$, is called a least indegree feasible set, if for any other realization with indegree $d(y_j)$, $j = 1, \ldots, m$, we have

$$d^*\left(y_j\right) \le d\left(y_j\right), \quad j = 1, \ldots, m. \tag{15}$$

We can use Lemma 6 to remove redundant variables and obtain a least indegree feasible set when the fuzzy rules are complete.

Assume a set of incomplete rules as

$$\begin{aligned} R^l\text{: IF } x_1 \text{ is } A_1^l, \ldots, \text{ and } x_n \text{ is } A_n^l, \\ \text{THEN } y_1 \text{ is } B_1^l, \ldots, \text{ and } y_m \text{ is } B_m^l, \\ l \in \{1, \ldots, t\}, \end{aligned} \tag{16}$$

where R^l denotes the lth fuzzy control rule, t is the number of the control rules, and $t < k$.

Consider the controls $y_j = M_j x$. Using this set of fuzzy rules, some columns of the structural matrix M_j can be determined. For instance

$$\begin{aligned} &M_j \\ &= \delta_{s_j}\left[\star \ \cdots \ \star \ c_{j_1} \ \star \ \cdots \ \star \ \cdots \ \star \ c_{j_k} \ \star \ \cdots \ \star\right], \end{aligned} \tag{17}$$

where "$\star$" stands for the uncertain columns. Equation (17) is called the uncertain structural matrix. Let

$$M_{j,i} := M_j W_{[k_i, \prod_{p=1}^{i-1} k_p]}, \quad i = 1, \ldots, n. \tag{18}$$

Then split it into k_i equal blocks as

$$M_{j,i} := \left[M_{j,i}^1 \ \ M_{j,i}^2 \ \ \cdots \ \ M_{j,i}^{k_i}\right]. \tag{19}$$

According to Lemma 6, we have the following result.

Proposition 10. *The fuzzy control y_j has an algebraic form which is independent of x_i, if and only if*

$$M_{j,i}^1 = M_{j,i}^2 = \cdots = M_{j,i}^{k_i} \tag{20}$$

has a solution for uncertain elements.

Proof. Consider the following.

Sufficiency. Assume that (20) holds. By Lemma 5, the fuzzy control y_j has an algebraic form which is independent of x_i.

Necessity. Assume the fuzzy control y_j is independent of x_i; then y_j remains invariant whenever $x_i = \delta_{k_i}^q$, $q = 1, \ldots, k_i$. Thus

$$M_j W_{[k_i, \prod_{p=1}^{i-1} k_p]} \delta_{k_i}^1 = \cdots = M_j W_{[k_i, \prod_{p=1}^{i-1} k_p]} \delta_{k_i}^{k_i}, \tag{21}$$

which implies that (20) holds. The proof is completed. □

Example 11. Consider a fuzzy controller, which has 4 inputs, $x_1, x_3 \in \mathscr{D}_2$, $x_2, x_4 \in \mathscr{D}_3$, and 2 outputs (controls), $y_1 \in \mathscr{D}_3$ and $y_2 \in \mathscr{D}_4$.

In the vector form, we assume that there are a set of control rules as follows:

IF $x_1 = \delta_2^1$, $x_2 = \delta_3^1$, $x_3 = \delta_3^1$ and $x_4 = \delta_2^1$, THEN $y_1 = \delta_3^2$ and $y_2 = \delta_4^2$.

IF $x_1 = \delta_2^1$, $x_2 = \delta_3^2$, $x_3 = \delta_2^2$ and $x_4 = \delta_3^1$, THEN $y_1 = \delta_3^2$ and $y_2 = \delta_4^4$.

IF $x_1 = \delta_2^1$, $x_2 = \delta_3^3$, $x_3 = \delta_2^2$ and $x_4 = \delta_3^3$, THEN $y_1 = \delta_3^1$ and $y_2 = \delta_4^1$.

IF $x_1 = \delta_2^2$, $x_2 = \delta_3^1$, $x_3 = \delta_2^1$ and $x_4 = \delta_3^1$, THEN $y_1 = \delta_3^1$ and $y_2 = \delta_4^3$.

IF $x_1 = \delta_2^2$, $x_2 = \delta_3^2$, $x_3 = \delta_2^1$ and $x_4 = \delta_3^3$, THEN $y_1 = \delta_3^3$ and $y_2 = \delta_4^2$.

IF $x_1 = \delta_2^2$, $x_2 = \delta_3^3$, $x_3 = \delta_2^2$ and $x_4 = \delta_3^2$, THEN $y_1 = \delta_3^3$ and $y_2 = \delta_4^4$.

Now, we would like to get a least indegree feasible set of controls. Some columns of M_1 and M_2 can be identified as

$$\begin{aligned} M_1 &= \delta_3\left[2 \ \star \ \star \ \star \ \star \ \star \ \star \ \star \ \star \ 2 \ \star \ \star \ \star \ \star \ \star \ \star \ \star \ 1 \ \star \ 1 \ \star \ \star \ \star \ \star \ \star \ \star \ 3 \ \star \ \star \ \star \ \star \ \star \ \star \ \star \ 3 \ \star\right], \\ M_2 &= \delta_4\left[2 \ \star \ \star \ \star \ \star \ \star \ \star \ \star \ \star \ 4 \ \star \ \star \ \star \ \star \ \star \ \star \ \star \ 1 \ \star \ 3 \ \star \ \star \ \star \ \star \ \star \ \star \ 2 \ \star \ \star \ \star \ \star \ \star \ \star \ \star \ 4 \ \star\right], \end{aligned} \tag{22}$$

where "$\star$" denotes the unknown element. Now, we check whether x_1 can be a redundant variable of y_1. Split M_1 into two equal blocks as $M_1 = [M_1^1 \ \ M_1^2]$, and let $M_1^1 = M_1^2$ which yields the solution as

$$\begin{aligned} &M_1^1 = M_1^2 \\ &= \delta_3\left[2 \ 1 \ \star \ \star \ \star \ \star \ \vdots \ \star \ \star \ 3 \ 2 \ \star \ \star \ \vdots \ \star \ \star \ \star \ \star \ 3 \ 1\right]. \end{aligned} \tag{23}$$

Thus, the control can be simplified as $y_1 = M_1^1 x_2 x_3 x_4$. Now, we check x_2. Splitting M_1^1 into three equal blocks as $[M_1^{11} \ \ M_1^{12} \ \ M_1^{13}]$ and letting $M_1^{11} = M_1^{12} = M_1^{13}$, it can be updated as

$$M_1^{11} = M_1^{12} = M_1^{13} = \delta_3\left[2 \ 1 \ 3 \ 2 \ 3 \ 1\right], \tag{24}$$

which satisfies $y_1 = M_1^{11} x_3 x_4$. Next, check x_3 and x_4. Since $M_1^{11} W_{[2,3]} = \delta_3[2 \ 2 \ 1 \ 3 \ 3 \ 1]$, x_3 and x_4 are not fabricated variables. Finally, we obtain $y_1 = \delta_3[2 \ 2 \ 1 \ 3 \ 3 \ 1] x_3 x_4$. Similarly, we have $y_2 = \delta_4[2 \ 4 \ 3 \ 4 \ 2 \ 1] x_3 x_4$.

And then, the least indegree realization can finally be obtained as

$$\begin{aligned} y_1 &= \delta_3\left[2 \ 2 \ 1 \ 3 \ 3 \ 1\right] x_3 x_4, \\ y_2 &= \delta_4\left[2 \ 4 \ 3 \ 4 \ 2 \ 1\right] x_3 x_4. \end{aligned} \tag{25}$$

3.3. Controller Design of MFS for Mobile Robot OSL. A great deal of sensor information needs to be processed rapidly for a mobile robot during the real-time searching process,

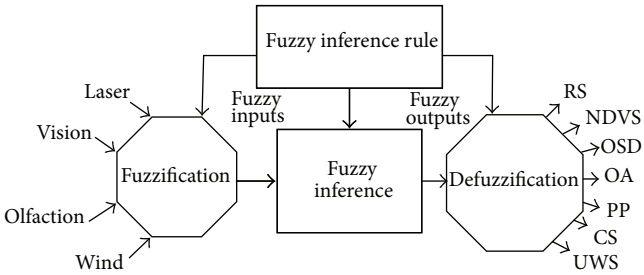


Figure 2: Fuzzy control based system for mobile robot OSL.

such as gas sensor (olfaction), camera (vision), wind sensor (wind speed and direction), laser sensor (distance), and electronic compass (position of robot). The mobile robot needs to make correct determination when different sensor information is required. In this paper, a MIMO fuzzy control based localization framework (shown in Figure 2) is set up in order to make full use of the diversity and complementary of multisensor information and obtain more detailed and accurate decision. The inputs are the multisensor information or the computed results of the sensor information. Here, the laser sensor information (LSI) is represented by the linguistic terms "near" and "far," vision information (VI) is "true" and "false," olfaction information (OI) is "too low," "normal," and "too high," and wind information (WI) is "true" and "false." And the outputs are several behaviors. In this paper, six behaviors are set up, including obstacle avoiding (OA), odor source declaration (OSD), nearest distance-based visual searching (NDVS), up-wind searching (UWS), path planning (PP), chemotaxis searching (CS), and random searching (RS).

We identify the following:

Inputs: LSI $\sim x_1$, OI $\sim x_2$, VI $\sim x_3$, WI $\sim x_4$.

Outputs: OA $\sim \delta_7^1$, GSD $\sim \delta_7^2$, NDVS $\sim \delta_7^3$, UWS $\sim \delta_7^4$, PP $\sim \delta_7^5$, CS $\sim \delta_7^6$, RS $\sim \delta_7^7$:

$$\begin{aligned} \text{near} &\sim \delta_2^1, \\ \text{far} &\sim \delta_2^2, \\ \text{true} &\sim \delta_2^1, \\ \text{false} &\sim \delta_2^2, \\ \text{too low} &\sim \delta_3^1, \\ \text{normal} &\sim \delta_3^2, \\ \text{too high} &\sim \delta_3^3. \end{aligned} \tag{26}$$

Then, the fuzzy rules can be expressed into the following form.

IF $x_1 = \delta_2^1$, $x_2 = \delta_3^1$, $x_3 = \delta_2^1$, and $x_4 = \delta_2^1$, THEN $y = \delta_7^1$; IF $x_1 = \delta_2^1$, $x_2 = \delta_3^1$, $x_3 = \delta_2^1$, and $x_4 = \delta_2^2$, THEN $y = \delta_7^1$; IF $x_1 = \delta_2^1$, $x_2 = \delta_3^1$, $x_3 = \delta_2^2$, and $x_4 = \delta_2^1$, THEN $y = \delta_7^1$; IF $x_1 = \delta_2^1$, $x_2 = \delta_3^1$, $x_3 = \delta_2^2$, and $x_4 = \delta_2^2$, THEN $y = \delta_7^1$; IF $x_1 = \delta_2^1$, $x_2 = \delta_3^2$, $x_3 = \delta_2^1$, and $x_4 = \delta_2^1$, THEN $y = \delta_7^1$; IF $x_1 = \delta_2^1$, $x_2 = \delta_3^2$, $x_3 = \delta_2^1$, and $x_4 = \delta_2^2$, THEN $y = \delta_7^1$; IF $x_1 = \delta_2^1$, $x_2 = \delta_3^2$, $x_3 = \delta_2^2$, and $x_4 = \delta_2^2$, THEN $y = \delta_7^1$; IF $x_1 = \delta_2^1$, $x_2 = \delta_3^2$, $x_3 = \delta_2^2$, and $x_4 = \delta_2^1$, THEN $y = \delta_7^1$; IF $x_1 = \delta_2^1$, $x_2 = \delta_3^3$, $x_3 = \delta_2^1$, and $x_4 = \delta_2^1$, THEN $y = \delta_7^2$; IF $x_1 = \delta_2^1$, $x_2 = \delta_3^3$, $x_3 = \delta_2^1$, and $x_4 = \delta_2^2$, THEN $y = \delta_7^2$; IF $x_1 = \delta_2^1$, $x_2 = \delta_3^3$, $x_3 = \delta_2^2$, and $x_4 = \delta_2^1$, THEN $y = \delta_7^1$; IF $x_1 = \delta_2^1$, $x_2 = \delta_3^3$, $x_3 = \delta_2^2$, and $x_4 = \delta_2^2$, THEN $y = \delta_7^1$; IF $x_1 = \delta_2^2$, $x_2 = \delta_3^1$, $x_3 = \delta_2^1$, and $x_4 = \delta_2^1$, THEN $y = \delta_7^3$; IF $x_1 = \delta_2^2$, $x_2 = \delta_3^1$, $x_3 = \delta_2^1$, and $x_4 = \delta_2^2$, THEN $y = \delta_7^3$; IF $x_1 = \delta_2^2$, $x_2 = \delta_3^1$, $x_3 = \delta_2^2$, and $x_4 = \delta_2^2$, THEN $y = \delta_7^7$; IF $x_1 = \delta_2^2$, $x_2 = \delta_3^1$, $x_3 = \delta_2^2$, and $x_4 = \delta_2^1$, THEN $y = \delta_7^7$; IF $x_1 = \delta_2^2$, $x_2 = \delta_3^2$, $x_3 = \delta_2^1$, and $x_4 = \delta_2^1$, THEN $y = \delta_7^5$; IF $x_1 = \delta_2^2$, $x_2 = \delta_3^2$, $x_3 = \delta_2^1$, and $x_4 = \delta_2^2$, THEN $y = \delta_7^5$; IF $x_1 = \delta_2^2$, $x_2 = \delta_3^2$, $x_3 = \delta_2^2$, and $x_4 = \delta_2^1$, THEN $y = \delta_7^4$; IF $x_1 = \delta_2^2$, $x_2 = \delta_3^2$, $x_3 = \delta_2^2$, and $x_4 = \delta_2^2$, THEN $y = \delta_7^6$; IF $x_1 = \delta_2^2$, $x_2 = \delta_3^3$, $x_3 = \delta_2^1$, and $x_4 = \delta_2^1$, THEN $y = \delta_7^5$; IF $x_1 = \delta_2^2$, $x_2 = \delta_3^3$, $x_3 = \delta_2^1$, and $x_4 = \delta_2^2$, THEN $y = \delta_7^5$; IF $x_1 = \delta_2^2$, $x_2 = \delta_3^3$, $x_3 = \delta_2^2$, and $x_4 = \delta_2^1$, THEN $y = \delta_7^4$; IF $x_1 = \delta_2^2$, $x_2 = \delta_3^3$, $x_3 = \delta_2^2$, and $x_4 = \delta_2^2$, THEN $y = \delta_7^6$; from the above form of the fuzzy rules, we can obtain the structure matrix:

$$M_f = \delta_7 \left[1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 2 \;\; 2 \;\; 1 \;\; 1 \;\; 3 \;\; 3 \;\; 7 \;\; 7 \;\; 5 \;\; 5 \;\; 4 \;\; 6 \;\; 5 \;\; 5 \;\; 4 \;\; 6\right]; \tag{27}$$

then $y = M_f x_1 x_2 x_3 x_4$.

Now using Lemma 5, we check whether x_1, x_2, x_3, or x_4 is a redundant variable of y. It is easy to verify that

$$\begin{aligned} M_f &= \delta_7 \left[1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 2 \;\; 2 \;\; 1 \;\; 1 \;\; \vdots \;\; 3 \;\; 3 \;\; 7 \;\; 7 \;\; 5 \;\; 5 \;\; 4 \;\; 6 \;\; 5 \;\; 5 \;\; 4 \;\; 6\right], \\ M_f W_{[3,2]} &= \delta_7 \left[1 \;\; 1 \;\; 1 \;\; 1 \;\; 3 \;\; 3 \;\; 7 \;\; 7 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 5 \;\; 5 \;\; 4 \;\; 6 \;\; 2 \;\; 2 \;\; 1 \;\; 1 \;\; 5 \;\; 5 \;\; 4 \;\; 6\right], \\ M_f W_{[2,6]} &= \delta_7 \left[1 \;\; 1 \;\; 1 \;\; 1 \;\; 2 \;\; 2 \;\; 3 \;\; 3 \;\; 5 \;\; 5 \;\; 5 \;\; 5 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 7 \;\; 7 \;\; 4 \;\; 6 \;\; 4 \;\; 6\right], \\ M_f W_{[2,12]} &= \delta_7 \left[1 \;\; 1 \;\; 1 \;\; 1 \;\; 2 \;\; 1 \;\; 3 \;\; 7 \;\; 5 \;\; 4 \;\; 5 \;\; 4 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 2 \;\; 1 \;\; 3 \;\; 7 \;\; 5 \;\; 6 \;\; 5 \;\; 6\right]. \end{aligned} \tag{28}$$

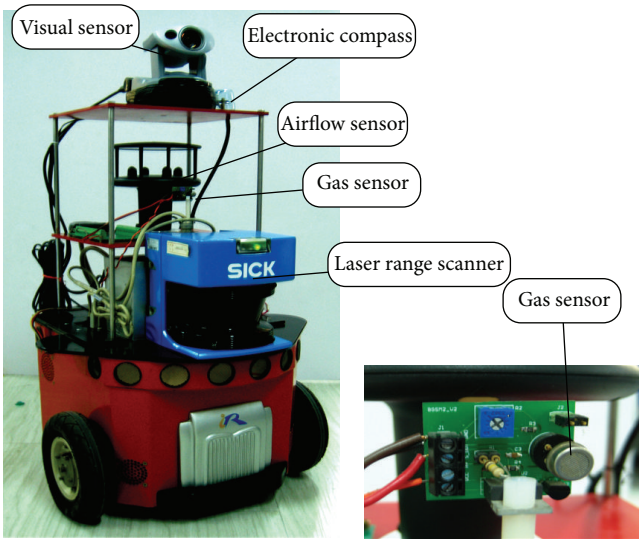


Figure 3: Experimental platform of mobile robot OSL.

Obviously, we know x_1, x_2, x_3, and x_4 are not redundant variables of y.

Assume that LSI is "near," OI is "too high," VI is "true," and WI is "true" or "false"; we have $y = M_f x_1 x_2 x_3 x_4 = M_f \delta_2^1 \delta_3^3 \delta_2^1 \delta_2^1 = \delta_7^2$ or $y = M_f x_1 x_2 x_3 x_4 = M_f \delta_2^1 \delta_3^3 \delta_2^1 \delta_2^2 = \delta_7^2$, which means GSD.

4. Experimental Results and Analysis

The proposed method is verified using real robot experiments. The mobile robot platform and the odor source are shown in Figure 3. A PTZ camera (EVI-D100P, Sony), a gas sensor (MiCS-5135, e2v Technologies (UK) Ltd.), an anemometer (WindSonic, Gill), a laser rangefinder (LMS200, Sick AG), and an electronic compass were mounded on the robot. The PTZ is 1.3 meters high from the ground. The size of each sampled image is 320240 pixels. The computer (CPU: 3.0 GHz, RAM: 1.0 GBytes) is used in this paper.

4.1. The Experimental Result with No Vision and Olfaction. The mobile robot searches the whole scene to find the odor plume using random searching (RS) methods when there is no vision and olfaction information.

LSI is "far," OI is "too low," VI is "false," and WI is "true" or "false"; we have

$$\begin{aligned} y &= M_f x_1 x_2 x_3 x_4 = M_f \delta_2^2 \delta_3^1 \delta_2^2 \delta_2^1 = \delta_7^7, \\ \text{or } y &= M_f x_1 x_2 x_3 x_4 = M_f \delta_2^2 \delta_3^1 \delta_2^2 \delta_2^2 = \delta_7^7, \end{aligned} \tag{29}$$

which means RS. Figure 4 shows the searching trajectory. The robot starts spiral surge with a certain radius (the radius is 377 mm in this paper) from the initial position (the black solid dot in Figure 4). The blue dots is the moving trajectory.

4.2. The Experimental Result with Vision. Traditionally, random searching methods are used for plume finding when there is no olfaction. However, these methods have the same hypothesis that the probabilities of the gas leakage source appearing in the scene are equal, which is obviously inconsistent with the actual situation. Because the probabilities of gas leakage source in some areas is big and others may be small, thus, these random searching methods have certain blindness. If some potential gas sources are determined using vision in advance and then drives the robot to check the relative plausible areas firstly, it would overcome the blindness of random searching efficiently in a certain degree.

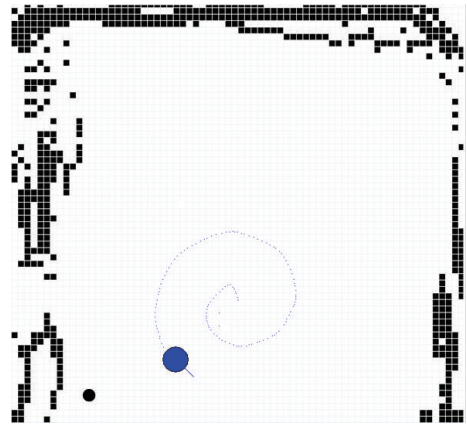


Figure 4: Random searching trajectory of mobile robot OSL.

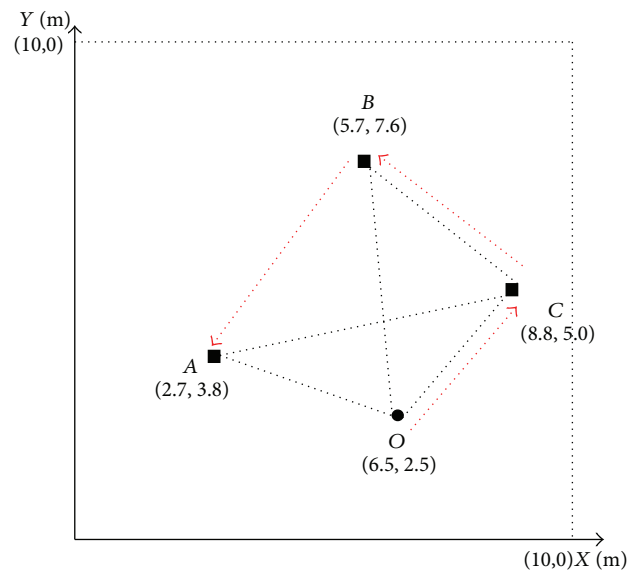
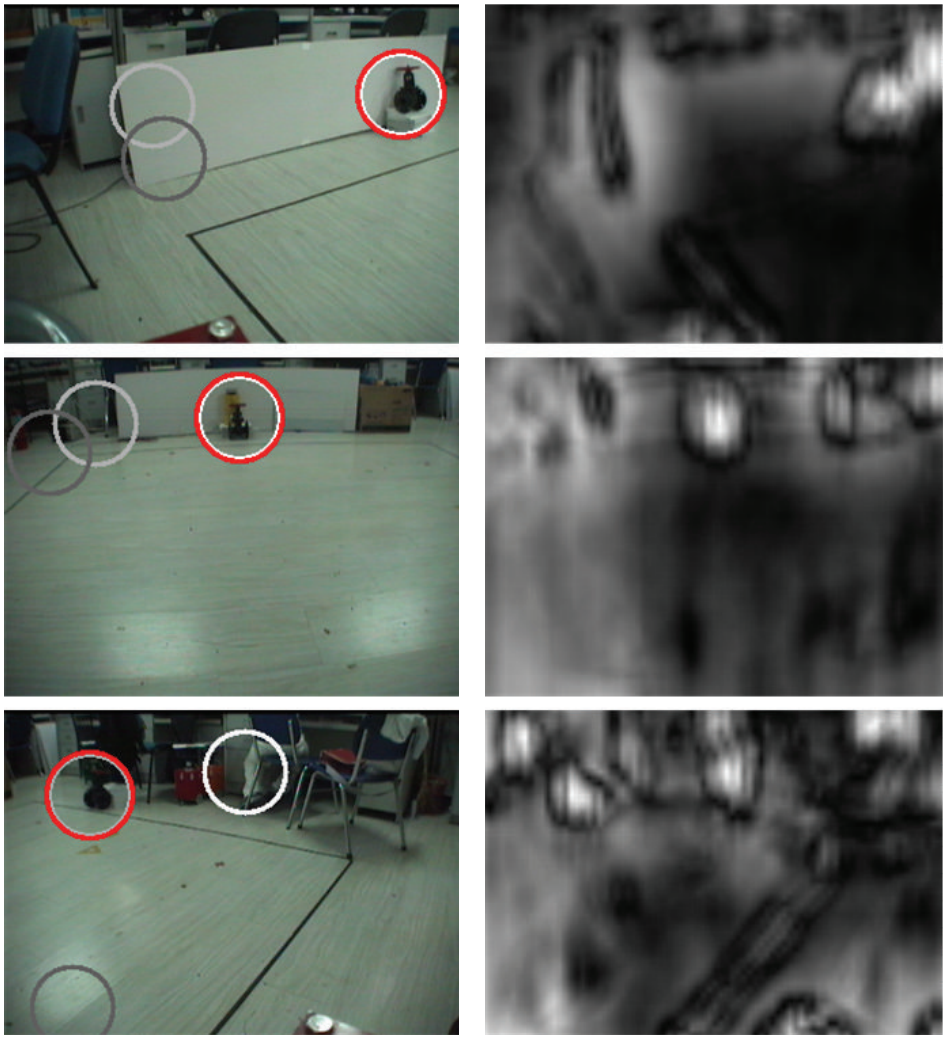


Figure 5: NDVS strategy of mobile robot OSL.

LSI is "far," OI is "too low," VI is "true," and WI is "true" or "false"; we have

$$\begin{aligned} y &= M_f x_1 x_2 x_3 x_4 = M_f \delta_2^2 \delta_3^1 \delta_2^1 \delta_2^1 = \delta_7^3, \\ \text{or } y &= M_f x_1 x_2 x_3 x_4 = M_f \delta_2^2 \delta_3^1 \delta_2^1 \delta_2^2 = \delta_7^3, \end{aligned} \tag{30}$$

which means NDVS. The optimal strategy is shown in Figure 5. If only one plausible area is existent in the scene, the robot moves to the area directly to check. If more



(a) (b)

Figure 6: Scene images and the relative saliency maps.

plausible areas are existent it needs to plan the searching path to improve the searching efficiency. A recursive optimal searching strategy (NDVS: nearest distance based visual searching) is proposed in this paper because it cannot be determined in advance which one will find the gas source.

In Figure 4, O is the initial position (6.5, 2.5) and A (2.7, 3.8), B (5.7, 7.6), and C (8.8, 5.0) are the plausible areas obtained using top-down visual attention mechanism and shape analysis [13] to the vision information. The distances between the initial position and the plausible areas are 4.02 m, 5.16 m, and 3.40 m, respectively. Thus, the robot moves to the nearest area (point C). If there is no gas source, the next target from A and B is selected according to the distance to C. L_{CB} (2.52 m) is less than L_{CA} (6.22 m). Thus, B is the next.

Figure 6(a) is the scene images in which the three white to gray circles represent the visual computing results (the most three saliency regions) and the red circle represents the potential gas source determined by using shape analysis. Figure 6(b) is the relative saliency map.

In Figure 7, point O (red solid dot) is the initial position of the robot, A, B, and C are the plausible areas, the big blue dot represents the robot, and the blue dots are the searching trajectory.

4.3. The Experimental Result with Vision and Olfaction. When vision, olfaction, and wind information are efficient the robot starts to make decision where to go, that is, path planning (PP). LSI is "far," OI is "normal," VI is "true," and WI is "true" or "false"; we have $y = M_f x_1 x_2 x_3 x_4 = M_f \delta_2^2 \delta_3^2 \delta_2^1 \delta_2^1 = \delta_7^5$ or $y = M_f x_1 x_2 x_3 x_4 = M_f \delta_2^2 \delta_3^2 \delta_2^1 \delta_2^2 = \delta_7^5$ which means PP. The searching result is shown in Figure 8.

The red dot line is the trajectory of the mobile robot and the big red round is the start position. At the beginning there is no gas concentration, the robot moves toward point A (points A, B, and C are the plausible areas) by using NDVS method. And in the moving process gas concentration is detected; then the robot adjusts the moving direction constantly according to the gas concentration and wind information. When both laser information and vision are

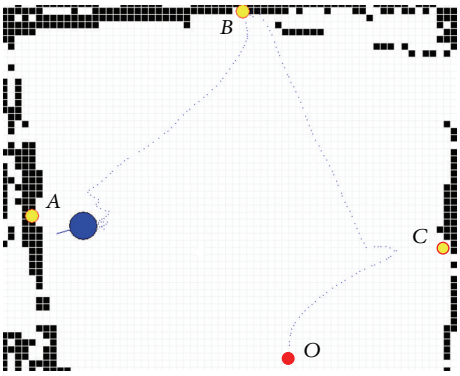


Figure 7: Real-time NDVS searching result.

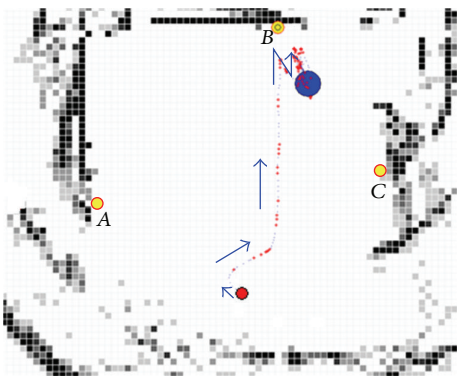


Figure 8: OSD trajectory of mobile robot OSL.

efficient and the gas concentration is detected constantly, the obstacle is declared as the gas source. LSI is "near," OI is "too high," VI is "true," and WI is "true" or "false"; we have $y = M_f x_1 x_2 x_3 x_4 = M_f \delta_2^1 \delta_3^3 \delta_2^1 \delta_2^1 = \delta_7^2$ or $y = M_f x_1 x_2 x_3 x_4 = M_f \delta_2^1 \delta_3^3 \delta_2^1 \delta_2^2 = \delta_7^2$, which means OSD.

In Figure 8, the point B is real source. Once the avoiding behavior actives, that is, the laser information is efficient, it will drive the mobile robot move away from the obstacle. But the computing results with olfaction and vision drive the robot toward the area of the obstacle. Thus, the robot will keep wandering near the area and the potential gas source is declared as the real source (B is real source).

5. Conclusion

In this paper, the multivariable fuzzy logic controller based on semitensor product (STP) for mobile robot OSL is designed. Using the basic properties of STP, the complex fuzzy control rules, and fuzzy logic inference are converted into an algebraic form. The multisensor information is the inputs of the fuzzy control system and the relative searching strategies are the outputs. The proposed multivariable fuzzy control system can activate relative searching strategies according to the timely multisensor information detected by the mobile robot, which makes the robot generate an optimization strategy to deal with the dynamic, complex, and unstructured environments. Compared with the classic olfaction-based odor source localization methods, the presented method can overcome the blindness of plume finding to a certain degree; that is, the traditional algorithms for plume finding are random searching without odor information and the mobile robot will check the scene with equal probability. Actually, the probability of suspected odor source in the scene is different. Thus, it will help to find the plume with the aid of more sensors, such as cameral. Therefore, the proposed method can make up the blindness of the olfaction-based ones to a certain degree. Equally important, any searching strategy can be updated in this framework and has no influence on the others whether based on a single sensor information or multisensor information. The proposed localization framework can degenerate into the traditional olfaction-based localization system. Most importantly, we gave an in-depth study on mobile robot odor source localization from the angle of mathematics which can reach the theory of the mobile robot odor source localization. The reliability and robustness of the proposed method are validated with the real robot experiments.

Conflict of Interests

The authors declare that they have no competing interests.

Acknowledgments

The research work of this paper is sponsored by National Natural Science Foundation (61374065) and the Research Fund for the Taishan Scholar Project of Shandong Province of China.

References

[1] M. A. Frye, M. Tarsitano, and M. H. Dickinson, "Odor localization requires visual feedback during free flight in *Drosophila melanogaster*," *Journal of Experimental Biology*, vol. 206, no. 5, pp. 843–855, 2003.

[2] J. Z. Guo and A. K. Guo, "Crossmodal interactions between olfactory and visual learning in *Drosophila*," *Science*, vol. 309, no. 5732, pp. 307–310, 2006.

[3] A. Mafra-Neto and R. T. Carde, "Fine-scale structure of pheromone plumes modulates upwind orientation of flying moths," *Nature*, vol. 369, no. 6476, pp. 142–144, 1994.

[4] R. A. Russell, "Survey of robotic applications for odor-sensing technology," *The International Journal of Robotics Research*, vol. 20, no. 2, pp. 144–162, 2001.

[5] Q.-H. Meng and F. Li, "Review of active olfaction," *Robot*, vol. 28, no. 1, pp. 89–96, 2006.

[6] A. J. Lilienthal, A. Loutfi, and T. Duckett, "Airborne chemical sensing with mobile robots," *Sensors*, vol. 6, no. 11, pp. 1616–1678, 2006.

[7] W. Naeem, R. Sutton, and J. Chudley, "Chemical plume tracing and odour source localisation by autonomous vehicles," *The Journal of Navigation*, vol. 60, no. 2, pp. 173–190, 2007.

[8] G. Kowadlo and R. A. Russell, "Robot odor localization: a taxonomy and survey," *The International Journal of Robotics Research*, vol. 27, no. 8, pp. 869–894, 2008.

[9] H. Ishida, Y. Wada, and H. Matsukura, "Chemical sensing in robotic applications: a review," *IEEE Sensors Journal*, vol. 12, no. 11, pp. 3163–3173, 2012.

[10] G. Kowadlo, D. Rawlinson, R. A. Russell, and R. Jarvis, "Bimodal search using complementary sensing (olfaction/vision) for odour source localisation," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA' 06)*, pp. 2041–2046, Piscataway, NJ , USA, May 2006.

[11] H. Ishida, H. Tanaka, H. Taniguchi, and T. Moriizumi, "Mobile robot navigation using vision and olfaction to search for a gas/odor source," *Autonomous Robots*, vol. 20, no. 3, pp. 231–238, 2006.

[12] P. Jiang, M. Zeng, Q.-H. Meng, F. Li, and Y.-H. Li, "A novel object recognition method for mobile robot localizing a single odor/gas source in complex environments," in *Proceedings of the IEEE International Conference on Robotics, Automation and Mechatronics (RAM '08)*, pp. 1–5, IEEE, Chengdu, China, September 2008.

[13] P. Jiang, Q.-H. Meng, and M. Zeng, "Mobile robot gas source localization via top-down visual attention mechanism and shape analysis," in *Proceedings of the 8th World Congress on Intelligent Control and Automation (WCICA '10)*, pp. 1818–1823, Jinan, China, July 2010.

[14] P. Jiang and Y. Zhang, "Least square estimation based mobile robot gas source localization in stable-airflow environment," in *Proceedings of the 31st Chinese Control Conference (CCC '12)*, pp. 7000–7004, IEEE, Hefei, China, July 2012.

[15] P. Jiang and Y. Zhang, "Subsumption architecture based mobile robot gas source localization in time-variant-airflow environment," in *Proceedings of the 31st Chinese Control Conference (CCC '12)*, pp. 4814–4819, Hefei, China, July 2012.

[16] B. Khaleghi, A. Khamis, F. O. Karray, and S. N. Razavi, "Multisensor data fusion: a review of the state-of-the-art," *Information Fusion*, vol. 14, no. 1, pp. 28–44, 2013.

[17] K. W. Chen, Z. P. Zhang, and J. Long, "Multisource information fusion: issues, research progress and new trends," *Computer Science*, vol. 40, no. 8, pp. 6–13, 2013.

[18] G. Feng, "A survey on analysis and design of model based fuzzy control systems," *IEEE Transactions on Fuzzy Systems*, vol. 14, no. 5, pp. 676–697, 2006.

[19] G. Feng, *Analysis and Synthesis of Fuzzy Control System: A Modelbased Approach*, CRC Press, Boca Raton, Fla, USA, 2010.

[20] D. Z. Cheng, H. S. Qi, and Z. Q. Li, *Analysis and Control of Boolean Networks: A Semi-Tensor Product Approach*, Springer, London, UK, 2011.

[21] H. T. Li and Y. Z. Wang, "Boolean derivative calculation with application to fault detection of combinational circuits via the semi-tensor product method," *Automatica*, vol. 48, no. 4, pp. 688–693, 2012.

[22] D. Z. Cheng, H. S. Qi, and Z. Q. Li, "Model construction of Boolean network via observed data," *IEEE Transactions on Neural Networks*, vol. 22, no. 4, pp. 525–536, 2011.

[23] Y. Z. Wang, C. H. Zhang, and Z. B. Liu, "A matrix approach to graph maximum stable set and coloring problems with application to multi-agent systems," *Automatica*, vol. 48, no. 7, pp. 1227–1236, 2012.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 756548, 9 pages
http://dx.doi.org/10.1155/2015/756548

Research Article

Normal Limiting Distribution of the Size of Binary Interval Trees

Jie Liu[1] and Yang Yang[2,3]

[1] *Department of Statistics and Finance, University of Science and Technology of China, Hefei 230026, China*
[2] *Department of Statistics, Nanjing Audit University, Nanjing 211815, China*
[3] *The Key Lab of Financial Engineering of Jiangsu Province, Nanjing Audit University, Nanjing 211815, China*

Correspondence should be addressed to Yang Yang; yyangmath@gmail.com

Received 1 August 2015; Revised 3 September 2015; Accepted 9 September 2015

Academic Editor: Xinguang Zhang

The limiting distribution of the size of binary interval tree is investigated. Our illustration is based on the contraction method, and it is quite different from the case in one-sided binary interval tree. First, we build a distributional recursive equation of the size. Then, we draw the expectation, the variance, and some high order moments. Finally, it is shown that the size (with suitable standardization) approaches the standard normal random variable in the Zolotarev metric space.

1. Introduction

Random trees are usually generated based on combinatorics and occur also in the context of algorithms from computer science. There are many kinds of random trees with different structures, such as recursive trees, search trees, binary trees, and interval trees. The asymptotic probability behavior of random variables in random trees has attracted more scholars' attention and has become a popular research area. Drmota [1] introduced some labelled and unlabelled random trees in his book. Devroye and Janson [2] studied the protected nodes in several random trees. Feng and Hu [3] researched the phase changes of scale-free trees. The limiting law for the height, size, and subtree of binary search trees was also considered (see [4–6]). There were also some researchers investigating the Zagreb index and nodes of random recursive trees (see [7–9]).

Binary interval tree is a random structure that underlies the process of random division of a line interval and parking problems. It has recently been a popular subject. Sibuya and Itoh [10] showed that the number of internal and external nodes in different levels of binary internal tree is asymptotically normal, from which the asymptotic normality of the size of the tree could not be achieved directly. Prodinger [11] looked into various parameters of the incomplete trie, a one-sided version of a random tree with a digital flavor. Fill et al. [12] followed with a study of the nonexistence of limit distribution for the height of the incomplete trie. Itoh and Mahmoud [13] considered five incomplete one-sided variants of binary interval trees and proved that their sizes all approach some normal random variables. Janson [14] drew the same result for a larger scale of one-sided interval trees by the renewal theory, and one kind of fragmentation trees was discussed by Janson and Neininger [15]. Javanian et al. [16] investigated the paths in m-ary interval trees. Su et al. [17] studied the complete binary interval trees and got the Law of Large Numbers. In addition, Pan et al. [18] considered the construction algorithm about binary interval trees.

The binary interval tree is a tree associated with repeated divisions of a line interval of length x. The process of divisions is as follows. If $x < 1$, there is no division in effect; the associated interval tree consists only of one terminal node. Supposing that $x \geq 1$, we begin with the interval $(0, x)$. Divide the interval $(0, x)$ into two subintervals by choosing U_x, a point uniformly distributed over the interval $(0, x)$. Then, we get two intervals, $(0, U_x)$ and (U_x, x). Each of the two subintervals is further divided at a uniform point of its length, and two smaller subintervals are got as before. If the length of the subinterval is less than 1, we stop the division. Repeat this

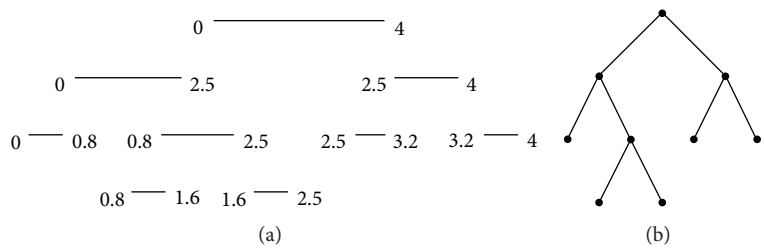


Figure 1: (a) The division process. (b) The binary interval tree.

process until the length of every interval (or subinterval) is less than 1.

We take $x = 4$, for instance. Figures 1(a) and 1(b) show how the above random division process of interval generates a binary interval tree.

If some different conditions are added and those intervals satisfying the conditions are not allowed to be divided (see [13, 14]), then we can get different incomplete interval trees. In particular, if we only divide one subinterval of every interval, then the interval tree we get is the so-called one-sided interval tree (see [13]).

It is obvious that interval tree could embody many properties of random division, so it can elicit lots of valuable subjects related to probability. For example, for $x > 0$, the height of the interval tree is the greatest level of all subintervals after the divisions, denoted by H_x; the total number of nodes of an interval tree is the total number of intervals that were got from the random division process, and so on. Let S_x be the size of the interval trees, that is, the total number of nodes of the binary interval trees. Our intention is to investigate the random variable S_x, the size of binary interval trees.

In this paper, the central limit theorem of the size of binary interval trees is investigated. In view of the difficulty to calculate the moment generating function of S_x, the method we used is completely different from that in the case of one-sided interval trees. In Section 2, we build a distributional recursive equation of S_x and give the expectation, the variance, and some high order moments of S_x. In Section 3, via the contraction method, the limit law of S_x is shown to approach the unique solution of a fixed-point distributional equation in the Zolotarev metric space. Finally, we demonstrate that S_x, with suitable standardization, converges to a normal limiting random variable, as $x \to \infty$.

2. The Moments of S_x

Compared with the one-sided interval trees, the properties of binary interval trees are much more complex. There are a lot of difficulties when it comes to obtaining the moment generating function of S_x. Therefore, the method used in the case of one-sided interval trees (see [13]) is no longer applicable. Here, we build a distributional recursive equation of S_x. We can calculate the expectation and the variance of S_x. Furthermore, we find that the order of the fourth central moment of S_x is $O(x^2)$ as x goes to infinity.

From the definition of binary interval tree, it is easy to see that $S_1 = 3$ and $S_x = 1$, for $x < 1$. For our purpose to investigate the case of $x \geq 1$, let U_x denote the point chosen uniformly from interval $(0, x)$; hence, $U_x \sim U(0, x)$. For any fixed real number $0 < u < x$, if $U_x = u$, we denote $S_u^{(1)}$ to be the size of the left subtree associated with the interval $(0, u)$. Correspondingly, $S_{x-u}^{(2)}$ denotes the size of the right subtree associated with the interval (u, x). According to the rule of division, we can see that $S_u^{(1)}$ and $S_{x-u}^{(2)}$ are mutually independent. Thus, we have

$$S_x\big|_{U=u} \overset{d}{=} 1 + S_u^{(1)} + S_{x-u}^{(2)}, \quad \forall 0 < u < x. \tag{1}$$

This formula implies that if $U_x = u$ is given, S_x has the same distribution as $1 + S_u^{(1)} + S_{x-u}^{(2)}$. Obviously, we can rewrite the above formula as

$$S_x \overset{d}{=} 1 + S_{U_x}^{(1)} + S_{x-U_x}^{(2)}. \tag{2}$$

Define

$$\begin{aligned} m_1(x) &:= \mathbf{E}S_x; \\ m_2(x) &:= \mathbf{E}(S_x)^2. \end{aligned} \tag{3}$$

It is easy to see that

$$\begin{aligned} m_1(1) &= \mathbf{E}S_1 = 3; \\ m_2(1) &= 9; \\ m_1(x) &= m_2(x) = 1, \quad 0 < x < 1. \end{aligned} \tag{4}$$

From the distributional recursive equation (2) and the above boundary conditions, Su et al. [17] calculated the expectation $\mathbf{E}S_x$ and the variance $\mathbf{Var}\, S_x$, for any $x \geq 0$.

Lemma 1. *Let S_x be the size of a binary interval tree. Then*

$$\mathbf{E}S_x = m_1(x) = 4x - 1, \quad x \geq 1. \tag{5}$$

Lemma 2. *Let S_x be the size of a binary interval tree. Then*

$$\mathbf{Var}\, S_x = \begin{cases} 32x \ln x - 16x^2 + 8x + 8, & 1 \leq x \leq 2, \\ (32 \ln 2 - 20)\, x, & x \geq 2. \end{cases} \tag{6}$$

In order to prove that the asymptotic distribution of S_x is normal, we also need the order of $\mathbf{E}(S_x - \mathbf{E}S_x)^4$ as $x \to \infty$. The following proposition shows the fourth central moment of S_x.

Proposition 3. *Let S_x be the size of a binary interval tree. Then*

$$\mathbf{E}\left(S_x - \mathbf{E}\left[S_x\right]\right)^4 = O\left(x^2\right), \quad x \longrightarrow \infty. \tag{7}$$

Proof. See the appendix. □

3. The CLT for S_x

In this section, we will prove the asymptotic normality of S_x as $x \to \infty$. The main method is the contraction method and some metrics are needed especially the Zolotarev metrics (see [19]).

First we introduce the Zolotarev metrics. Denote the distribution of the random variable X by $\mathscr{L}(X)$. Let $\mathscr{D}$ be the set of the distributions of all real random variables, and define

$$\begin{aligned}\mathscr{D}^* = \Big\{F : F \in \mathscr{D},\ \int_{\mathscr{R}} x dF(x) = 0,\ \int_{\mathscr{R}} x^2 dF(x) \\ = 1,\ \int_{\mathscr{R}} |x|^3 dF(x) < \infty\Big\}.\end{aligned} \tag{8}$$

It can be verified that random variable Z with $\mathscr{L}(Z) = \mathscr{N}(0, \sigma^2)$ satisfies the following formula. For any $u \in [0, 1]$,

$$Z \overset{d}{=} Z\sqrt{u} + \overline{Z}\sqrt{1-u}, \tag{9}$$

and more generally, we have the following lemma.

Lemma 4. *If Z and $\overline{Z}$ are standard normal random variables, U is uniformly distributed over interval $[0, 1]$, and $(U, Z, \overline{Z})$ are mutually independent and then one has*

$$Z \overset{d}{=} Z\sqrt{U} + \overline{Z}\sqrt{1-U}. \tag{10}$$

Proof. In fact, for any $u \in [0, 1]$, we have

$$\begin{aligned}&\mathbf{E}\exp\left\{\mathbf{i}t\left(\sqrt{u}Z + \sqrt{1-u}\,\overline{Z}\right)\right\} \\ &= \mathbf{E}\exp\left\{\mathbf{i}t\left(\sqrt{u}Z + \sqrt{1-u}\,\overline{Z}\right)\right\} \\ &= \mathbf{E}e^{\mathbf{i}(t\sqrt{u})Z}\mathbf{E}e^{\mathbf{i}(t\sqrt{1-u})\overline{Z}} \\ &= \exp\left\{-\frac{ut^2}{2}\right\}\exp\left\{-\frac{(1-u)t^2}{2}\right\} = \exp\left\{-\frac{t^2}{2}\right\}.\end{aligned} \tag{11}$$

Therefore,

$$\begin{aligned}&\mathbf{E}\exp\left\{\mathbf{i}t\left(\sqrt{U}Z + \sqrt{1-U}\,\overline{Z}\right)\right\} \\ &= \int_0^1 \mathbf{E}\exp\left\{\mathbf{i}t\left(\sqrt{u}Z + \sqrt{1-u}\,\overline{Z}\right)\right\} du \\ &= \int_0^1 \exp\left\{-\frac{t^2}{2}\right\} du = \exp\left\{-\frac{t^2}{2}\right\}.\end{aligned} \tag{12}$$

But, we can find that, in the set $\mathscr{D}^*$, there is only one distribution, the standard normal $\mathscr{N}(0, 1)$, satisfying (10). □

Suppose that m is a nonnegative integer. Denote $\mathfrak{F}^{(m)}$ by the set of all real functions that are m times continuous and differentiable, defined on the real line. Let

$$\begin{aligned}&\mathfrak{F}_\alpha^{(m)} \\ &:= \left\{f : f \in \mathfrak{F}^{(m)},\ \left|f^{(m)}(x) - f^{(m)}(y)\right| \le |x-y|^\alpha\right\},\end{aligned} \tag{13}$$

where $0 < \alpha \le 1$ is a fixed real number. Let $s = m + \alpha$ and

$$\begin{aligned}\zeta_s(X, Y) &:= \zeta_s(\mathscr{L}(X), \mathscr{L}(Y)) \\ &= \sup\left\{\left|\mathbf{E}f(X) - \mathbf{E}f(Y)\right| : f \in \mathfrak{F}_\alpha^{(m)}\right\},\end{aligned} \tag{14}$$

and then ζ_s is the Zolotarev metrics with order s on the set $\mathscr{D}$. According to the properties of the Zolotarev metric, we know

$$\begin{aligned}&\zeta_s(X, Y) < \infty \\ &\qquad \Longleftrightarrow \mathbf{E}|X|^s + \mathbf{E}|Y|^s < \infty, \\ &\mathbf{E}X^k = \mathbf{E}Y^k, \quad k = 1, \ldots, m.\end{aligned} \tag{15}$$

Therefore, we can choose ζ_3 as the metric we need on the subset $\mathscr{D}^*$ (see [20, 21]); that is, $m = 2$, $\alpha = 1$. This is due to the fact that, for any $\mathscr{L}(X) \in \mathscr{D}^*$ and $\mathscr{L}(Y) \notin \mathscr{D}^*$, we have $\zeta_3(X, Y) = \infty$, but if $\mathscr{L}(X), \mathscr{L}(Y) \in \mathscr{D}^*$, then $\zeta_3(X, Y) < \infty$.

The metric $\zeta_s(X, Y)$ has several properties as follows (see [20]):

(1) For any constant $c > 0$,

$$\zeta_s(cX, cY) = c^s\zeta_s(X, Y); \tag{16}$$

(2) if random variables Y and (X_1, X_2) are mutually independent, then

$$\zeta_s(X_1 + Y, X_2 + Y) \le \zeta_s(X_1, X_2); \tag{17}$$

(3) for random variables X and $\{X_n, n = 1, 2, 3, \ldots\}$,

$$\zeta_s(X_n, X) \longrightarrow 0 \Longrightarrow X_n \xrightarrow{d} X. \tag{18}$$

Now, we begin to prove the main result in this paper.

Theorem 5. *Let S_x be the size of a binary interval tree. Then, as $x \to \infty$,*

$$\frac{S_x - \mathbf{E}S_x}{\sqrt{\mathbf{Var}\,S_x}} \xrightarrow{d} \mathscr{N}(0, 1). \tag{19}$$

Proof. Denote

$$S_x^* := \frac{S_x - (4x-1)}{\sqrt{(32\ln 2 - 20)\,x}}, \quad x > 0,$$
$$h(x) := \sqrt{\frac{32x\ln x - 16x^2 + 8x + 8}{(32\ln 2 - 20)\,x}}, \quad x > 0. \tag{20}$$

Then from Lemmas 1 and 2, we know that

$$S_x^* = \begin{cases} \dfrac{S_x - \mathbf{E}S_x}{\sqrt{\mathbf{Var}\,S_x}}, & x \geq 2, \\ \dfrac{S_x - \mathbf{E}S_x}{\sqrt{\mathbf{Var}\,S_x}} \cdot h(x), & 1 \leq x < 2, \\ \dfrac{2-4x}{\sqrt{(32\ln 2 - 20)\,x}}, & 0 < x < 1. \end{cases} \tag{21}$$

So, we have $\mathscr{L}(S_x^*) \in \mathscr{D}^*$ for $x \geq 2$ and $\mathscr{L}(S_x^*) \notin \mathscr{D}^*$ for $0 < x < 2$.

According to the correlative inequality in [21], for any $\mathscr{L}(X) \in \mathscr{D}^*, \mathscr{L}(Y) \in \mathscr{D}^*$,

$$\zeta_3(X,Y) \leq \frac{\Gamma(2)}{\Gamma(4)} \int_{\mathscr{R}} |t|^3\, d\,|P(X<t) - P(Y<t)|, \tag{22}$$

where Γ is the gamma function. Assume that the distribution of random variable Z is $\mathscr{N}(0,1)$. It follows from Proposition 3 that

$$\sup_{x\geq 4} \mathbf{E}\left(S_x^*\right)^4 < \infty. \tag{23}$$

Therefore, there exists a constant $C > 0$ such that

$$\sup_{x\geq 4} \zeta_3\left(S_x^*, Z\right) \leq C\left(\sup_{x\geq 4} \mathbf{E}\left|S_x^*\right|^3 + \mathbf{E}\,|Z|^3\right) < \infty. \tag{24}$$

Denote

$$a_x := \zeta_3\left(\mathscr{L}\left(S_x^*\right), \Phi\right) = \zeta_3\left(S_x^*, Z\right), \tag{25}$$

where Φ is standard normal distribution and Z is standard normal random variable; then we can see that

$$0 \leq b := \limsup_{x\to\infty} a_x < \infty. \tag{26}$$

Now, we just need to prove that $b = 0$; then the theorem follows.

Suppose that $x \geq 4$; by (A.1) and (21), we have

$$S_x^*\big|_{U_x=t} = \frac{S_x - (4x-1)}{\sqrt{(32\ln 2 - 20)\,x}}\bigg|_{U_x=t}$$
$$\stackrel{d}{=} \begin{cases} \dfrac{S_t^{(1)} - \mathbf{E}\left[S_t^{(1)}\right]}{\sqrt{(32\ln 2 - 20)\,x}} + \dfrac{S_{x-t}^{(2)} - \mathbf{E}\left[S_{x-t}^{(2)}\right]}{\sqrt{(32\ln 2 - 20)\,x}}, & 2 \leq t \leq x-2, \\ \dfrac{S_t^{(1)} - \mathbf{E}\left[S_t^{(1)}\right]}{\sqrt{(32\ln 2 - 20)\,x}} + \dfrac{S_{x-t}^{(2)} - \mathbf{E}\left[S_{x-t}^{(2)}\right]}{\sqrt{(32\ln 2 - 20)\,x}}, & 1 \leq t < 2, \\ \dfrac{S_t^{(1)} - \mathbf{E}\left[S_t^{(1)}\right]}{\sqrt{(32\ln 2 - 20)\,x}} + \dfrac{S_{x-t}^{(2)} - \mathbf{E}\left[S_{x-t}^{(2)}\right]}{\sqrt{(32\ln 2 - 20)\,x}}, & x-2 < t \leq x-1, \\ \dfrac{S_{x-t}^{(2)} - \mathbf{E}\left[S_{x-t}^{(2)}\right] - (4t-2)}{\sqrt{(32\ln 2 - 20)\,x}}, & 0 < t < 1 \\ \dfrac{S_t^{(1)} - \mathbf{E}\left[S_t^{(1)}\right] - [4(x-t)-2]}{\sqrt{(32\ln 2 - 20)\,x}}, & x-1 < t < x \end{cases} \tag{27}$$
$$\stackrel{d}{=} \begin{cases} S_t^* \sqrt{\dfrac{t}{x}} + \widetilde{S}_{x-t}^* \sqrt{\dfrac{x-t}{x}}, & 2 \leq t \leq x-2; \\ S_t^* h(t) \sqrt{\dfrac{t}{x}} + \widetilde{S}_{x-t}^* \sqrt{\dfrac{x-t}{x}}, & 1 \leq t < 2, \\ S_t^* \sqrt{\dfrac{t}{x}} + \widetilde{S}_{x-t}^* h(x-t) \sqrt{\dfrac{x-t}{x}}, & x-2 < t \leq x-1, \\ \widetilde{S}_{x-t}^* \sqrt{\dfrac{x-t}{x}} - \dfrac{4t-2}{\sqrt{(32\ln 2 - 20)\,x}}, & 0 < t < 1; \\ S_t^* \sqrt{\dfrac{t}{x}} - \dfrac{4(x-t)-2}{\sqrt{(32\ln 2 - 20)\,x}}, & x-1 < t < x, \end{cases}$$

where $U_x = t$ is the first point chosen from interval $(0,x)$ and $\{\widetilde{S}_x^*, x > 0\}$ is an independent copy of $\{S_x^*, x > 0\}$.

If we denote $U := U_x/x$, then $U \sim U(0,1)$ and we can rewrite the above formula as

$$S_x^*\big|_{U=u} = \frac{S_x - (4x-1)}{\sqrt{(32\ln 2 - 20)\,x}}\bigg|_{U=u} \tag{28}$$
$$\stackrel{d}{=} \begin{cases} S_{ux}^* \sqrt{u} + \widetilde{S}_{(1-u)x}^* \sqrt{1-u}, & \dfrac{2}{x} \leq u \leq 1 - \dfrac{2}{x}; \\ S_{ux}^* h(ux) \sqrt{u} + \widetilde{S}_{(1-u)x}^* \sqrt{1-u}, & \dfrac{1}{x} \leq u < \dfrac{2}{x}, \\ S_{ux}^* \sqrt{u} + \widetilde{S}_{(1-u)x}^* h((1-u)x) \sqrt{1-u}, & 1 - \dfrac{2}{x} < u \leq 1 - \dfrac{1}{x}, \\ \widetilde{S}_{(1-u)x}^* \sqrt{1-u} - \dfrac{4ux-2}{\sqrt{(32\ln 2 - 20)\,x}}, & 0 < u < \dfrac{1}{x}; \\ S_{ux}^* \sqrt{u} - \dfrac{4(1-u)x-2}{\sqrt{(32\ln 2 - 20)\,x}}, & 1 - \dfrac{1}{x} < u < 1. \end{cases} \tag{29}$$

According to the definition of $\mathscr{D}^*$ and ζ_3, it could be found that $(S_x^* \mid U_x = t) \in \mathscr{D}^*$ for $2 < t < x-2$ and $S_x^* \in \mathscr{D}^*$.

If we define $S'_x := (S^*_x \mid U_x < 2 \text{ or } U_x > x-2)$, then we can also see that $S'_x \in \mathscr{D}^*$. Furthermore, $\mathbf{E}((S'_x)^4) \le C_1$ for some positive constant C_1 by conditioning on U_x and using the similar calculation in the appendix. Hence,

$$\zeta_3\left(S'_x, Z\right) \le \beta, \tag{30}$$

for some positive constant β.

As we had pointed out before, the standard normal distribution is the only distribution satisfying (10) in the set $\mathscr{D}^*$. From (25), (14), and Lemma 4, for $x > 4$, we have

$$\begin{aligned}
a_x &= \zeta_3\left(S^*_x, Z\right) \le \zeta_3\left(S'_x, Z\right) \cdot \frac{4}{x} \\
&\quad + \int_{2/x}^{1-2/x} \zeta_3\left(\left(S^*_x \mid U = u\right), Z\right) du \le \frac{4\beta}{x} \\
&\quad + \int_{2/x}^{1-2/x} \zeta_3\left(S^*_{xu}\sqrt{u} + \overline{S}^*_{x(1-u)}\sqrt{1-u}, Z\sqrt{u}\right. \\
&\quad \left. + \overline{Z}\sqrt{1-u}\right) du \quad (\text{By (15) and (29)}) \\
&\le \frac{4\beta}{x} + \int_{2/x}^{1-2/x} \zeta_3\left(S^*_{xu}\sqrt{u} + \overline{S}^*_{x(1-u)}\sqrt{1-u}, Z\sqrt{u}\right. \\
&\quad \left. + \overline{S}^*_{x(1-u)}\sqrt{1-u}\right) du + \int_{2/x}^{1-2/x} \zeta_3\left(Z\sqrt{u}\right. \\
&\quad \left. + \overline{S}^*_{x(1-u)}\sqrt{1-u}, Z\sqrt{u} + \overline{Z}\sqrt{1-u}\right) du \\
&\le \frac{4\beta}{x} + \int_{2/x}^{1-2/x} \zeta_3\left(S^*_{xu}\sqrt{u}, Z\sqrt{u}\right) du \\
&\quad + \int_{2/x}^{1-2/x} \zeta_3\left(\overline{S}^*_{x(1-u)}\sqrt{1-u}, \overline{Z}\sqrt{1-u}\right) du \\
&= \frac{4\beta}{x} + 2\int_{2/x}^{1-2/x} \zeta_3\left(S^*_{xu}\sqrt{u}, Z\sqrt{u}\right) du \\
&= \frac{4\beta}{x} + 2\int_{2/x}^{1-2/x} u^{3/2}\zeta_3\left(S^*_{xu}, Z\right) du \\
&= \frac{4\beta}{x} + 2\int_{2/x}^{1-2/x} u^{3/2} a_{xu} du.
\end{aligned} \tag{31}$$

Given $\varepsilon > 0$, let $\delta > 0$ be small enough such that $\beta\delta^{5/2} < \varepsilon/8$. For any fixed $\delta > 0$, when x is sufficiently large, then

$$\begin{aligned}
&\frac{4\beta}{x} < \frac{\varepsilon}{10}, \\
&\frac{2}{x} < \delta, \\
&\sup_{\delta \le u \le 1} a_{xu} < b + \varepsilon.
\end{aligned} \tag{32}$$

Thus,

$$\begin{aligned}
2\int_{2/x}^{\delta} u^{3/2} a_{xu} du &\le 2\beta \int_{2/x}^{\delta} u^{3/2} du \le 2\beta \int_0^{\delta} u^{3/2} du \\
&= \frac{4\beta\delta^{5/2}}{5} < \frac{\varepsilon}{10}; \\
2\int_{\delta}^{1-2/x} u^{3/2} a_{xu} du &\le 2(b+\varepsilon) \int_{\delta}^{1-2/x} u^{3/2} du \\
&\le 2(b+\varepsilon) \int_0^1 u^{3/2} du < \frac{4(b+\varepsilon)}{5},
\end{aligned} \tag{33}$$

where β is the constant as before and x is sufficiently large. It implies that

$$\begin{aligned}
a_x &\le \frac{4\beta}{x} + 2\int_{2/x}^{\delta} u^{3/2} a_{xu} du + 2\int_{\delta}^{1-2/x} u^{3/2} a_{xu} du \\
&< \frac{\varepsilon}{10} + \frac{\varepsilon}{10} + \frac{4(b+\varepsilon)}{5} < \frac{4b}{5} + \varepsilon,
\end{aligned} \tag{34}$$

when x is sufficiently large. Therefore,

$$b := \limsup_{x\to\infty} a_x \le \frac{4b}{5} + \varepsilon. \tag{35}$$

From this equation and the arbitrariness of $\varepsilon > 0$, we can conclude $b = 0$ and

$$\lim_{x\to\infty} \zeta_3\left(S^*_x, Z\right) = \lim_{x\to\infty} a_x = 0 \tag{36}$$

immediately. By (18), the theorem holds. □

Appendix

Proof of Proposition 3

From the process of generating the binary interval trees, it is obvious that, for given $U_x = t$, $t \in (0, x)$,

$$\begin{aligned}
&\left(S_x - \mathbf{E}\left[S_x\right]\right)\big|_{U_x=t} \\
&\stackrel{d}{=} \begin{cases}
\left(S^{(1)}_t - \mathbf{E}\left[S^{(1)}_t\right]\right) + \left(S^{(2)}_{x-t} - \mathbf{E}\left[S^{(2)}_{x-t}\right]\right), & 1 \le t \le x-1; \\
S^{(2)}_{x-t} - \mathbf{E}\left[S^{(2)}_{x-t}\right] - (4t-2), & 0 < t < 1; \\
S^{(1)}_t - \mathbf{E}\left[S^{(1)}_t\right] - [4(x-t)-2], & x-1 < t < x,
\end{cases}
\end{aligned} \tag{A.1}$$

where $U_x = t$ is the first point chosen from interval $(0, x)$. For $x \ge 1$, if we denote

$$\begin{aligned}
T_x &:= S^{(1)}_x - \mathbf{E}\left[S^{(1)}_x\right], \\
T^*_x &:= S^{(2)}_x - \mathbf{E}\left[S^{(2)}_x\right],
\end{aligned} \tag{A.2}$$

then we have

$$T_x\big|_{U_x=t} \stackrel{d}{=} \begin{cases}
T_t + T^*_{x-t}, & 1 \le t \le x-1; \\
T^*_{x-t} - (4t-2), & 0 < t < 1; \\
T_t - [4(x-t)-2], & x-1 < t < x.
\end{cases} \tag{A.3}$$

We need to calculate $\mathbf{E}T_x^3$ first before we get $\mathbf{E}T_x^4$. For $x > 3$, we have

$$\begin{aligned}
\mathbf{E}\left[T_x\right]^3 &= \mathbf{E}\left[\mathbf{E}\left(T_x^3 \mid U_x\right)\right] = \frac{1}{x}\int_0^1 \mathbf{E}\left[T_{x-t}\right. \\
&\quad \left.-(4t-2)\right]^3 dt + \frac{1}{x}\int_{x-1}^{x} \mathbf{E}\left\{T_t\right. \\
&\quad \left.-[4(x-t)-2]\right\}^3 dt + \frac{1}{x}\int_1^{x-1} \mathbf{E}\left[T_t + T_{x-t}^*\right]^3 dt \\
&= \frac{2}{x}\int_{x-1}^{x} \mathbf{E}\left\{T_t - [4(x-t)-2]\right\}^3 dt + \frac{1}{x} \\
&\quad \cdot \int_1^{x-1} \mathbf{E}\left[T_t + T_{x-t}^*\right]^3 dt = \left(-\frac{2}{x}\right. \\
&\quad \cdot \int_{x-1}^{x} [4(x-t)-2]^3 dt + \frac{6}{x} \\
&\quad \cdot \int_{x-1}^{x} [4(x-t)-2]^2 \mathbf{E}\left[T_t\right] dt - \frac{6}{x} \\
&\quad \cdot \int_{x-1}^{x} [4(x-t)-2] \mathbf{E}\left[T_t^2\right] dt + \frac{2}{x} \\
&\quad \left.\cdot \int_{x-1}^{x} \mathbf{E}\left[T_t^3\right] dt\right) + \frac{1}{x}\int_1^{x-1} \left(\mathbf{E}\left[T_t\right]^3\right. \\
&\quad + 3\mathbf{E}\left[\left(T_t\right)^2 T_{x-t}^*\right] + 3\mathbf{E}\left[T_t\left(T_{x-t}^*\right)^2\right] \\
&\quad \left.+ \mathbf{E}\left[T_{x-t}^*\right]^3\right) dt.
\end{aligned} \tag{A.4}$$

In view of the independence between T_t and T_{x-t}^* and that $\mathbf{E}[T_t] = \mathbf{E}[T_t^*] = 0$ holds for any $1 \le t \le x-1$, we have

$$\begin{aligned}
\mathbf{E}\left[T_x\right]^3 &= -\frac{2}{x}\int_{x-1}^{x} [4(x-t)-2]^3 dt \\
&\quad - \frac{6}{x}\int_{x-1}^{x} [4(x-t)-2] \mathbf{E}\left[T_t^2\right] dt \\
&\quad + \frac{2}{x}\int_{x-1}^{x} \mathbf{E}\left[T_t^3\right] dt + \frac{2}{x}\int_1^{x-1} \mathbf{E}\left[T_t^3\right] dt \\
&= -\frac{2}{x}\int_{x-1}^{x} [4(x-t)-2]^3 dt \\
&\quad - \frac{6}{x}\int_{x-1}^{x} [4(x-t)-2] \mathbf{E}\left[T_t^2\right] dt \\
&\quad + \frac{2}{x}\int_1^{x} \mathbf{E}\left[T_t^3\right] dt \\
&:= M_1 + M_2 + \frac{2}{x}\int_1^{x} \mathbf{E}\left[T_t^3\right] dt.
\end{aligned} \tag{A.5}$$

It is easy to see that

$$M_1 = -\frac{1}{2x}\int_{-2}^{2} u^3 du = 0, \tag{A.6}$$

and when $x > 3$, for the part M_2, we have

$$\begin{aligned}
M_2 &= -\frac{6}{x}\int_{x-1}^{x} [4(x-t)-2] \mathbf{Var}\left[S_t\right] dt \\
&= -\frac{6}{x}\int_{x-1}^{x} [4(x-t)-2] [(32\ln 2 - 20)t] dt \\
&= -\frac{6(32\ln 2 - 20)}{x}\int_0^1 (4t-2)(x-t) dt \\
&= \frac{2(32\ln 2 - 20)}{x}.
\end{aligned} \tag{A.7}$$

Therefore,

$$\mathbf{E}\left[T_x^3\right] = \frac{2}{x}\int_1^{x} \mathbf{E}\left[T_t^3\right] dt + \frac{2(32\ln 2 - 20)}{x}, \quad x > 3. \tag{A.8}$$

That is,

$$x\mathbf{E}\left[T_x^3\right] = 2\int_1^{x} \mathbf{E}\left[T_t^3\right] dt + 2(32\ln 2 - 20), \quad x > 3. \tag{A.9}$$

Via differentiation with respect to x, we get the differential equation:

$$\left(\mathbf{E}\left[T_x^3\right]\right)' - \frac{1}{x}\mathbf{E}\left[T_x^3\right] = 0, \quad x > 3. \tag{A.10}$$

The solution to this differential equation is

$$\mathbf{E}\left[T_x^3\right] = k_0 x, \quad x > 3, \tag{A.11}$$

where k_0 is a constant real number.

Similarly, for $\mathbf{E}[T_x]^4$, when $x > 4$, we have

$$\begin{aligned}
\mathbf{E}\left[T_x\right]^4 &= \frac{2}{x}\int_{x-1}^{x} \mathbf{E}\left\{T_t - [4(x-t)-2]\right\}^4 dt \\
&\quad + \frac{1}{x}\int_1^{x-1} \mathbf{E}\left[T_t + T_{x-t}^*\right]^4 dt.
\end{aligned} \tag{A.12}$$

Because T_t is independent of T_{x-t}^*, and $\mathbf{E}T_t = 0$ holds for any $1 \le t \le x-1$, we get

$$\begin{aligned}
\mathbf{E}\left[T_x\right]^4 &= \frac{2}{x}\int_{x-1}^{x} [4(x-t)-2]^4 dt \\
&\quad + \frac{12}{x}\int_{x-1}^{x} [4(x-t)-2]^2 \mathbf{E}\left[T_t^2\right] dt \\
&\quad - \frac{8}{x}\int_{x-1}^{x} [4(x-t)-2] \mathbf{E}\left[T_t^3\right] dt
\end{aligned}$$

$$
\begin{aligned}
&+\frac{2}{x}\int_{x-1}^{x}\mathbf{E}\left[T_t^4\right]dt\\
&+\frac{1}{x}\int_{1}^{x-1}\left(\mathbf{E}\left[T_t^4\right]+\mathbf{E}\left[T_{x-t}^*\right]^4+6\mathbf{E}\left[T_tT_{x-t}^*\right]^2\right)dt\\
&=\frac{2}{x}\int_{x-1}^{x}\left[4(x-t)-2\right]^4dt\\
&+\frac{12}{x}\int_{x-1}^{x}\left[4(x-t)-2\right]^2\mathbf{E}\left[T_t^2\right]dt\\
&-\frac{8}{x}\int_{x-1}^{x}\left[4(x-t)-2\right]\mathbf{E}\left[T_t^3\right]dt\\
&+\frac{2}{x}\int_{1}^{x}\mathbf{E}\left[T_t\right]^4dt+\frac{6}{x}\int_{1}^{x-1}\mathbf{E}\left[T_tT_{x-t}^*\right]^2dt:=I_1\\
&+I_2+I_3+I_4+I_5.
\end{aligned}
\tag{A.13}
$$

In particular, for the part I_1, we have

$$
I_1=\frac{32}{5x}. \tag{A.14}
$$

When $x>3$, for the part I_2, we have

$$
\begin{aligned}
I_2&=\frac{12}{x}\int_{x-1}^{x}\left[4(x-t)-2\right]^2\mathbf{E}\left[T_t^2\right]dt\\
&=\frac{12}{x}\int_{x-1}^{x}\left[4(x-t)-2\right]^2\mathbf{Var}\left[S_t\right]dt\\
&=\frac{12}{x}\int_{x-1}^{x}\left[4(x-t)-2\right]^2\left[(32\ln 2-20)t\right]dt\\
&=\frac{12(32\ln 2-20)}{x}\int_{0}^{1}(4t-2)^2(x-t)\,dt\\
&=\frac{48}{3}(32\ln 2-20)-\frac{24(32\ln 2-20)}{3x}\\
&:=12k_1-\frac{6k_1}{x},
\end{aligned}
\tag{A.15}
$$

where $k_1:=4(32\ln 2-20)/3$ is a constant.

When $x>4$, for the part I_3, we have

$$
\begin{aligned}
I_3&=-\frac{8}{x}\int_{x-1}^{x}\left[4(x-t)-2\right]\mathbf{E}\left[T_t^3\right]dt\\
&=-\frac{8}{x}\int_{x-1}^{x}\left[4(x-t)-2\right]k_0t\,dt\\
&=-\frac{8k_0}{x}\int_{0}^{1}(4t-2)(x-t)\,dt=\frac{8k_0}{3x},
\end{aligned}
\tag{A.16}
$$

where k_0 is the same as that in (A.11).

When $x>4$, for the part I_5, we have

$$
\begin{aligned}
I_5&=\frac{6}{x}\int_{1}^{x-1}\mathbf{E}\left[T_tT_{x-t}^*\right]^2dt\\
&=\frac{6}{x}\int_{1}^{2}\mathbf{E}\left[T_t^2\right]\mathbf{E}\left[T_{x-t}^*\right]^2dt\\
&\quad+\frac{6}{x}\int_{x-2}^{x-1}\mathbf{E}\left[T_t^2\right]\mathbf{E}\left[T_{x-t}^*\right]^2dt\\
&\quad+\frac{6}{x}\int_{2}^{x-2}\mathbf{E}\left[T_t^2\right]\mathbf{E}\left[T_{x-t}^*\right]^2dt\\
&=\frac{12}{x}\int_{1}^{2}\mathbf{E}\left[T_t^2\right]\mathbf{E}\left[T_{x-t}^*\right]^2dt\\
&\quad+\frac{6}{x}\int_{2}^{x-2}\mathbf{E}\left[T_t^2\right]\mathbf{E}\left[T_{x-t}^*\right]^2dt.
\end{aligned}
\tag{A.17}
$$

Noting that $\mathbf{E}[T_t^2]=\mathbf{Var}[S_t]$ and (6), we can see that

$$
\begin{aligned}
&\int_{1}^{x}\mathbf{E}\left[T_t^2\right]\mathbf{E}\left[T_{x-t}^*\right]^2dt\\
&=2\int_{1}^{2}\left(32t\ln t-16t^2+8t+8\right)\\
&\cdot\left((32\ln 2-20)(x-t)\right)dt\\
&+\int_{2}^{x-2}\left((32\ln 2-20)t\right)\\
&\cdot\left((32\ln 2-20)(x-t)\right)dt=\frac{1}{6}(32\ln 2-20)^2\\
&\cdot x^3-\frac{4}{3}(32\ln 2-20)^2x^2+\frac{1}{3}(32\ln 2-20)\\
&\cdot(256\ln 2-168)x+\frac{16}{9}(32\ln 2-20):=a_3x^3\\
&+a_2x^2+a_1x+a_0,
\end{aligned}
\tag{A.18}
$$

where

$$
\begin{aligned}
a_0&=\frac{16}{9}(32\ln 2-20);\\
a_1&=\frac{1}{3}(32\ln 2-20)(256\ln 2-168);\\
a_2&=-\frac{4}{3}(32\ln 2-20)^2;\\
a_3&=\frac{1}{6}(32\ln 2-20)^2.
\end{aligned}
\tag{A.19}
$$

Therefore,

$$\mathbf{E}\left[T_x\right]^4 = \frac{2}{x}\int_1^x \mathbf{E}\left[T_t\right]^4 dt + I_1 + I_2 + I_3 + I_5, \quad x > 4. \tag{A.20}$$

That is,

$$x\mathbf{E}\left[T_x^4\right] = 2\int_1^x \mathbf{E}\left[T_t\right]^4 dt + \left(I_1 + I_2 + I_3 + I_5\right)x, \quad x > 4. \tag{A.21}$$

Via differentiation with respect to x, we get the differential equation:

$$\left(\mathbf{E}T_x^4\right)' - \frac{1}{x}\mathbf{E}T_x^4 = 12\frac{k_1}{x} + 6\left(3a_3x + 2a_2 + a_1\frac{1}{x}\right), \quad x > 4. \tag{A.22}$$

The solution to this differential equation is

$$\mathbf{E}T_x^4 = 18a_3x^2 + 12a_2x\ln x + cx - 6a_1 - 12k_1, \quad x > 4, \tag{A.23}$$

where c is a constant and the constants k_1, a_1, a_2, a_3 are real numbers as defined before. From this equation, Proposition 3 follows.

Consent

Informed consent was obtained from all individual participants included in the study.

Conflict of Interests

The authors declare that they have no conflict of interests.

Acknowledgments

The authors are most grateful to the referee and the editor for their very thorough reading of the paper and valuable suggestions, which greatly improve the original results and presentation of this paper. Jie Liu's work was supported by the National Natural Science Foundation of China (nos. 11101394, 71471168, and 71520107002), China Postdoctoral Science Foundation Funded Project (nos. 201104312 and 20100480688), Fund for the Doctoral Program of Higher Education Foundation (no. 20113402120005), and the Fundamental Research Funds for the Central Universities of China (no. WK2040160008). Yang Yang's work was supported by the National Natural Science Foundation of China (no. 71471090), the Humanities and Social Sciences Foundation of the Ministry of Education of China (no. 14YJCZH182), China Postdoctoral Science Foundation (nos. 2014T70449 and 2012M520964), Natural Science Foundation of Jiangsu Province of China (no. BK20131339), the Major Research Plan of Natural Science Foundation of the Jiangsu Higher Education Institutions of China (no. 15KJA110001), Qing Lan Project, PAPD, Program of Excellent Science and Technology Innovation Team of the Jiangsu Higher Education Institutions of China, Project of Construction for Superior Subjects of Statistics of Jiangsu Higher Education Institutions, and Project of the Key Lab of Financial Engineering of Jiangsu Province.

References

[1] M. Drmota, *Random Trees*, Springer, New York, NY, USA, 2009.

[2] L. Devroye and S. Janson, "Protected nodes and fringe subtrees in some random trees," *Electronic Communications in Probability*, vol. 19, article 6, 2014.

[3] Q. Feng and Z. Hu, "Phase changes in the topological indices of scale-free trees," *Journal of Applied Probability*, vol. 50, no. 2, pp. 516–532, 2013.

[4] N. Broutin and P. Flajolet, "The distribution of height and diameter in random non-plane binary trees," *Random Structures and Algorithms*, vol. 41, no. 2, pp. 215–252, 2012.

[5] F. Dennert and R. Grübel, "On the subtree size profile of binary search trees," *Combinatorics, Probability and Computing*, vol. 19, no. 4, pp. 561–578, 2010.

[6] H. M. Mahmoud, "One-sided variations on binary search trees," *Annals of the Institute of Statistical Mathematics*, vol. 55, no. 4, pp. 885–900, 2003.

[7] Q. Feng and Z. Hu, "On the Zagreb index of random recursive trees," *Journal of Applied Probability*, vol. 48, no. 4, pp. 1189–1196, 2011.

[8] R. Grübel and I. Michailow, "Random recursive trees: a boundary theory approach," *Electronic Journal of Probability*, vol. 20, article 37, 2015.

[9] H. M. Mahmoud and M. D. Ward, "Asymptotic properties of protected nodes in random recursive trees," *Journal of Applied Probability*, vol. 52, no. 1, pp. 290–297, 2015.

[10] M. Sibuya and Y. Itoh, "Random sequential bisection and its associated binary tree," *Annals of the Institute of Statistical Mathematics*, vol. 39, no. 1, pp. 69–84, 1987.

[11] H. Prodinger, "How to select a loser," *Discrete Mathematics*, vol. 120, no. 1–3, pp. 149–159, 1993.

[12] J. A. Fill, H. M. Mahmoud, and W. Szpankowski, "On the distribution for the duration of a randomized leader election algorithm," *The Annals of Applied Probability*, vol. 6, no. 4, pp. 1260–1283, 1996.

[13] Y. Itoh and H. M. Mahmoud, "One-sided variations on interval trees," *Journal of Applied Probability*, vol. 40, no. 3, pp. 654–670, 2003.

[14] S. Janson, "One-sided interval trees," *Journal of the Iranian Statistical Society*, vol. 3, pp. 149–164, 2004.

[15] S. Janson and R. Neininger, "The size of random fragmentation trees," *Probability Theory and Related Fields*, vol. 142, no. 3-4, pp. 399–442, 2008.

[16] M. Javanian, H. M. Mahmoud, and M. Vahidi-Asl, "Paths in m-ary interval trees," *Discrete Mathematics*, vol. 287, no. 1–3, pp. 45–53, 2004.

[17] C. Su, J. Liu, and Z. S. Hu, "The Laws of large numbers for the size of complete interval trees," *Advances in Mathematics*, vol. 36, no. 2, pp. 181–188, 2007.

[18] W. Pan, S.-K. Li, X. Cai, and L. Zeng, "The construction algorithm of binary interval tree nodes based on lattice partition," *Journal of Computer-Aided Design & Computer Graphics*, vol. 23, pp. 1115–1122, 2011.

[19] V. M. Zolotarev, *Modern Theory of Summation of Random Variables*, VSP, Utrecht, The Netherlands, 1997.

[20] V. M. Zolotarev, "Approximation of distributions of sums of independent random variables with values in infinite-dimensional spaces," *Theory of Probability and Its Applications*, vol. 21, no. 4, pp. 721–737, 1976.

[21] S. T. Rachev and L. Rüschendorf, "Probability metrics and recursive algorithms," *Advances in Applied Probability*, vol. 27, no. 3, pp. 770–799, 1995.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 268286, 9 pages
http://dx.doi.org/10.1155/2015/268286

Research Article

Decomposition and Decoupling Analysis of Energy-Related Carbon Emissions from China Manufacturing

Qingchun Liu, Shufang Liu, and Lingqun Kong

School of Economics, Shandong University of Finance and Economics, Jinan 250014, China

Correspondence should be addressed to Qingchun Liu; lqc7919@163.com

Received 25 July 2015; Accepted 3 September 2015

Academic Editor: Xinguang Zhang

The energy-related carbon emissions of China's manufacturing increased rapidly, from 36988.97×10^4 tC in 1996 to 74923.45×10^4 tC in 2012. To explore the factors to the change of the energy-related carbon emissions from manufacturing sector and the decoupling relationship between energy-related carbon emissions and economic growth, the empirical research was carried out based on the LMDI method and Tapio decoupling model. We found that the production scale contributed the most to the increase of the total carbon emissions, while the energy intensity was the most inhibiting factor. And the effects of the intrastructure and fuel mix on the change of carbon emissions were relatively weak. At a disaggregative level within manufacturing sector, EI subsector had a greater impact on the change of the total carbon emissions, with much more potentiality of energy conservation and emission reduction. Weak decoupling of manufacturing sector carbon emissions from GDP could be observed in the manufacturing sector and EI subsector, while strong decoupling state appeared in NEI subsector. Several advices were put forward, such as adjusting the fuel structure and optimizing the intrastructure and continuing to improve the energy intensity to realize the manufacturing sustainable development in low carbon pattern.

1. Introduction

Carbon emissions amount from China has already surpassed the United States since 2007 and has been the number one in the world [1]. Increasing trend of carbon emissions from China has received the great attention with the global warming. Both energy saving and emission reduction are becoming more and more important for Chinese government, who was committed to reducing carbon dioxide emissions per unit of GDP by 40–45% in 2020 to be less than 2005 levels on the Copenhagen Climate Change Conference in 2009. Manufacturing sector's, as the core of China economy, product value surpassed the United States for the first time in 2011 and became the number one in the world. But the development of China manufacturing sector has been depending on the high energy consumption for a long time, and it was responsible for approximately 82.54% of China's final energy demand in 1995–2006 [2]. Therefore, how to realize the manufacturing sustainable development in low carbon pattern is the biggest challenge for China in future. However, China's manufacturing sector has great energy saving potentiality and space due to excessive dependence on resources and energy consumption [3]. Thus, it is necessary to study the changes of energy-related carbon emissions over time and to explore the main driving factors to increase carbon emissions from China manufacturing sector, as well as the relationship between carbon emissions and economic growth in order to help meet the government target.

2. Literature Review

2.1. Decomposition Carbon Emissions. Decomposition analysis can divide the changes in carbon emissions over time into a number of different factors, which help us better understand the reasons for the changes observed. The broad technique of decomposition analysis undertaken here is often known as IDA, due to the advantage of its simplicity, the availability of statistical data, and the ease of historical comparison. There are a number of different methods available within IDA, and the Log Mean Divisia Index method I (LMDI I) is perfect in decomposition, having no residual term, which

is used here in decomposition carbon emissions due to the adaptability and ease of use [4, 5]. Using LMDI I method, many studies were carried out on carbon emissions from China in the last decade. Dong and Zhang [6] applied LMDI to decompose the energy-related carbon emissions from China industry into the production scale, carbon emissions intensity, energy structure, and energy intensity, and the findings showed that industrial production scale plays direct role in increasing carbon emissions. Wei and Xia [7] studied per-capita carbon emissions from the world and found that reducing the energy intensity and developing renewable clear energy were the only two major ways to realize low carbon economy under the background of increasing income and high energy dependence on coal. Song [8] also studied energy-related carbon emissions of Shandong Province in China and decomposed the carbon emissions into population, average wealth, intrastructure, and energy intensity, and pointed out that only the energy intensity had the negative effect on carbon emissions. Hammond and Norman [9] used the LMDI method to make decomposition analysis of energy-related carbon emissions from UK manufacturing sector and separated the contributions of changes in output, intrastructure, energy intensity, fuel mix, and electricity emission factor to the reduction in carbon emissions.

2.2. Decoupling Analysis between Carbon Emissions and Economic Growth. Many studies show the carbon emissions are strongly connected to the economic growth [10–13], and the target of low carbon economy is to sustain the decoupling state between the carbon emissions and economic growth. The concept of decoupling is first proposed by Organization for Economic Cooperation and Development (OECD), who divided it into absolute decoupling and relative decoupling [14]. Tapio [15] further expanded the decoupling theory, distinguished eight logical possibilities of decoupling, and studied the decoupling situations relationships between GDP, traffic volumes, and CO_2 emissions from transport in the EU15 countries in 1970–2001; Zhuang [16] applied Tapio model and studied the decoupling relationship between CO_2 emission and economic growth in Taiwan. Wang et al. [1] explored the decoupling relationship between energy-related carbon emissions and economic growth in Guangdong and found that its decoupling state turns from weak decoupling in 1996 to strong decoupling state in 2011. Li et al. [17] employed Tapio decoupling index to analyze the relationship between rural and urban construction land. Compared with OECD decoupling index, Tapio decoupling model has been widely used in empirical verification owing to the advantage of ease of adaptability, ease of use, and ease of understanding.

In summary, previous studies mainly focused on the carbon emissions from industry or the total carbon emissions amount in a certain nation or region, while few of them focused on the carbon emissions from manufacturing sector and the relationship between energy-related carbon emissions and economic growth in China [18]. Moreover, the changes of carbon emissions at disaggregative level lack analysis, but it is necessary for government to consider the differences of subsectors when making energy saving policies and measures. Therefore, the paper studies the changes of energy-related carbon emissions from China manufacturing over the years from 1996 to 2012, decomposes it in production scale, intrastructure, energy intensity and fuel mix, and carbon emissions coefficient by the means of LMDI, makes the decoupling analysis between energy-related carbon emissions and economic growth, and decomposes the total decoupling elasticity into energy conservation, carbon emissions reduction, output value, and industrial development. In addition, the manufacturing sector is split into two subsectors, energy-intensive (EI) subsector and non-energy-intensive (NEI) subsector; the carbon emissions of each subsector is decomposed, as well as the decoupling analysis.

3. Methods and Data Resource

3.1. Defining Energy-Intensive and Energy-Intensive Subsector. The manufacturing sector here is defined by Standard Industrial Classification (SIC) codes 1–30 according to *China Statistical Yearbook*, excluding the artwork and other subsectors (SIC code 29) and recycling subsector (SIC code 30) omitted for the limited data. The details of remaining 28 subsectors are shown in Table 1. Manufacturing sectors can be divided into EI and NEI subsectors, based on the extent of energy dependence and the potential strength of drivers to energy intensity improvement [9], and there are three criteria including the aggregate energy intensity of a subsector, the proportion of total financial costs represented by energy and water for a subsector, and the mean energy use per enterprise in a subsector. Given the limited data, the paper chooses the aggregate energy intensity value of 64.6 TJ/10^8 Yuan as the division criteria [9]. And five subsectors whose values are over the threshold are classified as the EI sector; they are manufacture of pulp, paper, and paper products (SCI code 10), manufacture of petroleum processing, coking, and nuclear fuel (SCI code 13), manufacture of chemicals and chemical products (SCI code 14), manufacture of other nonmetallic mineral products (SCI code 19), and manufacture of ferrous metal smelting and rolling (SCI code 20), and the remaining 23 subsectors are classified as the NEI sectors; all of them are labeled in Table 1. The average energy intensity in the EI subsector is 365.25 TJ/10^8 Yuan, which is 9.19 times of that in the NEI subsector.

3.2. Calculation of Energy-Related Carbon Emissions from Manufacturing Sector. Energy consumption includes the end-use energy consumption by manufacturing sector and energy consumption by production of thermal power and heat power. The paper just calculated the end-use energy consumption for the limited data. There are 16 types of energy mainly consumed in the manufacturing sector, including coal, crude oil, natural gas, and other fossil fuels from *Energy Balance Sheet of China Energy Statistical Yearbook*. Carbon emissions coefficient of each fuel could be calculated referenced by 2006 IPCC Guidelines for National Greenhouse Gas Inventories [19]. Based on these, the carbon emissions of 28 manufacturing subsectors could be calculated, as well as the total manufacturing sector.

Table 1: Subsector split of the manufacturing sector.

SCI code	Manufacturing subsector	EI/NEI
1	Manufacture of agricultural and sideline products	NEI
2	Manufacture of food products	NEI
3	Manufacture of beverages	NEI
4	Manufacture of tobacco products	NEI
5	Manufacture of textiles	NEI
6	Manufacture of garment, shoes, and hat	NEI
7	Manufacture of leather, fur, and feather	NEI
8	Manufacture of wood and wood products	NEI
9	Manufacture of furniture	NEI
10	Manufacture of pulp, paper, and paper products	EI
11	Publishing, printing, and reproduction of recorded media	NEI
12	Manufacture of stationery and sporting goods	NEI
13	Manufacture of petroleum processing, coking, and nuclear fuel	EI
14	Manufacture of chemicals, chemical products	EI
15	Manufacture of pharmaceutical	NEI
16	Manufacture of chemicals fibers	NEI
17	Manufacture of rubber	NEI
18	Manufacture of plastic products	NEI
19	Manufacture of other nonmetallic mineral products	EI
20	Manufacture of ferrous metal smelting and rolling	EI
21	Manufacture of nonferrous metal smelting and rolling	NEI
22	Manufacture of mineral product	NEI
23	Manufacture of general equipment	NEI
24	Manufacture of special equipment	NEI
25	Manufacture of transport equipment	NEI
26	Manufacture of electrical machinery	NEI
27	Manufacture of communication equipment computer and other electronic equipment	NEI
28	Manufacture of instrument and office machinery	NEI

3.3. Data Source and Processing. The energy data used in this paper are derived from *Energy Balance Sheet of China Energy Statistical Yearbook* (1996–2013). Other data come from the *Statistical Yearbook of China* (1996–2013). To eliminate the effect of price changes, we converted the GDP at current price to the GDP at constant price in 1995 with index reduction method, and the price indexes come from *China Statistical Yearbook*.

3.4. Decomposition Carbon Emissions Based on LMDI Model. According to LMDI model, the total change in carbon emissions over the period (0 to T) (ΔC_T) is the sum of the changes, including the changes in production scale (ΔC_p), the changes in intrastructure (ΔC_s), the changes in energy intensity (ΔC_i), the changes in fuel mix (ΔC_m), and the changes in carbon emissions coefficient (ΔC_e). Carbon emissions coefficients of different basic fuels are approximately constant in China in the actual application; therefore $\Delta C_e = 0$.

The total change in carbon emissions can be expressed as

$$\Delta C_T = \Delta C_p + \Delta C_s + \Delta C_i + \Delta C_m. \tag{1}$$

For i manufacturing subsector consuming j fuels, the total carbon emissions are given by

$$C = \sum_{ij} C_{ij} = \sum_{ij} P \frac{P_i}{P} \frac{E_i}{P_i} \frac{E_{ij}}{E_i} \frac{C_{ij}}{E_{ij}} = \sum_{ij} P S_i I_i M_{ij} U_{ij}, \tag{2}$$

where P is the output of manufacturing sector, P_i is the output of subsector i, E_i is the energy consumption of subsector i, E_{ij} is the consumption of fuel j in subsector i, C_{ij} is the carbon emissions of fuel j in subsector i, $S_i = (P_i/P)$ is the output share occupied by subsector i, and $I_i = (E_i/P_i)$ is the energy intensity of subsector i. And $M_{ij} = (E_{ij}/E_i)$ represents the proportion of energy in subsector i supplied by fuel j and $U_{ij} = (C_{ij}/E_{ij})$ is carbon emissions coefficient factor of fuel j in subsector i.

The components of change in (1) are the following:

$$\begin{aligned}
\Delta C_p &= \sum_{ij} L\left(C_{ij}^{t-1}, C_{ij}^{t}\right) \ln\left[\frac{P(t)}{P(t-1)}\right] \\
\Delta C_s &= \sum_{ij} L\left(C_{ij}^{t-1}, C_{ij}^{t}\right) \ln\left[\frac{S_i(t)}{S_i(t-1)}\right] \\
\Delta C_i &= \sum_{ij} L\left(C_{ij}^{t-1}, C_{ij}^{t}\right) \ln\left[\frac{I_{ij}(t)}{I_{ij}(t-1)}\right] \\
\Delta C_m &= \sum_{ij} L\left(C_{ij}^{t-1}, C_{ij}^{t}\right) \ln\left[\frac{M_{ij}(t)}{M_{ij}(t-1)}\right],
\end{aligned} \tag{3}$$

where

$$L\left(C_{ij}^{t-1}, C_{ij}^{t}\right) = \begin{cases} \dfrac{C_{ij}^{t} - C_{ij}^{t-1}}{\ln\left(C_{ij}^{t}/C_{ij}^{t-1}\right)} & \left(C_{ij}^{t-1} \neq C_{ij}^{t}\right) \\ C_{ij}^{t-1} \text{ or } C_{ij}^{t} & \left(C_{ij}^{t-1} = C_{ij}^{t}\right). \end{cases} \tag{4}$$

Equation (3) denotes the production effect, intrastructure effect, energy intensity effect, and fuel mix effect, respectively.

Table 2: Eight decoupling states divided by Tapio [15].

Decoupling states	$\Delta C/C$	ΔGDP/GDP	Decoupling elasticity values (D)
Negative decoupling			
Expansive negative decoupling	>0	>0	$D > 1.2$
Strong negative decoupling	>0	<0	$D < 0$
Weak negative decoupling	<0	<0	$0 < D < 0.8$
Decoupling			
Weak decoupling	>0	>0	$0 < D < 0.8$
Strong decoupling	<0	>0	$D < 0$
Recessive decoupling	<0	<0	$D > 1.2$
Coupling			
Expansive coupling	>0	>0	$0.8 < D < 1.2$
Recessive coupling	<0	<0	$0.8 < D < 1.2$

To measure the effect contribution of each factor, we define them as follows:

$$g_p = \frac{\Delta C_p}{\Delta C}, \quad g_s = \frac{\Delta C_s}{\Delta C}, \quad g_i = \frac{\Delta C_i}{\Delta C}, \quad g_m = \frac{\Delta C_m}{\Delta C}, \tag{5}$$

where g_p, g_s, g_i, and g_m indicate the effect contribution values of production scale, intrastructure, energy intensity, and fuel mix, respectively.

3.5. Decoupling Model between Carbon Emissions and Economic Growth. In order to study the decoupling relation between energy relationship between carbon emissions and economic growth further, the decoupling model could be decomposed into five elasticity values, and the model is given as follows:

$$\begin{aligned} D_i &= \frac{\Delta C_i/C_i}{\Delta \mathrm{GDP}/\mathrm{GDP}} \\ &= \frac{\Delta C_i/C_i}{\Delta E_i/E_i} \cdot \frac{\Delta E_i/E_i}{\Delta P_i/P_i} \cdot \frac{\Delta P_i/P_i}{\Delta \mathrm{IGDP}/\mathrm{IGDP}} \\ &\quad \cdot \frac{\mathrm{IGDP}/\mathrm{IGDP}}{\Delta \mathrm{GDP}/\mathrm{GDP}}, \end{aligned} \tag{6}$$

where D_i denotes the decoupling elasticity between carbon emissions from subsector i and economic growth, C is the carbon emissions from subsector i at base year, GDP and IGDP are the gross domestic product and the secondary industry product value at base year, and E_i and P_i denote energy consumption and output value of subsector i, respectively. ΔC_i, E_i, and ΔP_i denote increment of carbon emissions, energy consumption, output value of subsector i, respectively, and ΔIGDP and ΔGDP denote the increment of secondary industry product value and the gross domestic product.

$(\Delta C_i/C_i)/(\Delta E_i/E_i)$ is the carbon reduction decoupling elasticity, reflecting the elasticity between carbon emissions of subsector i and energy consumption, $(\Delta E_i/E_i)/(\Delta P_i/P_i)$ is the energy conservation decoupling elasticity, describing the elasticity between energy consumption of subsector i and manufacturing output value, $(\Delta P_i/P_i)/(\Delta \mathrm{IGDP}/\mathrm{IGDP})$ is the output value decoupling elasticity, describing the elasticity between product value of subsector i and that of the secondary industry, and $(\Delta \mathrm{IGDP}/\mathrm{IGDP})/(\Delta \mathrm{GDP}/\mathrm{GDP})$ is the industrial development decoupling elasticity, describing the elasticity between the secondary industry product value and the GDP. Additionally, the decoupling states defined by Tapio are presented in Table 2 according to the decoupling elasticity value.

4. Results and Discussion

4.1. Changes of Energy-Related Carbon Emissions. As is shown in Figure 1 and Table 3, the energy-related carbon emissions of China manufacturing increased rapidly from 36988.97×10^4 tons of carbon in 1996 to 74923.45×10^4 tons of carbon in 2012 by 6.41% per annum and present obvious change during the different period. The average carbon emissions of EI subsector from 1996 to 2012 account for 83% of the total amount of carbon emissions from manufacturing sector, but its output value only takes up 28% of total manufacturing sector. The carbon emissions of NEI subsector change little during the period, while both the carbon emissions from EI subsector and the total manufacturing sector fluctuate from 1996 to 2012.

(1) Low carbon emissions and falling trend period (1996–2001): the energy consumption from manufacturing sector and EI subsector in this period decreased gradually, showing falling trend with the year changing. Mainly because of the financial crisis from Asia in 1997, China economy, especially the manufacturing, was affected, and the slow growth rate of manufacturing led to the stable low level of carbon emissions in the whole country.

(2) Rapid growth period (2002–2012): the feature of carbon emissions from manufacturing sector and EI subsector in this period was much higher increasing speed and much bigger amount of carbon emissions. Except the fall in 2010,

Table 3: The proportion of output value and carbon emissions of EI subsector, NEI subsector to the total amount of manufacturing sector in 1996–2012.

Year	Output value proportion		Carbon emissions proportion	
	EI	NEI	EI	NEI
1996	28%	72%	78%	22%
1997	28%	72%	79%	21%
1998	26%	74%	79%	21%
1999	26%	74%	78%	22%
2000	27%	73%	80%	20%
2001	27%	73%	80%	20%
2002	26%	74%	80%	20%
2003	27%	73%	82%	18%
2004	29%	71%	84%	16%
2005	29%	71%	85%	15%
2006	29%	71%	86%	14%
2007	29%	71%	86%	14%
2008	30%	70%	86%	14%
2009	28%	72%	87%	13%
2010	29%	71%	87%	13%
2011	30%	70%	88%	12%
2012	29%	71%	89%	11%
Average	28%	72%	83%	17%

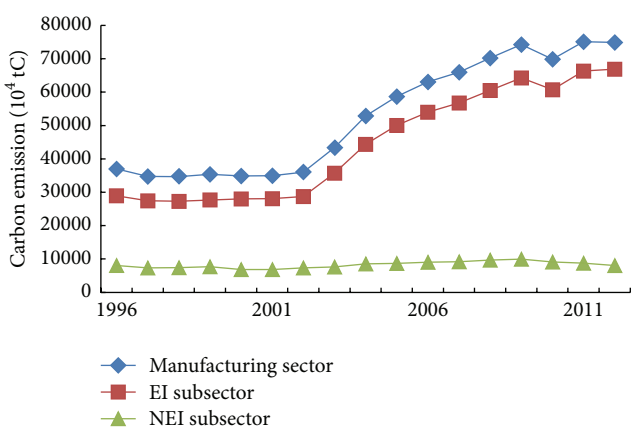


Figure 1: Change trend of carbon emissions from China's manufacturing sector and subsector in 1996–2012.

the total emission amount showed significant upward trend. Because manufacturing sector developed rapidly, the dependence on high energy consumption brought about the rapid growth of carbon emissions, after China joined WTO in 2001. Although China has been advocating transforming economic growth pattern, adjusting and optimizing the intrastructure since 2006, the carbon emissions amount has not changed correspondingly. The policy of energy conservation and reduction still faces great pressure in future to achieve the government targets in 2020.

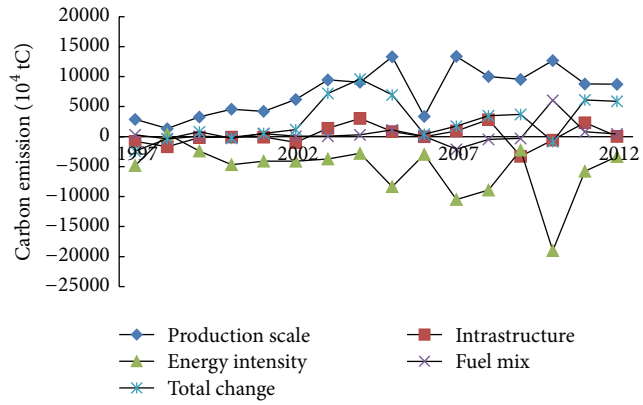


Figure 2: Annual effect of each factor on carbon emissions in the China manufacturing sector, 1997–2012.

Table 4: Cumulative effect contribution value of each factor on carbon emissions in the China manufacturing sector, EI, and NEI in 2012.

Factor	Manufacturing sector	EI	NEI
Production scale	2.672	2.46	3.205
Intrastructure	0.087	0.005	−0.194
Energy intensity	−1.994	−1.608	−2.158
Fuel mix	0.145	0.142	0.147

4.2. Decomposition Analysis of Carbon Emissions. The annual effect of the production scale, intrastructure, energy intensity, and fuel mix on carbon emissions from manufacturing sector is shown in Figure 2. Taking 1996 as the base year, the energy intensity effect was always negative and the production scale effect was positive; both intrastructure and fuel mix had slight changes over the study period. In terms of cumulative effect contribution value of each factor in 2012, as shown in Table 4, the value of production scale was 276.2% and made the largest contribution to the changes of carbon emissions among all factors, indicating the production scale had positive effect on the increase of carbon emissions, energy intensity was an important inhibition factor of increasing carbon emissions, and its cumulative effect contribution value reached −199.3%. The results showed that although energy intensity decreases, the expansion of production from manufacturing sector could result in the increase of carbon emissions.

While the value of intrastructure and fuel mix was 8.6% and 14.4%, respectively, both the change of intrastructure and fuel mix had no significant impact on curbing carbon emissions of manufacturing sector. Manufacturing sector consumed the coal for approximately over 50% of the total energy consumption during the period, and the coal was still the main fuel source of the energy consumption. Besides, within manufacturing sector, the share of each subsector did not change significantly; the output proportion of EI is still too high with few changes. Thus, both fuel structure and intrastructure of manufacturing sector need to be further optimized in order to curb the increase of carbon emissions in future.

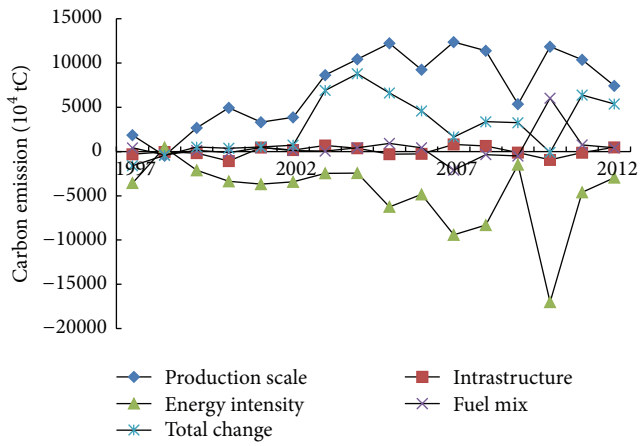


Figure 3: Annual effect of each factor on carbon emissions in EI subsector, 1997–2012.

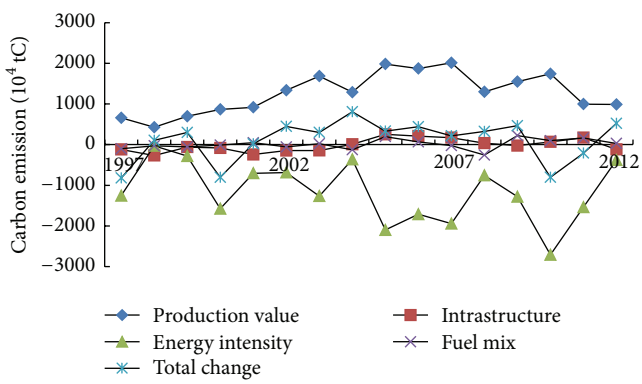


Figure 4: Annual effect of each factor on carbon emissions in NEI subsectors, 1997–2012.

Figures 3 and 4 represent the annual effect of each factor on the changes of carbon emissions from EI and NEI subsectors, respectively. The results showed that the effect value of EI subsector was higher than that of NEI subsector. For each subsector, both production scale and energy intensity had relatively great impact on the changes of carbon emissions during the study period. The effect of production scale on the changes of carbon emissions was positive in most of years during the study period, while energy intensity effect was negative. In comparison, both intrastructure and fuel mix had relative smaller impacts; their cumulative effect contribution values were small, showing relatively weak pushing effects on carbon emissions.

For EI subsector, in terms of cumulative effect contribution values in 2012, production scale made the largest positive contribution to the changes of carbon emissions and the value reached 246%, energy intensity is the second factor, and its contribution value reached −160.8%, while the value of intrastructure and fuel mix was 0.5% and 14.2%, respectively. As to NEI subsector, all the absolute cumulative effects of contribution value of factors except fuel mix were relatively higher than EI subsector, whose contribution value of the production scale, intrastructure, energy intensity, and fuel mix was 320.5%, −19.4%, −215.8%, and 14.7%, respectively.

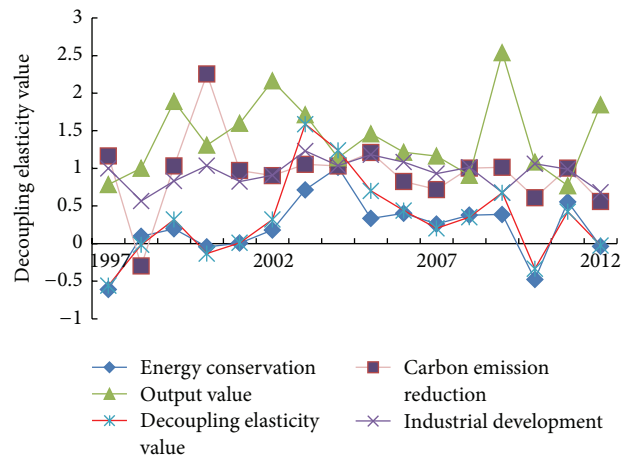


Figure 5: Change trends of decoupling elasticity values between energy-related carbon emissions and economic growth from manufacturing sector, 1997–2012.

4.3. Analysis of Decoupling Elasticity. In order to analyze the relationship between carbon emissions of manufacturing sector and economic growth, the paper calculated the decoupling elasticity value according to (6). And five decoupling elasticity values of manufacturing sector and two subsectors from 1996 to 2012 could be obtained. As to the manufacturing sector, the total decoupling elasticity value was 0.163, expressing the weak decoupling state. Comparing the four decoupling elasticity values, the output value had the biggest value, which was 1.819 presenting the expansive negative decoupling state between the output value of manufacturing sector and that of the secondary industry, while the energy conservation had the smallest value, which was 0.091 expressing the weak decoupling state between the energy consumption and the output value of manufacturing sector. For two subsectors, the weak decoupling state occurred in EI subsector, its total decoupling elasticity value was 0.208, and the strong decoupling state appeared in NEI subsector, and its total decoupling elasticity value was −0.0003.

The annual changes of all decoupling elasticity values from manufacturing sector, EI subsector, and NEI subsector were shown in Figures 5, 6, and 7, respectively. From the change trends of all the decoupling elasticity values during the study period, the total decoupling elasticity of carbon emissions from economic growth had the same trend as the energy conservation decoupling elasticity, it increased between 1997 and 2003 but showed a clear decline after 2003, and the decoupling state transformed from the weak decoupling state to strong decoupling state in 2012.

Decoupling elasticity of energy conservation describes the elasticity of manufacturing output value from energy consumption. For manufacturing sector, it had the same trend as the total decoupling elasticity and presented the weak or the strong decoupling state during the study period, indicating the effective policy of energy conservation in

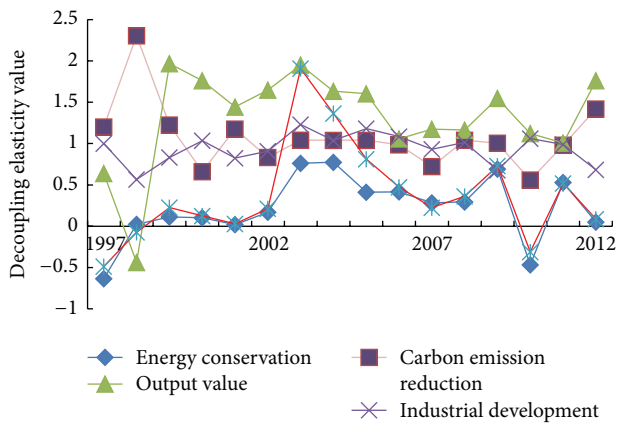


Figure 6: Change trends of each of the decoupling elasticity values between energy-related carbon emissions and economic growth from EI subsectors in 1997–2012.

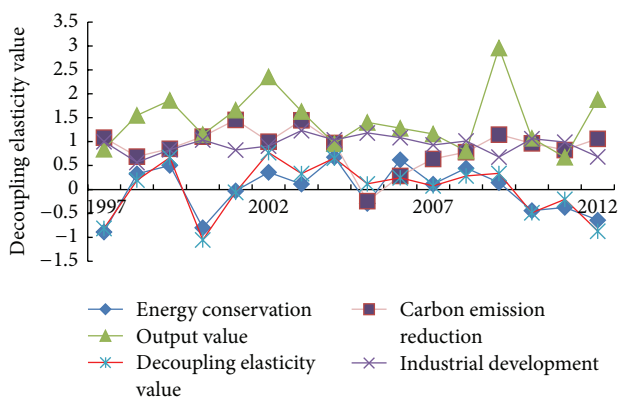


Figure 7: Change trends of each of the decoupling elasticity values between energy-related carbon emissions and economic growth from NEI subsectors in 1997–2012.

China owing to falling energy consumption per unit of output. Analyzing EI and NEI subsector, respectively, both two subsectors had slight change, and the state of weak decoupling often appeared in EI subsector almost every year, but the strong decoupling often occurred in NEI subsector. Either weak decoupling or strong decoupling state occurred in manufacturing sector during the study period except in 2003 and 2004, showing significant decoupling effects.

Decoupling elasticity of carbon emission reduction reflects the energy consumption elasticity of carbon emissions. For China manufacturing, it increased between 1997 and 2000 but began to decline after 2000, and the expansive coupling state appeared many times during the study period, which indicated that the policy of carbon emissions reduction in China manufacturing had little effect. Analyzing two subsectors, respectively, the expansive coupling state mainly appeared in EI subsector, while recessive coupling state mainly occurred in NEI subsector during the study period.

Decoupling elasticity of output value describes the secondary industry value of elasticity of the manufacturing sector output value. For the manufacturing sector, it expressed the state of expansive coupling or expansive negative decoupling from 1997 to 2012, indicating the production value of manufacturing sector increased more than the secondary industry. Like the total manufacturing sector, both EI and NEI subsectors were also in the state of expansive coupling state or expansive negative decoupling during the study period, but the growth rate of EI sector was lower than that of NEI subsector.

Decoupling elasticity of industrial development describes the GDP elasticity of the secondary industry output value. Either the state of expansive coupling or expansive negative decoupling mainly appeared in 1997–2012, indicating the growth rate of the secondary industry increased more than GDP. Although our government has been encouraging the development of the third industry, the secondary industry took up higher proportion of GDP, and it was 45.3% in 2012.

5. Conclusions and Implications

Energy-related carbon emissions from 1996 to 2012 were calculated from 28 subsectors of China manufacturing consuming 16 types of fuels, and the results indicated that carbon emissions from China manufacturing sector have been increasing rapidly since 2002. The paper makes the analysis of the reasons of the increase of carbon emissions, and the relationship between carbon emissions and economic growth. The conclusions are as follows:

(i) Production scale was the major positive contribution factor affecting energy-related carbon emissions, which had closely relationships with the expansion of manufacturing output in recent years. However, curbing the product of manufacturing is not the feasible approach to decrease the carbon emissions for the economic growth. The way to solve the issue in the long run is to develop low carbon economy and decrease energy intensity to improve energy efficiency to coordinate the carbon emissions and economic development.

(ii) Energy intensity always had strong inhibiting effects on the increase of carbon emissions, showing downward trend during the study periods. This is connected with the energy conservation and emission reduction advocated in recent years, and particularly since the 11th Five-Year Plan period, China has been developing environment-friendly industries and the new low carbon industries, strengthening the assessment of energy conservation and the elimination of backward production capacity. So we should continue strengthening the policies of energy conservation and emission reduction, and technological innovations, in order to promote the industry upgrading and improve the energy use efficiency.

(iii) Intrastructure had weak positive influences on carbon emissions, which showed irrational structure existing within manufacturing sector, too many extensive subsectors with high energy consumption but low output. The policies of industrial structure adjustment

advocated in China in recent years have little effects on reducing carbon emissions. So keeping optimizing intrastructure and converting extensive subsectors into intensive ones should be taken as main strategies to drive intrastructure adjustments.

(iv) Compared with the other factors, the effect of fuel mix on carbon emissions was relatively weak, which was related with the long-term dependence on coal and petrol consumption. So the government should promote energy structure adjustment and develop reproducible energy such as nuclear energy, water energy, wind energy, solar energy, and bioenergy, to optimize fuel structure and decrease the dependence of energy consumption on fossil energy in future.

(v) Only 5 subsectors in the 28 subsectors of manufacturing were classified as EI subsector, which accounted for 28% of total manufacturing output value but contributed 83% to the total carbon emissions of manufacturing sector. Both EI and NEI subsectors had two same factors of affecting emission change, production scale, and energy intensity, while the effect values of the two factors in EI subsector were much higher than that in NEI subsector, but the cumulative effect contribution values were relatively lower. Therefore, the EI subsector still has much more potentiality of energy conservation and emission reduction than NEI subsector.

(vi) Weak decoupling of manufacturing sector carbon emissions from GDP could be observed during the study period. Weak decoupling state in EI subsector and strong decoupling state in NEI subsector were observed in most years of the study period. In order to sustain the decoupling relationship between carbon emissions of manufacturing sector and GDP, our government need to strengthen the decoupling elasticity of energy conservation and carbon emissions reduction.

There are some limitations in the paper. Carbon emissions are only involved in the final energy consumptions without considering the production process which also generates carbon emissions. In addition, carbon emissions from manufacturing sector at the provincial level should be studied so as to establish appropriate and specific low carbon policies and measures for the large difference of province, which is the direction of the research in future.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors gratefully acknowledge the Natural Science Foundation of China for Grant Support (no. 41401163) and Shandong Province Soft Science for Grant Support (no. 2014RZB01009, no. 2013RZB01023).

References

[1] W.-X. Wang, Y.-Q. Kuang, N. Huang, and D. Zhao, "Empirical research on decoupling relationship between energy-related carbon emission and economic growth in Guangdong province based on extended kaya identity," *The Scientific World Journal*, vol. 2014, Article ID 782750, 11 pages, 2014.

[2] Z.-W. Gong and J. Wu, "The energy consumption and it's intensity forecast of China's manufacturing industry," in *Proceedings of the 4th Symposium of Chinese Scientifically and Technological Policies and the Annual Conference of 4th Science and Technology Policy Studies of China (II)*, October 2008.

[3] B. W. Ang, "Decomposition analysis for policymaking in energy: which is the preferred method?" *Energy Policy*, vol. 32, no. 9, pp. 1131–1139, 2004.

[4] L. F. Wang, W. J. Duan, M. Y. Lai et al., "Study on the regional differences and industry differences of energy-saving potentiality in China's manufacturing industry," *Geographical Research*, vol. 34, no. 1, pp. 101–121, 2015.

[5] C. Liu, "An overview for decomposition of industry energy consumption," *American Journal of Applied Science*, vol. 2, no. 7, pp. 1166–11687, 2005.

[6] J. Dong and X. Zhang, "Decomposition of carbon emissions and low carbon strategies for industrial sector energy consumption in China," *Resources Science*, vol. 32, no. 10, pp. 1856–1862, 2010.

[7] C. Wei and D. Xia, "A decomposition analysis of per capita CO_2 emission of China: based on cross-country comparison," *Management Review*, vol. 22, no. 8, pp. 1856–1862, 2010.

[8] J.-K. Song, "Factor decomposition of carbon emissions from energy consumption of Shandong Province based on LMDI," *Resources Science*, vol. 34, no. 3, pp. 35–41, 2012.

[9] G. P. Hammond and J. B. Norman, "Decomposition analysis of energy-related carbon emissions from UK manufacturing," *Energy*, vol. 41, no. 1, pp. 220–227, 2012.

[10] M. E. H. Arouri, A. Ben Youssef, H. M'henni, and C. Rault, "Energy consumption, economic growth and CO_2 emissions in Middle East and North African countries," *Energy Policy*, vol. 45, pp. 342–349, 2012.

[11] B. Saboori and J. Sulaiman, "Environmental degradation, economic growth and energy consumption: evidence of the environmental Kuznets curve in Malaysia," *Energy Policy*, vol. 60, pp. 892–905, 2013.

[12] S. Niu, Y. Ding, Y. Niu, Y. Li, and G. Luo, "Economic growth, energy conservation and emissions reduction: a comparative analysis based on panel data for 8 Asian-Pacific countries," *Energy Policy*, vol. 39, no. 4, pp. 2121–2131, 2011.

[13] K. Alkhathlan and M. Javid, "Energy consumption, carbon emissions and economic growth in saudi arabia: an aggregate and disaggregate analysis," *Energy Policy*, vol. 62, pp. 1525–1532, 2013.

[14] OECD, "Indicators to measure decoupling of environmental pressures from economic growth," Tech. Rep., OECD, Paris, France, 2002.

[15] P. Tapio, "Towards a theory of decoupling: degrees of decoupling in the EU and the case of road traffic in Finland between 1970 and 2001," *Transport Policy*, vol. 12, no. 2, pp. 137–151, 2005.

[16] M. F. Zhuang, *Decoupling index and evaluation of industrial and transportation departments in Taiwan [Ph.D. thesis]*, National Taipei University, Taipei, Taiwan, 2006.

[17] X.-S. Li, F.-T. Qu, and Z.-X. Guo, "Decoupling between urban and rural construction land," *China Population, Resources and Environment*, vol. 18, no. 5, pp. 179–184, 2008.

[18] X. Pan, F.-S. Tao, and D.-W. Xu, "On the changes in the carbon emissions intensity of China's manufacturing industry and its factors decomposition," *China Population, Resources and Environment*, vol. 21, no. 5, pp. 101–105, 2011.

[19] Intergovernmental Panel on Climate Change (IPCC), "IPCC guidelines for national greenhouse gas inventories," 2006, http://www.ipcc-nggip.iges.or.jp.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 917137, 7 pages
http://dx.doi.org/10.1155/2015/917137

Research Article

Chaotification for a Class of Delay Difference Equations Based on Snap-Back Repellers

Zongcheng Li,[1,2] Shutang Liu,[1] Wei Li,[3] and Qingli Zhao[2]

[1] *College of Control Science and Engineering, Shandong University, Jinan, Shandong 250061, China*
[2] *School of Science, Shandong Jianzhu University, Jinan, Shandong 250101, China*
[3] *Department of Public Foundation, Shandong Radio and TV University, Jinan, Shandong 250010, China*

Correspondence should be addressed to Zongcheng Li; chengzi_0905@163.com

Received 29 April 2015; Accepted 12 July 2015

Academic Editor: Xinguang Zhang

We study the chaotification problem for a class of delay difference equations by using the snap-back repeller theory and the feedback control approach. We first study the stability and expansion of fixed points and establish a criterion of chaos. Then, based on this criterion of chaos and the feedback control approach, we establish a chaotification scheme such that the controlled system is chaotic in the sense of both Devaney and Li-Yorke when the parameters of this system satisfy some mild conditions. For illustrating the theoretical result, we give some computer simulations.

1. Introduction

Research on chaos control has attracted a lot of interest from many scientists and mathematicians. There are two directions in chaos control, that is, control of chaos and anticontrol of chaos (or called chaotification). The former regarded chaos as harmful. So many earlier works focused on stabilizing a chaotic system, which was regarded as the traditional control. The reader is referred to the monographs [1–3] for more details. However, in recent years, it has been found that chaos can actually be very useful in some applications; a typical example is chaos-based cryptography [4]. Hence, sometimes it is useful and even important to make a nonchaotic system chaotic, or to make a chaotic system produce a stronger or different type of chaos. This progress is called chaotification or anticontrol of chaos.

In research on chaotification for discrete dynamical systems, a mathematically rigorous and effective chaotification method was first proposed by Chen and Lai [5–7], where they first used the feedback control technique. This method plays an important role in studying chaotification problems of discrete dynamical systems. For a survey on chaotification of discrete dynamical systems, one can see [8] and some references therein.

To the best of our knowledge, although there already exist many works on chaotification of discrete dynamical systems, there are few results on chaotification of delay difference equations. Motivated by the feedback control approach, we have succeeded in studying the chaotification problems on linear delay difference equations [9] and a class of delay difference equations [10]. In the two papers, we use the sine functions as controllers to establish some chaotification schemes. The reason of using this type of controllers is that the sine function has some favorable properties and this designed controller is also simple, cheap, and implementable in real engineering applications (see [8–10] and the references therein). In the chaotification theorem of [10], the delay difference equations need to have at least two fixed points. However, there are also many delay difference equations with only one fixed point, which cannot satisfy the above condition. This motivates us to study this case. In this paper, we will apply the feedback control approach and the snap-back repeller theory to study chaotification for a class of delay difference equations with at least one fixed point.

This paper is organized as follows. In Section 2, we give some basic concepts and one lemma. In Section 3, we study the stability and expansion of fixed points and establish a criterion of chaos. Based on this criterion of chaos, we

establish a chaotification scheme for a class of delay difference equations with at least one fixed point. Then, we give some computer simulations to illustrate the theoretical result. Finally, we conclude this paper in Section 4.

2. Preliminaries

Up to now, there is no unified definition of chaos in mathematics. For convenience, we present two definitions of chaos, which will be used in this paper.

Definition 1 (see [11]). Let (X, d) be a metric space, let $F : X \to X$ be a map, and let S be a set of X with at least two distinct points. Then S is called a scrambled set of F if, for any two different points $x, y \in S$,

$$\begin{aligned} \liminf_{n\to\infty} d\left(F^n(x), F^n(y)\right) &= 0, \\ \limsup_{n\to\infty} d\left(F^n(x), F^n(y)\right) &> 0. \end{aligned} \tag{1}$$

The map F is said to be chaotic in the sense of Li-Yorke if there exists an uncountable scrambled set S of F.

Remark 2. The term "chaos" was first used by Li and Yorke [12] for a map on a compact interval. Following the work of Li and Yorke, Zhou [11] gave the above definition of chaos for a topological dynamical system on a general metric space.

Definition 3 (see [13]). Let (X, d) be a metric space. A map $F : V \subset X \to V$ is said to be chaotic on V in the sense of Devaney if

(i) F is topologically transitive in V;

(ii) the periodic points of F are dense in V;

(iii) F has sensitive dependence on initial conditions in V.

Remark 4. In [14], Huang and Ye showed that chaos in the sense of Devaney is stronger than that in the sense of Li-Yorke under some conditions.

The following criterion of chaos is established by Shi et al., which plays an important role in the present paper.

Lemma 5 (see [15, Theorem 2.1]; [16, Theorem 4.4]). *Let $F : \mathbf{R}^n \to \mathbf{R}^n$ be a map with a fixed point $z \in \mathbf{R}^n$. Assume that*

(i) *F is continuously differentiable in a neighborhood of z and all the eigenvalues of $DF(z)$ have absolute values larger than 1, which implies that there exist a positive constant r and a norm $\|\cdot\|$ in $\mathbf{R}^n$ such that F is expanding in $\overline{B}_r(z)$ in $\|\cdot\|$, where $\overline{B}_r(z)$ is the closed ball of radius r centered at z in $(\mathbf{R}^n, \|\cdot\|)$;*

(ii) *z is a snap-back repeller of F with $F^m(x_0) = z$, $x_0 \neq z$, for some $x_0 \in B_r(z)$ and some positive integer m, where $B_r(z)$ is the open ball of radius r centered at z in $(\mathbf{R}^n, \|\cdot\|)$. Furthermore, F is continuously differentiable in some neighborhoods of $x_0, x_1, \ldots, x_{m-1}$, respectively, and $\det DF(x_j) \neq 0$ for $0 \le j \le m-1$, where $x_j = F(x_{j-1})$ for $1 \le j \le m-1$.*

Then for each neighborhood U of z, there exist a positive integer $k > m$ and a Cantor set $\Lambda \subset U$ such that $F^k : \Lambda \to \Lambda$ is topologically conjugate to the symbolic dynamical system $\sigma : \sum_2^+ \to \sum_2^+$. Consequently, there exists a compact and perfect invariant set $V \subset \mathbf{R}^n$, containing the Cantor set Λ, such that F is chaotic on V in the sense of Devaney as well as in the sense of Li-Yorke and has a dense orbit in V.

Remark 6. In 1978, Marotto [17] first gave the concept of snap-back repeller for maps in $\mathbf{R}^n$. Later, in 2004, Shi and Chen [18] extended this concept to general metric spaces. According to the classifications of snap-back repellers for maps in metric spaces in [18], the snap-back repeller given by Marotto [17] is regular and nondegenerate. For more details on snap-back repeller, we refer to [16–19] and the references therein. We can easily conclude that the point z in Lemma 5 is a regular and nondegenerate snap-back repeller. Hence, Lemma 5 can be summed as a single word: "a regular and nondegenerate snap-back repeller in $\mathbf{R}^n$ implies chaos in the sense of both Devaney and Li-Yorke." For more details, one can see [15, 16].

3. Chaotification Based on Snap-Back Repellers

In this paper, we will study the chaotification problem of a delay difference equation, chaotic or not, in the form of

$$x(n+1) = f(x(n-k), x(n)), \quad n \ge 0, \tag{2}$$

where $k \ge 1$ is a fixed integer and $f : D \subset \mathbf{R}^2 \to \mathbf{R}$ is a map. Equation (2) is a discrete analogue of many one-dimensional delay differential equations, such as the well known Mackey-Glass equation.

The objective here is to design a control input sequence $\{v(n)\}$ such that the output of the controlled system

$$x(n+1) = f(x(n-k), x(n)) + v(n), \quad n \ge 0, \tag{3}$$

is chaotic in the sense of both Devaney and Li-Yorke. In our earlier paper [10], by using the result that heteroclinic cycles connecting repellers imply chaos established in [20], we have studied the chaotification problem of (2) for the case where (2) has at least two fixed points. However, there are also many delay discrete dynamical systems which only have one fixed point. Then, the chaotification scheme established in [10] cannot be used. In this paper, we will study the chaotification problem for the case where (2) has at least one fixed point. We design the controller as follows:

$$v(n) = \alpha \operatorname{saw}_\varepsilon(\beta x(n-k)), \tag{4}$$

where $\varepsilon > 0$ is any given constant, α and β are two undetermined parameters, and $\operatorname{saw}_\varepsilon(\cdot)$ is the classical sawtooth function; that is,

$$\begin{aligned} \operatorname{saw}_\varepsilon(x) &= (-1)^m (x - 2m\varepsilon), \\ &\qquad (2m-1)\varepsilon \le x < (2m+1)\varepsilon, \ m \in \mathbf{Z}, \end{aligned} \tag{5}$$

while $\mathbf{Z}$ denotes the integer set. Many researchers have succeeded in using the sawtooth function as a controller

to chaotify discrete dynamical systems (see [15, 21] and the references therein).

Set

$$u_j(n) := x(n+j-k-1), \quad 1 \le j \le k+1, \ n \ge 0. \tag{6}$$

Then (2) and the controlled system (3) with controller (4) can be transformed into the following $k+1$-dimensional discrete systems on $\mathbf{R}^{k+1}$:

$$u(n+1) = F(u(n)), \tag{7}$$

$$u(n+1) = G(u(n)), \tag{8}$$

respectively, where $u = (u_1, u_2, \ldots, u_{k+1})^T \in \mathbf{R}^{k+1}$ and the maps $F, G : \mathbf{R}^{k+1} \to \mathbf{R}^{k+1}$.

As defined in [10], the maps F and G are called the maps induced by f and g, respectively, where $g(x,y) := f(x,y) + \alpha \mathrm{saw}_\varepsilon(\beta x)$. Systems (7) and (8) are called the systems induced by (2) and (3) in the Euclidean space $\mathbf{R}^{k+1}$, respectively. System (3) is said to be chaotic in the sense of Devaney (or Li-Yorke) on $V \subset \mathbf{R}^{k+1}$ if its induced system (8) is chaotic in the sense of Devaney (or Li-Yorke) on $V \subset \mathbf{R}^{k+1}$.

In the following, without loss of generality and for simplicity, we can suppose that the origin $O := (0, \ldots, 0)^T \in \mathbf{R}^{k+1}$ is always a fixed point of the induced system (7). Otherwise, if none of the fixed points is the origin O, then we can choose a transformation of coordinates such that one of the fixed points becomes the origin O in a new coordinate system. Then the map f in (2) satisfies $f(0,0) = 0$, throughout the rest of the paper.

It is well known that the stability and expansion of a map at a fixed point has a close relationship with the modulus of the eigenvalues of its derivative operator when the map is differentiable at the fixed point. Suppose that f is differentiable at $(0,0)$; then the induced map F is differentiable at O. Let $f_x(x,y)$ and $f_y(x,y)$ denote the first partial derivatives of f with respect to the first and the second variables at the point (x,y), respectively. Then we can get the following results on stability and expansion of the fixed point O of the induced system (7).

Theorem 7. *Assume that $k < \infty$. Denote $a := f_y(0,0)$, $b := f_x(0,0)$.*

(i) *If f is differentiable at $(0,0)$, then, for $a = 0$, the fixed point O of system (7) is asymptotically stable if and only if $|b| < 1$; and for $a \neq 0$, the fixed point O of system (7) is asymptotically stable if and only if $|a| < (k+1)/k$, and*

$$\begin{aligned} &|a| - 1 < -b < \left(a^2 + 1 - 2|a|\cos\phi\right)^{1/2}, \quad \text{for } k \text{ odd}, \\ &|a+b| < 1, \\ &|b| < \left(a^2 + 1 - 2|a|\cos\phi\right)^{1/2}, \\ &\qquad\qquad\qquad\qquad \text{for } k \text{ even}, \end{aligned} \tag{9}$$

where ϕ is the solution in $(0, \pi/(k+1))$ of equation $\sin(k\theta)/\sin[(k+1)\theta] = 1/|a|$.

(ii) *If f is continuously differentiable in a neighborhood of $(0,0)$ and $|b| - |a| > 1$, then the fixed point O of system (7) is a regular expanding fixed point in some norm in $\mathbf{R}^{k+1}$.*

Proof. When $a = 0$, it is easy to obtain that all the eigenvalues of $DF(O)$ have absolute values less than 1 if and only if $|b| < 1$. So, the result in (i) holds. When $a \neq 0$, the result in (i) can be directly derived by using Theorem 3 in [22]. Result (ii) can be derived from Lemma 2.1 of [10]. This completes the proof. □

Now, we establish a criterion of chaos for the induced system (7).

Theorem 8. *Let $f : D \subset \mathbf{R}^2 \to \mathbf{R}$ be a map and let it be continuously differentiable in a neighborhood of $(0,0)$ with $f(0,0) = 0$. Assume that*

(i) *$|f_x(0,0)| - |f_y(0,0)| > 1$, which implies that there exist a positive constant r^* and a norm $\|\cdot\|^*$ in $\mathbf{R}^{k+1}$ such that F is continuously differentiable in $\overline{B}_{r^*}(O)$ and O is a regular expanding fixed point of F in $\overline{B}_{r^*}(O)$ in the norm $\|\cdot\|^*$, where $\overline{B}_{r^*}(O)$ is the closed ball of radius r^* centered at O in $(\mathbf{R}^{k+1}, \|\cdot\|^*)$;*

(ii) *there exists a point $u^* \in D$ with $u^* \neq 0$, such that f is continuously differentiable in a neighborhood of $(u^*, 0)$ with $f(u^*, 0) = 0$, $f_x(u^*, 0) \neq 0$,*

(iia) *when $k = 1$, there exist $x_1, x_2 \in (-r^*, r^*)$ such that $x_1^2 + x_2^2 \neq 0$, $(x_1, x_2)^T \in B_{r^*}(O)$, f is continuously differentiable in a neighborhood of (x_2, u^*), and*

$$\begin{aligned} f(x_2, u^*) &= 0, \\ f(x_1, x_2) &= u^*, \\ f_x(x_2, u^*) &\neq 0, \\ f_x(x_1, x_2) &\neq 0; \end{aligned} \tag{10}$$

(iib) *when $k > 1$, there exist $x_1, x_2 \in (-r^*, r^*)$ such that $x_1^2 + x_2^2 \neq 0$, $(x_1, x_2, 0, \ldots, 0)^T \in B_{r^*}(O)$, f is continuously differentiable in a neighborhood of (x_2, u^*), and*

$$\begin{aligned} f(x_2, u^*) &= 0, \\ f(x_1, 0) &= u^*, \\ f_x(x_2, u^*) &\neq 0, \\ f_x(x_1, 0) &\neq 0. \end{aligned} \tag{11}$$

Then the induced system (7), and consequently system (2), is chaotic in the sense of both Devaney and Li-Yorke.

Proof. We will apply Lemma 5 to prove this theorem. So, we only need to show that all the assumptions in Lemma 5 are satisfied.

It follows from assumption (i) and the second conclusion of Theorem 7 that F is continuously differentiable in $\overline{B}_{r^*}(O)$, all the eigenvalues of $DF(O)$ have absolute values larger than 1, and O is a regular expanding fixed point of F in $\overline{B}_{r^*}(O)$ in some norm $\|\cdot\|^*$ of $\mathbf{R}^{k+1}$. Therefore, condition (i) in Lemma 5 is satisfied.

Next, we will show that O is a snap-back repeller of F in the norm $\|\cdot\|^*$. In the following, we will show that there exists a point $O_0 \in W$ with $O_0 \neq O$ satisfying

$$F^{k+2}(O_0) = O, \tag{12}$$

which implies that O is a snap-back repeller of F.

For the case where $k = 1$, it follows from condition (iia) that there exists a point $O_0 = (x_1, x_2)^T \in B_{r^*}(O)$, $O_0 \neq O$, such that $O_1 = F(O_0) = (x_2, u^*)^T$, $O_2 = F^2(O_0) = (u^*, 0)^T$, and $F^3(O_0) = O$.

For the case where $k > 1$, it follows from condition (iib) that there exists a point $O_0 = (x_1, x_2, 0, \ldots, 0)^T \in B_{r^*}(O)$, $O_0 \neq O$, such that $O_1 = F(O_0) = (x_2, 0, \ldots, 0, u^*)^T$, $O_j = F^j(O_0) = (0, \ldots, 0, \underbrace{u^*, 0, \ldots, 0}_{j})^T$ for $2 \leq j \leq k+1$, and $F^{k+2}(O_0) = O$.

It is obvious that F is continuously differentiable in some neighborhoods of $O_j := F(O_{j-1})$ for $1 \leq j \leq k+1$. So, we need to show that the following holds:

$$\det DF\left(O_j\right) \neq 0, \quad 0 \leq j \leq k+1. \tag{13}$$

If F is differentiable at $u = (u_1, \ldots, u_{k+1})^T \in \mathbf{R}^{k+1}$, then a direct calculation shows that

$$\det DF(u) = (-1)^k f_x(u_1, u_{k+1}). \tag{14}$$

From condition (i), we get that $|f_x(0,0)| > 1 + |f_y(0,0)| > 0$, which together with condition (ii) and (14) implies that conclusion (13) holds for $k = 1$ and $k > 1$.

Therefore, all the assumptions in Lemma 5 are satisfied. Then the induced system (7), and consequently system (2), is chaotic in the sense of both Devaney and Li-Yorke. This completes the proof. □

Remark 9. Since f is a function of two variables, the conditions in (ii) of Theorem 8 are not very strict conditions.

Based on Theorem 8, a chaotification scheme for the controlled system (3) with controller (4) is established in the following.

Theorem 10. *Consider the controlled system (3) with controller (4). Assume that*

(i) *f is continuously differentiable in $[-r, r]^2$ for some $r > 0$ with $f(0,0) = 0$, which implies that there exist positive constants M and N such that for any $(x, y) \in [-r, r]^2$*

$$\begin{aligned} |f(x,y)| &\leq M, \\ |f_x(x,y)| &\leq N, \\ |f_y(x,y)| &\leq N; \end{aligned} \tag{15}$$

(ii) *there exists a point $u^* \in (-r, r)$ with $u^* \neq 0$ such that $f(u^*, 0) = 0$.*

Then there exist two positive constants α_0 and β_0 satisfying

$$\begin{aligned} \alpha_0 &> \frac{M + |u^*|}{\varepsilon}, \\ \beta_0 &:= \frac{2m_0\varepsilon}{u^*} > \max\left\{\frac{1+2N}{\alpha_0}, \frac{3\varepsilon}{r}\right\}, \end{aligned} \tag{16}$$

where $\varepsilon > 0$ is any given constant and m_0 is some integer, such that, for any $\alpha > \alpha_0$ and $\beta = \beta_0$, the controlled system (3) with controller (4) is chaotic in the sense of both Devaney and Li-Yorke.

Proof. We will use Theorem 8 to prove this theorem. So, it suffices to show that the map $g(x, y) := f(x, y) + \alpha \mathrm{saw}_\varepsilon(\beta x)$ satisfies all the assumptions in Theorem 8.

For convenience, let $\alpha > \alpha_0$, $\beta = 2m\varepsilon/u^* > \max\{(1+2N)/\alpha_0, 3\varepsilon/r\}$ throughout the proof, where m is an undetermined integer. Let G denote the induced map of g.

It is obvious that the function $\mathrm{saw}_\varepsilon(\beta x)$ is continuously differentiable in $(-\varepsilon/\beta, \varepsilon/\beta)$. Then, from assumption (i), we obtain that g is continuously differentiable in $(-\varepsilon/\beta, \varepsilon/\beta)^2$ with $g(0,0) = 0$, O is a fixed point of the map G, and G is continuously differentiable in $(-\varepsilon/\beta, \varepsilon/\beta)^{k+1}$. It follows from the last two relations of (15) that

$$\begin{aligned} |g_x(0,0)| &= |f_x(0,0) + \alpha\beta| \geq \alpha\beta - |f_x(0,0)| \\ &\geq \alpha\beta - N > 1 + N \geq 1 + |f_y(0,0)| \\ &= 1 + |g_y(0,0)|. \end{aligned} \tag{17}$$

So condition (i) in Theorem 8 holds. Consequently, there exist a positive constant r^* and a norm $\|\cdot\|^*$ in $\mathbf{R}^{k+1}$ such that G is continuously differentiable in $\overline{B}_{r^*}(O)$ and O is a regular expanding fixed point of G in $\overline{B}_{r^*}(O)$ in the norm $\|\cdot\|^*$, where $\overline{B}_{r^*}(O) \subset (-\varepsilon/\beta, \varepsilon/\beta)^{k+1}$ is the closed ball of radius r^* centered at O in $(\mathbf{R}^{k+1}, \|\cdot\|^*)$. Further, suppose that $W \subset B_{r^*}(O)$ is an arbitrary neighborhood of O in $\mathbf{R}^{k+1}$. Then there exists a neighborhood U of 0 such that $U \times U \times \cdots \times U \subset W$.

Next, we need to show that g satisfies assumption (ii) in Theorem 8. It is obvious that $\mathrm{saw}_\varepsilon(\beta u^*) = 0$ and $\mathrm{saw}_\varepsilon(\beta x)$ is continuously differentiable in a neighborhood of u^*. So, $g(x, y)$ is continuously differentiable in a neighborhood of $(u^*, 0)$. From assumption (ii) and condition (15), it follows that

$$\begin{aligned} g(u^*, 0) &= f(u^*, 0) + \alpha \mathrm{saw}_\varepsilon(\beta u^*) = 0, \\ |g_x(u^*, 0)| &= |f_x(u^*, 0) + (-1)^m \alpha\beta| \\ &\geq \alpha\beta - |f_x(u^*, 0)| > 1 + N > 0. \end{aligned} \tag{18}$$

For $k = 1$, let

$$h_1(x) := f(x, u^*) + \alpha \mathrm{saw}_\varepsilon(\beta x). \tag{19}$$

It follows from assumption (i) and the definition of sawtooth function that h_1 is continuous in $[-\varepsilon/\beta, 3\varepsilon/\beta]$. From the first relation of (15), we get that

$$h_1\left(\frac{\varepsilon}{\beta}\right) = f\left(\frac{\varepsilon}{\beta}, u^*\right) + \alpha\varepsilon \geq \alpha\varepsilon - M > 0,$$
$$h_1\left(\frac{3\varepsilon}{\beta}\right) = f\left(\frac{3\varepsilon}{\beta}, u^*\right) - \alpha\varepsilon \leq M - \alpha\varepsilon < 0. \tag{20}$$

Therefore, by the intermediate value theorem, there exists a point x_2 with $\varepsilon/\beta < x_2 < 3\varepsilon/\beta$, such that $h_1(x_2) = 0$; that is, $g(x_2, u^*) = 0$. Similarly, let

$$h_2(x) := f(x, x_2) + \alpha \mathrm{saw}_\varepsilon(\beta x) - u^*. \tag{21}$$

It is also clear that h_2 is continuous in $[-\varepsilon/\beta, 3\varepsilon/\beta]$. It also follows from the first relation of (15) that

$$h_2\left(\frac{\varepsilon}{\beta}\right) = f\left(\frac{\varepsilon}{\beta}, x_2\right) + \alpha\varepsilon - u^* \geq \alpha\varepsilon - M - |u^*| > 0,$$
$$h_2\left(\frac{3\varepsilon}{\beta}\right) = f\left(\frac{3\varepsilon}{\beta}, x_2\right) - \alpha\varepsilon - u^* \leq M + |u^*| - \alpha\varepsilon < 0. \tag{22}$$

By the intermediate value theorem again, there exists a point x_1 with $\varepsilon/\beta < x_1 < 3\varepsilon/\beta$, such that $h_2(x_1) = 0$; that is, $g(x_1, x_2) = u^*$. It is clear that x_1 and x_2 are both in $(\varepsilon/\beta, 3\varepsilon/\beta) = (u^*\varepsilon/2mr, u^*\varepsilon/2mr)$. So we can take a sufficiently large integer $m_1 > 0$, such that $x_1, x_2 \in U$ with $x_1^2 + x_2^2 \neq 0$, and $(x_1, x_2)^T \in W$ for any $|m| \geq m_1$. It can easily be proved that g is continuously differentiable in some neighborhoods of (x_2, u^*) and (x_1, x_2). Now, we show $g_x(x_2, u^*) \neq 0$. Otherwise, if $g_x(x_2, u^*) = 0$, then the following equality holds:

$$f_x(x_2, u^*) + (-1)^m \alpha\beta = 0. \tag{23}$$

Hence, $\alpha\beta = |f_x(x_2, u^*)| \leq N$, which is a contradiction. Similarly, we can prove that $g_x(x_1, x_2) \neq 0$. Hence, condition (iia) in Theorem 8 holds.

For $k > 1$, the determination of x_2 can be derived from the proof of the above paragraph as $k = 1$. That is, there exists a point x_2 in $(\varepsilon/\beta, 3\varepsilon/\beta)$, such that $h_1(x_2) = 0$; that is, $g(x_2, u^*) = 0$. Set

$$h_3(x) := f(x, 0) + \alpha \mathrm{saw}_\varepsilon(\beta x) - u^*. \tag{24}$$

With a similar method to the above paragraph, we can also get that there exists a point x_1 in $(\varepsilon/\beta, 3\varepsilon/\beta)$ such that $h_3(x_1) = 0$, which implies that $g(x_1, 0) = u^*$. So we can also take a sufficiently large integer $m_2 > 0$, such that $x_1, x_2 \in U$ with $x_1^2 + x_2^2 \neq 0$, and $(x_1, x_2, 0, \ldots, 0)^T \in W$ for any $|m| \geq m_2$. It can also easily be proved that g is continuously differentiable in some neighborhoods of (x_2, u^*) and (x_1, x_2). The proofs of $g_x(x_2, u^*) \neq 0$ and $g_x(x_1, 0) \neq 0$ are similar to the above paragraph. So, the details are omitted.

Finally, let $|m_0| = \max\{m_1, m_2\}$. Then condition (ii) in Theorem 8 is satisfied for $m = m_0$. Therefore, for any $\alpha > \alpha_0$ and $\beta = \beta_0$, the controlled system (3) with controller (4) is chaotic in the sense of both Devaney and Li-Yorke. The proof is complete. □

Remark 11. It is clear that the classical sinusoidal function $\sin x$ has similar geometric properties to the sawtooth function. So the following function

$$v(n) = \alpha \sin(\beta x(n-k)) \tag{25}$$

can also be used as a controller to chaotify system (2), where β is some constant to be determined and $\alpha > 0$ is the controlled parameter. In fact, with a similar argument to the proof of Theorem 10, one can show that there also exist two positive constants β_0 and α_0 such that for any constant $\alpha > \alpha_0$ and $\beta = \beta_0$ the result in Theorem 10 holds.

Remark 12. In [10], a similar result is given for a class of maps with at least two fixed points. In such a case, the two chaotification schemes obtained in [10] and this paper can be used. However, there will be many chaotic invariant sets as pointed out in Lemma 2.2 of [10] when using the chaotification scheme in [10]. It seems that the chaotic behaviors induced by a heteroclinic cycle connecting repellers are more complex than that induced by a single snap-back repeller. The difference between them will be our further research. But when the original system only has one fixed point, the chaotification scheme obtained in [10] cannot be used. Then, we can use the chaotification scheme obtained in this paper to chaotify this system.

Remark 13. Since the point u^* in assumption (ii) of Theorem 10 can be negative, the value of m_0 determined in this paper can be a negative integer. In addition, it is very difficult to determine the concrete value m_0 since the concrete expanding area of a fixed point is not easy to obtain. To the best of our knowledge, there are few methods to determine the concrete expanding area of a fixed point in the existing literatures. So, in practical problems, we can take $|m_0|$ large enough such that the chaotification scheme can be effective.

In the last part of this section, we give an example to illustrate the theoretical result of Theorem 10.

Example 14. We take the map f in (2) as the following:

$$f(x, y) = 0.01x(x-1) - 0.01y^2. \tag{26}$$

It is clear that f is continuously differentiable in $\mathbf{R}^2$ and satisfies $f(0,0) = 0$. Without loss of generality, we take $r = 3$ in Theorem 10. Then, for any $(x, y) \in [-3, 3]^2$, we get that

$$|f(x, y)| \leq 0.21,$$
$$|f_x(x, y)| \leq 0.07, \tag{27}$$
$$|f_y(x, y)| \leq 0.06.$$

Hence, we take $M = 0.21$, $N = 0.07$, and $r = 3$ in assumption (i) of Theorem 10. It is also clear that the equation $f(x, 0) = 0$

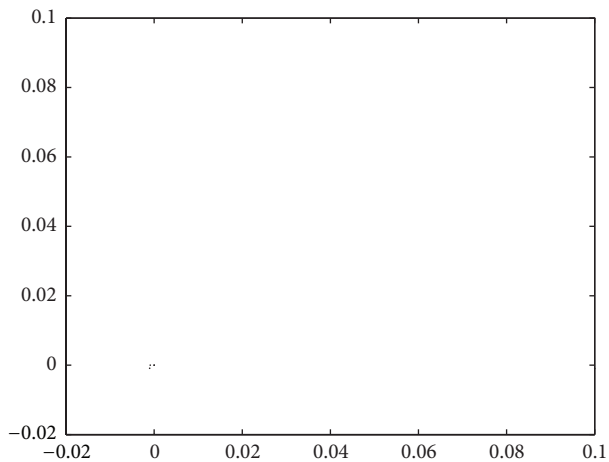


Figure 1: Simple dynamical behaviors of uncontrolled system (7) for $k = 1$, where the initial value is taken as $u(0) = (0.1, 0.1)^T$.

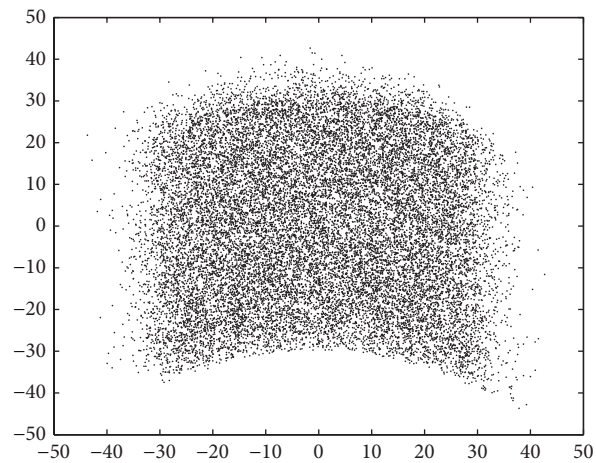


Figure 2: Complex dynamical behaviors of controlled system (8) for $\alpha = 30$, $\beta = 20$, $\varepsilon = 1$, and $k = 1$, where the initial value is taken as $u(0) = (0.1, 0.1)^T$.

has a nonzero solution $u^* = 1$, which lies in $(-3, 3)$. Therefore, all the assumptions in Theorem 10 are satisfied. Here, we take the constant $\varepsilon = 1$ in controller (4). Then, it follows from Theorem 10 that there exist two positive constants

$$\begin{aligned} \alpha_0 &> \frac{M + |u^*|}{\varepsilon} = 1.21, \\ \beta_0 &= \frac{2m_0\varepsilon}{u^*} = 2m_0 > \max\left\{\frac{1+2N}{\alpha_0}, \frac{3\varepsilon}{r}\right\} \\ &= \max\left\{\frac{1.14}{\alpha_0}, 1\right\} = 1, \end{aligned} \tag{28}$$

where m_0 is some positive integer, such that, for any $\alpha > \alpha_0$ and $\beta = \beta_0$, the controlled system (3) with controller (4) is chaotic in the sense of both Devaney and Li-Yorke.

In fact, there is only one fixed point $O := (0, \ldots, 0)^T \in \mathbf{R}^{k+1}$ in the uncontrolled system (7). It is obvious that $f_y(0,0) = 0$ and $f_x(0,0) = -0.01$, which imply that O is asymptotically stable from result (i) in Theorem 7. It is also clear that all the solutions of the uncontrolled system (7) are bounded if the initial values are taken from $[-3, 3]^{k+1}$. Therefore, if we take an initial condition $u(0) = (0.1, \ldots, 0.1)^T \in \mathbf{R}^{k+1}$, then the solution $u(n)$ of the uncontrolled system (7) should tend to the asymptotically stable fixed point O when n tends to infinity. This is confirmed in Figures 1 and 3.

Here, we take $\varepsilon = 1$, $m_0 = 10$, $\alpha = 30$, $\beta = 20$, $k = 1, 2$, and n from 0 to 20000 for computer simulations. The simulated results show that the original system (7) has simple dynamical behaviors, and the controlled system (8) has complex dynamical behaviors; see Figures 1–4.

It should be pointed out that the relative existing chaotification scheme in [10] is not available for this map since there is only one fixed point.

4. Conclusion

In this paper, we study the chaotification problem for a class of delay difference equations with at least one fixed

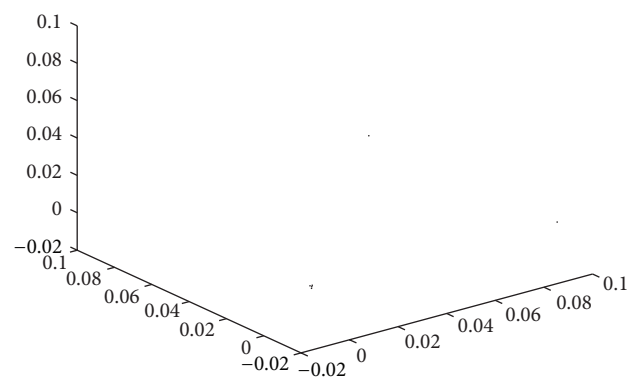


Figure 3: Simple dynamical behaviors of uncontrolled system (7) for $k = 2$, where the initial value is taken as $u(0) = (0.1, 0.1, 0.1)^T$.

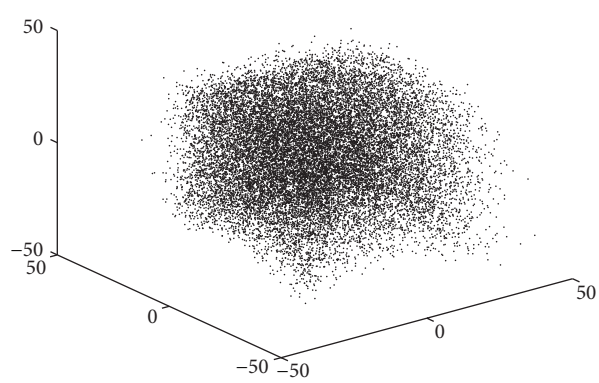


Figure 4: Complex dynamical behaviors of controlled system (8) for $\alpha = 30$, $\beta = 20$, $\varepsilon = 1$, and $k = 2$, where the initial value is taken as $u(0) = (0.1, 0.1, 0.1)^T$.

point. We first establish a criterion of chaos by using the snap-back repeller theory. Then, based on this criterion of chaos and the feedback control approach, we establish a chaotification scheme. We have proved that the controlled system is chaotic in the sense of both Devaney and Li-Yorke when the parameters of the system satisfy some mild conditions. Numerical simulations confirm the theoretical analysis. The chaotification problem for more general maps in the original system will be our further research.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grants 11101246, 61273088, and 10971120), the Nature Science Foundation of Shandong Province (Grant ZR2010FM010), and the Postdoctoral Science Foundation of China (Grant 2014M561908).

References

[1] G. R. Chen and X. N. Dong, *From Chaos to Order: Methodologies, Perspectives, and Applications*, World Scientific Publishing, Singapore, 1998.

[2] T. Kapitaniak, *Chaos for Engineers: Theory, Applications, and Control*, Springer, New York, NY, USA, 1998.

[3] A. L. Fradkov and A. Pogromsky, *Introduction to Control of Oscillations and Chaos*, World Scientific, Singapore, 1999.

[4] G. Jakimoski and L. Kocarev, "Chaos and cryptography: block encryption ciphers based on chaotic maps," *IEEE Transactions on Circuits and Systems. I. Fundamental Theory and Applications*, vol. 48, no. 2, pp. 163–169, 2001.

[5] G. R. Chen and D. J. Lai, "Feedback control of Lyapunov exponents for discrete-time dynamical systems," *International Journal of Bifurcation and Chaos in Applied Sciences and Engineering*, vol. 6, no. 7, pp. 1341–1349, 1996.

[6] G. R. Chen and D. J. Lai, "Anticontrol of chaos via feedback," in *Proceedings of the IEEE Conference on Decision and Control*, pp. 367–372, San Diego, Calif, USA, 1997.

[7] G. R. Chen and D. J. Lai, "Feedback anticontrol of discrete chaos," *International Journal of Bifurcation and Chaos in Applied Sciences and Engineering*, vol. 8, no. 7, pp. 1585–1590, 1998.

[8] G. R. Chen and Y. M. Shi, "Introduction to anti-control of discrete chaos: theory and applications," *Philosophical Transactions of the Royal Society of London. Series A. Mathematical, Physical and Engineering Sciences*, vol. 364, no. 1846, pp. 2433–2447, 2006.

[9] Z. C. Li, "Chaotification for linear delay difference equations," *Advances in Difference Equations*, vol. 2013, article 59, 11 pages, 2013.

[10] Z. C. Li, "Anticontrol of chaos for a class of delay difference equations based on heteroclinic cycles connecting repellers," *Abstract and Applied Analysis*, vol. 2014, Article ID 260150, 8 pages, 2014.

[11] Z. L. Zhou, *Symbolic Dynamics*, Shanghai Scientific and Technological Education Publishing House, Shanghai, China, 1997.

[12] T. Y. Li and J. A. Yorke, "Period three implies chaos," *The American Mathematical Monthly*, vol. 82, no. 10, pp. 985–992, 1975.

[13] R. L. Devaney, *An Introduction to Chaotic Dynamical Systems*, Addison-Wesley, New York, NY, USA, 1987.

[14] W. Huang and X. D. Ye, "Devaney's chaos or 2-scattering implies Li-Yorke's chaos," *Topology and Its Applications*, vol. 117, no. 3, pp. 259–272, 2002.

[15] Y. M. Shi, P. Yu, and G. R. Chen, "Chaotification of discrete dynamical systems in Banach spaces," *International Journal of Bifurcation and Chaos in Applied Sciences and Engineering*, vol. 16, no. 9, pp. 2615–2636, 2006.

[16] Y. M. Shi and G. R. Chen, "Discrete chaos in Banach spaces," *Science in China Series A: Mathematics*, vol. 34, pp. 595–609, 2004 (Chinese), English version: vol. 48, pp. 222–238, 2005.

[17] F. R. Marotto, "Snap-back repellers imply chaos in R^n," *Journal of Mathematical Analysis and Applications*, vol. 63, no. 1, pp. 199–223, 1978.

[18] Y. M. Shi and G. R. Chen, "Chaos of discrete dynamical systems in complete metric spaces," *Chaos, Solitons and Fractals*, vol. 22, no. 3, pp. 555–571, 2004.

[19] F. R. Marotto, "On redefining a snap-back repeller," *Chaos, Solitons and Fractals*, vol. 25, no. 1, pp. 25–28, 2005.

[20] Z. C. Li, Y. M. Shi, and W. Liang, "Discrete chaos induced by heteroclinic cycles connecting repellers in Banach spaces," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 72, no. 2, pp. 757–770, 2010.

[21] X. F. Wang and G. R. Chen, "Chaotification via arbitrary small feedback controls," *International Journal of Intelligent Systems*, vol. 10, pp. 549–570, 2000.

[22] S. A. Kuruklis, "The asymptotic stability of $x_{n+1} - ax_n + bx_{n-k} = 0$," *Journal of Mathematical Analysis and Applications*, vol. 188, no. 3, pp. 719–731, 1994.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 207427, 8 pages
http://dx.doi.org/10.1155/2015/207427

Research Article

The Stability of Solutions for the Generalized Degasperis-Procesi Equation with Variable Coefficients

Jing Chen

School of Science, Southwest University of Science and Technology, Mianyang 621000, China

Correspondence should be addressed to Jing Chen; chenjing_math@163.com

Received 19 May 2015; Accepted 25 June 2015

Academic Editor: Xinguang Zhang

A generalized Degasperis-Procesi equation with variable coefficients is investigated. The $L^1(R)$ stability of the strong solution for the equation is established under certain assumptions.

1. Introduction

The Degasperis-Procesi (DP) equation

$$v_t - v_{txx} + 4vv_x = 3v_x v_{xx} + vv_{xxx}, \quad t > 0,\ x \in R, \tag{1}$$

was discovered by Degasperis and Procesi [1] in a search for integrable equations similar to the Camassa-Holm equation. Degasperis and Procesi [1] studied a family of third order dispersive nonlinear equations

$$v_t + v_x + \gamma v_{xxx} - \alpha^2 v_{txx} = \left(c_1 v^2 + c_2 v_x^2 + c_3 v v_{xx}\right)_x, \tag{2}$$

where $\alpha, \gamma, c_0, c_1, c_2, c_3 \in R$. It is found in [1] that there are only three equations that satisfy asymptotic integrability conditions within this family. By rescaling and applying a Galilean transformation, the three equations are Korteweg-de Vries equation

$$v_t + v_{xxx} + vv_x = 0, \quad t > 0,\ x \in R, \tag{3}$$

the Camassa-Holm equation

$$v_t - v_{txx} + 3vv_x = 2v_x v_{xx} + vv_{xxx}, \quad t > 0,\ x \in R, \tag{4}$$

and the Degasperis-Procesi equation (1). Degasperis et al. [2] proved the formal integrability of (1) and the existence of the nonsmooth solutions by constructing a Lax pair.

In recent years, (1) which plays a similar role in water wave theory as the Camassa-Holm equation has caused extensive concern of many scholars (see [1–11]). For example, Coclite and Karlsen [3] established the well-posedness of $L^1 \cap BV$ weak solutions for (1). They proved uniqueness within a class of discontinuous solutions to (1) in [4]. Escher et al. [5] established the precise blow-up rate and proved the existence and uniqueness of global weak solutions to (1) in which the initial data satisfied appropriate conditions. Lai and Wu [7] investigated the local well-posedness of solutions to a generalization of both (1) and (4) in the Sobolev space $H^s(R)$ with $s > 3/2$. Lenells [8] classified all weak traveling wave solutions of the Degasperis-Procesi equation (1). Ai and Gui [9] proved global existence of solutions for the viscous Degasperis-Procesi equation and showed that the blow-up phenomena occurs in finite time. Fu et al. [11] studied the orbital stability of the peakons for the Degasperis-Procesi equation with a strong dispersive term on the line and proved that the shapes of these peakons were stable under small perturbations.

As we know, their coefficients play an important role to study the fundamental dynamical properties of the Degasperis-Procesi models. It prompts us to study the following generalized Degasperis-Procesi equation:

$$\begin{aligned} v_t - v_{txx} &+ m(t,x) f'(v) v_x \\ &= b(t,x)\left(f'(v) v_{xxx} + 3f''(v) v_x v_{xx} + f'''(v) v_x^3\right) \\ &\quad + 2b_x(t,x)\left(f'(v) v_{xx} + f''(v) v_x^2\right) \\ &\quad + b_{xx}(t,x) f'(v) v_x, \end{aligned} \tag{5}$$

where $m(t,x) \in C_0(R)$, $b(t,x) \in C_0^3(R)$, and function $f(\cdot)$ is a polynomial of order n ($n \geq 2$). Letting $m = 4$, $b = 1$, $f(v) = v^2/2$, (5) reduces to the Degasperis-Procesi equation (1). We consider the Cauchy problem of (5) with an initial condition $v_0(x)$. Namely,

$$v(0,x) = v_0(x), \quad x \in R. \tag{6}$$

Assume that (5) possesses a bounded strong solution in its maximum existence time interval $[0,T)$ and v_0 lies in $L^1(R) \cap H^s(R)$ ($s > 3/2$). We use the approaches of Kružkov doubling the variables presented in [12] to prove the L^1 stability of the solution for the variable coefficients equation (5). From our knowledge, it has not been acquired in the literature.

This paper is organized as follows. Section 2 gives several lemmas. The proof of local solution stability is presented in Section 3.

2. Preliminaries

Applying the operator $\Lambda^{-2} = (1-\partial_x^2)^{-1}$ to (5), we obtain its equivalent form

$$\begin{aligned} &v_t + b(t,x) f'(v) v_x + \Lambda^{-2}(m(t,x) - b(t,x)) \partial_x f(v) \\ &\quad = 0, \end{aligned} \tag{7}$$

where $\Lambda^{-2}p(t,x) = (1-\partial_x^2)^{-1}p(t,x) = (1/2)\int_R e^{-|x-y|}p(t,y)dy$.

Equations (5) and (6) are equivalent to the problem

$$\begin{aligned} v_t + \partial_x P(t,x,v) + \Psi(t,x,v) &= 0, \\ v(0,x) &= v_0(x), \end{aligned} \tag{8}$$

where $P(t,x,v) = b(t,x)f(v)$ and $\Psi(t,x,v) = \Lambda^{-2}(m(t,x) - b(t,x))\partial_x f(v) - b_x(t,x)f(v)$. Notice that $\partial_x P = P_x + P_v v_x$.

Remark. According to the statements presented in [7] or [12], we know that problem (8) has a unique local solution in the space $C([0,T), H^s(R))$ if we assume $v_0(x) \in H^s(R)$ ($s > 3/2$).

Assume that $v_1(t,x)$ and $v_2(t,x)$ are solutions of problem (8) in the domain $[0,T) \times R$ with initial functions $v_{10}(x)$ and $v_{20}(x) \in L^1(R) \cap H^s(R)$ ($s > 3/2$), where T is the maximum existence time of solutions. For simplicity, we denote by c any positive constants. Now we give several lemmas.

Lemma 1. *Let $v(t,x)$ be the solution of problem (8) and $\|v\|_{L^\infty(R)} \leq M$. Then*

$$\|\Psi(t,x,v)\|_{L^\infty(R)} \leq cM^n, \tag{9}$$

where positive constant c depends on $\|v_0\|_{L^\infty}$ and T.

Proof. We have

$$\begin{aligned} &|\Psi(t,x,v)| \\ &\quad = \left|\Lambda^{-2}(m(t,x) - b(t,x))\partial_x f(v) - b_x(t,x) f(v)\right| \\ &\quad \leq \left|\frac{1}{2}\int_R e^{-|x-y|} m(t,y)\partial_y f(v)\,dy\right| \\ &\qquad + \left|\frac{1}{2}\int_R e^{-|x-y|} b(t,y)\partial_y f(v)\,dy\right| + \left|b_x f(v)\right| \\ &\quad \leq c\left|\int_R e^{-|x-y|}\operatorname{sign}(y-x) f(v)\,dy\right| + c\left|f(v)\right| \\ &\quad \leq cM^n, \end{aligned} \tag{10}$$

in which we have used $\int_R e^{-|x-y|}dx = 2$ to complete the proof. □

Lemma 2. *Assume that $v_1(t,x)$ and $v_2(t,x)$ are solutions of problem (8) in the domain $[0,T) \times R$, $\|v_1\|_{L^\infty(R)} \leq M$, and $\|v_2\|_{L^\infty(R)} \leq M$. Then*

$$\begin{aligned} &\int_{-\infty}^{+\infty} \left|\Psi(t,x,v_1) - \Psi(t,x,v_2)\right| dx \\ &\qquad \leq c\int_{-\infty}^{+\infty} \left|v_1 - v_2\right| dx, \end{aligned} \tag{11}$$

where $c > 0$ depends on $\|v_{10}\|_{L^\infty(R)}$, $\|v_{20}\|_{L^\infty(R)}$ and T.

Proof. Using the property of the operator Λ^{-2}, we get

$$\begin{aligned} &\int_{-\infty}^{+\infty} \left|\Psi(t,x,v_1) - \Psi(t,x,v_2)\right| dx \\ &\quad \leq \int_{-\infty}^{+\infty} \left|\Lambda^{-2}(m-b)\partial_x(f(v_1) - f(v_2))\right| dx \\ &\quad + \int_{-\infty}^{+\infty} \left|(f(v_1) - f(v_2))\right| dx \leq c\int_{-\infty}^{+\infty} dx \\ &\quad \cdot \int_{-\infty}^{+\infty} \left|e^{-|x-y|}\operatorname{sign}(y-x)(f(v_1) - f(v_2))\right| dy \\ &\quad + c\int_{-\infty}^{+\infty} \left|v_1 - v_2\right| dx \leq c\int_{-\infty}^{+\infty} \left|f(v_1) - f(v_2)\right| dy \\ &\quad \cdot \int_{-\infty}^{+\infty} e^{-|x-y|}dx + c\int_{-\infty}^{+\infty} \left|v_1 - v_2\right| dx \\ &\quad \leq c\int_{-\infty}^{+\infty} \left|v_1 - v_2\right| dx, \end{aligned} \tag{12}$$

in which we apply the Tonelli Theorem to complete the proof. □

Let $\delta(\sigma) \geq 0$, $\delta(\sigma) \equiv 0$, for $|\sigma| \geq 1$; $\int_{-\infty}^{+\infty}\delta(\sigma)d\sigma = 1$ and $\delta(\sigma)$ is infinitely differential on $(-\infty,+\infty)$. Set

$\delta_\varepsilon(\sigma) = \delta(\varepsilon^{-1}\sigma)/\varepsilon$, where ε is an arbitrary positive constant. It is found that $\delta_\varepsilon(\sigma) \in C_0^\infty(-\infty, +\infty)$ and

$$\begin{aligned} &\delta_\varepsilon(\sigma) \geq 0, \\ &\delta_\varepsilon(\sigma) = 0 \\ &\qquad\qquad\qquad\qquad \text{for } |\sigma| \geq \varepsilon, \\ &|\delta_\varepsilon(\sigma)| \leq \frac{c}{\varepsilon}, \\ &\int_{-\infty}^{+\infty} \delta_\varepsilon(\sigma)\, d\sigma = 1. \end{aligned} \tag{13}$$

Let the function $\phi(x)$ be defined and locally integrable on $(-\infty, +\infty)$. Set $\phi^\varepsilon(x)$; denote the approximation function of $\phi(x)$ as

$$\phi^\varepsilon(x) = \frac{1}{\varepsilon}\int_{-\infty}^{+\infty} \delta\left(\frac{x-y}{\varepsilon}\right)\phi(y)\, dy. \tag{14}$$

We call x_0 a Lebesgue point of the function $\phi(x)$ if

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_{x_0-\varepsilon}^{x_0+\varepsilon} |\phi(x) - \phi(x_0)|\, dx = 0. \tag{15}$$

At any Lebesgue point x_0, we get

$$\lim_{\varepsilon\to 0}\phi^\varepsilon(x_0) = \phi(x_0). \tag{16}$$

Since the set of points which are not Lebesgue points of $\phi(x)$ has measure zero, we have $\phi^\varepsilon(x) \to \phi(x)$ as $\varepsilon \to 0$ almost everywhere.

For any $T > 0$, we denote the band $\{(t,x) : [0,T] \times R\}$ by ϖ_T. Let $K_r = \{x : |x| \leq r\}$ and

$$\begin{aligned} \Pi = \Big\{(t,x,\tau,y) : &\left|\frac{t-\tau}{2}\right| \leq \varepsilon,\ \rho \leq \frac{t+\tau}{2} \leq T \\ &- \rho,\ \left|\frac{x-y}{2}\right| \leq \varepsilon,\ \left|\frac{x+y}{2}\right| \leq r-\rho\Big\}, \end{aligned} \tag{17}$$

where $r > 0$, $\rho > 0$.

We state the concept of a characteristic cone. Let $\|v\|_{L^\infty(R)} \leq M$, for any $R_1 > 0$; we define

$$N > \max_{(t,x)\in[0,T]\times K_{R_1}} |f'(v)|. \tag{18}$$

Let Ω represent the cone $\{(t,x) : |x| \leq R_1 - Nt, 0 \leq t \leq T_1 = \min(T, R_1 N^{-1})\}$ and let S_τ designate the cross section of the cone Ω by the plane $t = \tau, \tau \in [0, T_1]$.

Lemma 3 (see [12]). *Let the function $\phi(t,x)$ be bounded and measurable in cylinder $[0,T_1]\times K_r$. For any $\rho \in (0, \min[r, T_1])$ and any $\varepsilon \in (0,\rho)$, the function*

$$V_\varepsilon = \frac{1}{\varepsilon^2}\iiiint_\Pi |\phi(t,x) - \phi(\tau,y)|\, dt\, dx\, d\tau\, dy \tag{19}$$

satisfies

$$\lim_{\varepsilon\to 0} V_\varepsilon = 0. \tag{20}$$

Lemma 4 (see [12]). *If $\partial F(\phi)/\partial\phi$ is bounded, the function $H(\phi,\psi) = \operatorname{sign}(\phi-\psi)(F(\phi) - F(\psi))$ satisfies the Lipschitz condition in ϕ and ψ.*

Lemma 5. *If $v(t,x)$ is the solution of problem (11) on ϖ_T, $\eta(t,x) \in C_0^\infty(\varpi_T)$, it holds that*

$$\begin{aligned} \iint_{\varpi_T} \{&|v-k|\,\eta_t \\ &+ \operatorname{sign}(v-k)\,[P(t,x,v) - P(t,x,k)]\,\eta_x \\ &- \operatorname{sign}(v-k)\,(P_x(t,x,k) + \Psi(t,x,v))\,\eta\}\, dt\, dx \\ &= 0, \end{aligned} \tag{21}$$

where k is an arbitrary constant.

Proof. Suppose that $\Phi(v)$ is a twice differential function. Multiplying the first equation of problem (8) by $\Phi'(v)\eta(t,x)$ and integrating over ϖ_T, we get

$$\begin{aligned} \iint_{\varpi_T} \Big\{&\Phi'(v)\,\eta v_t + \Phi'(v)\,\eta P_x + \Phi'(v)\,\eta P_v v_x \\ &+ \Phi'(v)\,\eta\Psi(t,x,v)\Big\}\, dt\, dx = 0. \end{aligned} \tag{22}$$

Using the method of integration by parts, we get

$$\iint_{\varpi_T} \Phi'(v)\,\eta v_t dt\, dx = -\iint_{\varpi_T} \Phi(v)\,\eta_t dt\, dx. \tag{23}$$

Notice that

$$\begin{aligned} &\left(\int_k^v \Phi'(z)\,P_z(t,x,z)\,dz\right)'_x \\ &\quad = \Phi'(v)\,P_v(v)\,v_x \\ &\qquad + \int_k^v \left(\Phi''(z)\,P_z z_x + \Phi'(z)\,P_{zx}\right) dz. \end{aligned} \tag{24}$$

So

$$\begin{aligned} &\int_{-\infty}^{+\infty}\left(\int_k^v \Phi'(z)\,P_z dz\right)\eta_x dx \\ &\quad = -\int_{-\infty}^{+\infty} \Phi'(v)\,P_v v_x \eta\, dx \\ &\qquad - \int_{-\infty}^{+\infty}\left[\int_k^v \left(\Phi''(z)\,P_z z_x + \Phi'(z)\,P_{zx}\right) dz\right]\eta\, dx. \end{aligned} \tag{25}$$

Then we have

$$\begin{aligned} \iint_{\varpi_T} \Phi'(v)\,P_v v_x \eta\, dt\, dx = &-\iint_{\varpi_T}\left(\int_k^v v'(z)\,P_z dz\right) \\ &\cdot \eta_x dt\, dx \\ &+ \iint_{\varpi_T}\left[\int_k^v \left(\Phi''(z)\,P_z z_x + \Phi'(z)\,P_{zx}\right) dz\right]\eta\, dt\, dx. \end{aligned} \tag{26}$$

Substituting (23) and (26) into (22), we get

$$\begin{aligned}\iint_{\varpi_T} &\Big\{\Phi'(v)\eta_t + \Big(\int_k^v \Phi'(z)P_z dz\Big)\eta_x \\ &+ \Big(\int_k^v \Phi''(z)P_z z_x dz\Big)\eta + \Big(\int_k^v \Phi'(z)P_{zx} dz\Big)\eta \\ &- \Phi'(v)\eta P_x - \Phi'(v)\eta\Psi\Big\}\,dt\,dx = 0.\end{aligned} \tag{27}$$

Let $\Phi^\varepsilon(v)$ be an approximation of the function $|v-k|$. When $\varepsilon \to 0$, $\Phi^\varepsilon(v) \to \Phi(v)$. Setting $\Phi(v) = |v-k|$, then $\Phi'(v) = \operatorname{sign}(v-k)$, $\Phi''(v) = 0$. Hence,

$$\begin{aligned}&\int_k^v \Phi'(z)P_z dz \\ &\quad = \operatorname{sign}(v-k)\,(P(t,x,v) - P(t,x,k)),\end{aligned} \tag{28}$$

$$\begin{aligned}&\int_k^v \Phi'(z)P_{zx} dz \\ &\quad = \operatorname{sign}(v-k)\,(P_x(t,x,v) - P_x(t,x,k)),\end{aligned} \tag{29}$$

combining with (27), we complete the proof. □

3. Main Result

Set function $\eta(t,x) \in C_0^\infty(\varpi_T)$, $\eta(t,x) \equiv 0$, outside the cylinder $\{(t,x)\} = [\rho, T-2\rho] \times K_{r-2\rho}$, where $K_{r-2\rho} = \{|x| : |x| \le r-2\rho\}$, $r > 0$, $0 < 2\rho < \min(T,r)$. Now we give the main result of this work.

Theorem 6. *Assume that $v_1(t,x)$ and $v_2(t,x)$ are two strong solutions of problem (8) with initial data $v_{10}(x), v_{20}(x) \in L^1(R) \cap H^s(R)$ ($s > 3/2$). Let $T > 0$ be the maximum existence time of $v_1(t,x)$ and $v_2(t,x)$. If $\|v_1(t,x)\|_{L^\infty} \le M$ and $\|v_2(t,x)\|_{L^\infty} \le M$, for any $t \in [0,T]$, it holds that*

$$\begin{aligned}&\|v_1(t,x) - v_2(t,x)\|_{L^1} \\ &\quad \le e^{ct}\int_{-\infty}^{+\infty} |v_{10}(x) - v_{20}(x)|\,dx,\end{aligned} \tag{30}$$

where c is a positive constant depending on $\|v_{10}\|_{L^\infty(R)}$ and $\|v_{20}\|_{L^\infty(R)}$.

Proof. Set $k = \psi(\tau,y)$, $\eta(t,x) = h(t,x,\tau,y)$. Using the Kružkov device of doubling the variables in [12] and Lemma 5, we get

$$\begin{aligned}&\iiiint_{\varpi_T\times\varpi_T} \{|v_1(t,x) - v_2(\tau,y)|\,h_t \\ &\quad + \operatorname{sign}(v_1(t,x) - v_2(\tau,y)) \\ &\quad \cdot [P(t,x,v_1(t,x)) - P(\tau,y,v_2(\tau,y))]\,h_x \\ &\quad - \operatorname{sign}(v_1(t,x) - v_2(\tau,y)) \\ &\quad \cdot [P_x(t,x,v_2(\tau,y)) + \Psi(t,x,v_1(t,x))]\,h\}\,dt\,dx\,d\tau\,dy \\ &\quad = 0.\end{aligned} \tag{31}$$

Similarly, we have

$$\begin{aligned}&\iiiint_{\varpi_T\times\varpi_T} \{|v_2(\tau,y) - v_1(t,x)|\,h_\tau \\ &\quad + \operatorname{sign}(v_2(\tau,y) - v_1(t,x)) \\ &\quad \cdot [P(\tau,y,v_2(\tau,y)) - P(t,x,v_1(t,x))]\,h_y \\ &\quad - \operatorname{sign}(v_2(\tau,y) - v_1(t,x)) \\ &\quad \cdot [P_y(\tau,y,v_1(t,x)) + \Psi(\tau,y,v_2(\tau,y))]\,h\}\,dt\,dx\,d\tau\,dy \\ &\quad = 0.\end{aligned} \tag{32}$$

Adding (31) and (32), we obtain

$$\begin{aligned}&\iiiint_{\varpi_T\times\varpi_T} \{|v_1(t,x) - v_2(\tau,y)|\,(h_t + h_\tau) \\ &\quad + \operatorname{sign}(v_1(t,x) - v_2(\tau,y))\,[P(t,x,v_1(t,x)) \\ &\quad - P(t,x,v_2(\tau,y))]\,(h_x + h_y) \\ &\quad + \operatorname{sign}(v_1(t,x) - v_2(\tau,y)) \\ &\quad \cdot ([P(\tau,y,v_2(\tau,y)) - P(t,x,v_2(\tau,y))]\,h_x \\ &\quad + [P(\tau,y,v_1(t,x)) - P(t,x,v_1(t,x))]\,h_y \\ &\quad + [P_y(\tau,y,v_1(t,x)) - P_x(t,x,v_2(\tau,y))]\,h) \\ &\quad - \operatorname{sign}(v_1(t,x) - v_2(\tau,y))\,[\Psi(t,x,v_1(t,x)) \\ &\quad - \Psi(\tau,y,v_2(\tau,y))]\,h\}\,dt\,dx\,d\tau\,dy \\ &\quad = \iiiint_{\varpi_T\times\varpi_T} (A_1 + A_2 + A_3 + A_4)\,dt\,dx\,d\tau\,dy \\ &\quad = 0.\end{aligned} \tag{33}$$

Set function

$$\begin{aligned}&h(t,x,\tau,y) \\ &\quad = \eta\Big(\frac{t+\tau}{2}, \frac{x+y}{2}\Big)\delta_\varepsilon\Big(\frac{t-\tau}{2}\Big)\delta_\varepsilon\Big(\frac{x-y}{2}\Big) \\ &\quad = \eta(\cdots)\,\lambda_\varepsilon\begin{pmatrix}\vdots\end{pmatrix},\end{aligned} \tag{34}$$

in which $(\cdots) = ((t+\tau)/2, (x+y)/2)$ and $\begin{pmatrix}\vdots\end{pmatrix} = ((t-\tau)/2, (x-y)/2)$. Thus, we obtain

$$\begin{aligned}h_t + h_\tau &= \eta_t(\cdots)\,\lambda_\varepsilon\begin{pmatrix}\vdots\end{pmatrix}, \\ h_x + h_y &= \eta_x(\cdots)\,\lambda_\varepsilon\begin{pmatrix}\vdots\end{pmatrix}.\end{aligned} \tag{35}$$

We will prove that the form A_3 in (33) approaches zero as $\varepsilon \to 0$. In fact, the coefficients of η_x and η_y in A_3 vanish for

$|t-\tau|+|x-y|=0$. Thus the integrals of A_3 can be rewritten as the following concrete form:

$$\begin{aligned}&\iiiint_{\omega_T\times\omega_T}\operatorname{sign}\left(v_1(t,x)-v_2(\tau,y)\right)\\&\quad\cdot\left\{\left[P\left(\tau,y,v_2(\tau,y)\right)-P\left(t,x,v_2(\tau,y)\right)\right](\lambda_\varepsilon)_x\right.\\&\quad-P_x\left(t,x,v_2(\tau,y)\right)\lambda_\varepsilon\\&\quad+\left[P\left(\tau,y,v_1(t,x)\right)-P\left(t,x,v_1(t,x)\right)\right](\lambda_\varepsilon)_y\\&\quad\left.+P_y\left(\tau,y,v_1(t,x)\right)\lambda_\varepsilon\right\}\eta(\cdots)\,dt\,dx\,d\tau\,dy=A_\varepsilon.\end{aligned}\tag{36}$$

In the following computations, we omit the index ε of function λ. Applying the Taylor formula, we have the relations

$$\begin{aligned}&\left[P\left(\tau,y,v_2(\tau,y)\right)-P\left(t,x,v_2(\tau,y)\right)\right]\lambda_x\\&\quad-P_x\left(t,x,v_2(\tau,y)\right)\lambda\\&\quad=\left(\left[P\left(\tau,y,v_2(\tau,y)\right)-P\left(t,x,v_2(\tau,y)\right)\right]\lambda\right)_x\\&\quad=P_\tau\left(\tau,y,v_2(\tau,y)\right)(\tau-t)\lambda_x\\&\quad+P_y\left(\tau,y,v_2(\tau,y)\right)(y-x)\lambda_x\\&\quad-P_y\left(\tau,y,v_2(\tau,y)\right)\lambda+\alpha_1\lambda_x+\alpha_0\lambda\\&\quad=P_\tau\left(\tau,y,v_2(\tau,y)\right)\left((\tau-t)\lambda\right)_x\\&\quad+P_y\left(\tau,y,v_2(\tau,y)\right)\left((y-x)\lambda\right)_x+\alpha_1\lambda_x+\alpha_0\lambda.\end{aligned}\tag{37}$$

It is seen that the identity $\lambda_y=-\lambda_x$. In a similar way, we obtain

$$\begin{aligned}&\left[P\left(\tau,y,v_1(t,x)\right)-P\left(t,x,v_1(t,x)\right)\right]\lambda_y\\&\quad+P_y\left(\tau,y,v_1(t,x)\right)\lambda\\&\quad=P_\tau\left(\tau,y,v_1(t,x)\right)(\tau-t)\lambda_y\\&\quad+P_y\left(\tau,y,v_1(t,x)\right)(y-x)\lambda_y\\&\quad+P_y\left(\tau,y,v_1(t,x)\right)\lambda+\beta_1\lambda_y\\&\quad=P_\tau\left(\tau,y,v_1(t,x)\right)\left((t-\tau)\lambda\right)_x\\&\quad+P_y\left(\tau,y,v_1(t,x)\right)\left((x-y)\lambda\right)_x+\beta_1\lambda_y.\end{aligned}\tag{38}$$

The functions α_0, α_1, and β_1 in (37) and (38) satisfy

$$\left|\alpha_0\right|+\left|\alpha_1\right|+\left|\beta_1\right|\le d\varepsilon(d),\tag{39}$$

where $d=|t-\tau|+|x-y|$ and $\varepsilon(d)\to 0$ as $d\to 0$. There are $\lambda=\lambda_\varepsilon\equiv 0$ for $|t-\tau|\ge 2\varepsilon$ or $|x-y|\ge 2\varepsilon$ and

$$\begin{aligned}&\left|\lambda_x\right|+\left|\lambda_y\right|\le\frac{c}{\varepsilon^3},\\&\left|\eta(\cdots)-\eta(\tau,y)\right|\le c\left(|t-\tau|+|x-y|\right).\end{aligned}\tag{40}$$

Hence, we get

$$\begin{aligned}A_\varepsilon=&\iiiint_{\omega_T\times\omega_T}\operatorname{sign}\left(v_1(t,x)-v_2(\tau,y)\right)\\&\cdot\left\{\left[P_\tau\left(\tau,y,v_2(\tau,y)\right)-P_\tau\left(t,x,v_2(\tau,y)\right)\right]\right.\\&\cdot\left((\tau-t)\lambda\right)_x\\&+\left[P_y\left(\tau,y,v_2(t,x)\right)-P_y\left(\tau,y,v_1(t,x)\right)\right]\\&\left.\cdot\left((y-x)\lambda\right)_x\right\}\eta(\tau,y)\,dt\,dx\,d\tau\,dy+I(\varepsilon),\end{aligned}\tag{41}$$

where $I(\varepsilon)\to 0$ as $\varepsilon\to 0$.

We denote the integrand in (41) as

$$\begin{aligned}B_\varepsilon=&K_1\left(\tau,y,v_1(t,x),v_2(\tau,y)\right)\left((\tau-t)\eta(\tau,y)\lambda\right)_x\\&+K_2\left(\tau,y,v_1(t,x),v_2(\tau,y)\right)\left((y-x)\eta(\tau,y)\lambda\right)_x\\&=\left\{K_1\left(\tau,y,v_1(t,x),v_2(\tau,y)\right)\right.\\&\left.-K_1\left(\tau,y,v_1(\tau,y),v_2(\tau,y)\right)\right\}\left((\tau-t)\eta(\tau,y)\lambda\right)_x\\&+\left\{K_2\left(\tau,y,v_1(t,x),v_2(\tau,y)\right)\right.\\&\left.-K_2\left(\tau,y,v_1(\tau,y),v_2(\tau,y)\right)\right\}\\&\cdot\left((y-x)\eta(\tau,y)\lambda\right)_x\\&+K_1\left(\tau,y,v_1(\tau,y),v_2(\tau,y)\right)\left((\tau-t)\eta(\tau,y)\lambda\right)_x\\&+K_2\left(\tau,y,v_1(\tau,y),v_2(\tau,y)\right)\left((y-x)\eta(\tau,y)\lambda\right)_x,\end{aligned}\tag{42}$$

where K_1 and K_2 satisfy the Lipschitz condition in v. Applying the property of the function $\eta(\tau,y)\lambda_\varepsilon$ and the method of integration by parts, we have

$$\begin{aligned}&\iiiint_{\omega_T\times\omega_T}\left\{K_1\left(\tau,y,v_1(\tau,y),v_2(\tau,y)\right)\right.\\&\quad\cdot\left((\tau-t)\eta(\tau,y)\lambda\right)_x\\&\quad+K_2\left(\tau,y,v_1(\tau,y),v_2(\tau,y)\right)\\&\quad\left.\cdot\left((y-x)\eta(\tau,y)\lambda\right)_x\right\}dt\,dx\,d\tau\,dy=0.\end{aligned}\tag{43}$$

Hence

$$\begin{aligned}\left|A_\varepsilon-I(\varepsilon)\right|=&\left|\iiiint_{\omega_T\times\omega_T}B_\varepsilon dt\,dx\,d\tau\,dy\right|\le c\\&\cdot\iiiint_{\omega_T\times\omega_T}\left|v_1(t,x)-v_2(\tau,y)\right|\\&\cdot\left\{\left(|\tau-t|+|y-x|\right)\eta(\tau,y)\left|\lambda_x\right|+\lambda_\varepsilon\right\}dt\,dx\,d\tau\,dy\\&\le\frac{c}{\varepsilon^2}\iiiint_{\Pi}\left|v_1(t,x)-v_2(\tau,y)\right|dt\,dx\,d\tau\,dy.\end{aligned}\tag{44}$$

Using Lemma 3, $A_\varepsilon-I(\varepsilon)\to 0$ as $\varepsilon\to 0$. Therefore, we have

$$\lim_{\varepsilon\to 0}\iiiint_{\omega_T\times\omega_T}A_3\,dt\,dx\,d\tau\,dy=0.\tag{45}$$

It follows from (33) to (45) that

$$\begin{aligned}
&\iiiint_{\omega_T\times\omega_T}\{|v_1(t,x)-v_2(\tau,y)|\eta_t \\
&\quad+\operatorname{sign}(v_1(t,x)-v_2(\tau,y)) \\
&\quad\cdot[P(t,x,v_1(t,x))-P(t,x,v_2(\tau,y))]\eta_x\} \\
&\quad\cdot\lambda_\varepsilon dt\,dx\,d\tau\,dy \\
&\quad+\left|\iiiint_{\omega_T\times\omega_T}\operatorname{sign}(v_1(t,x)-v_2(\tau,y))\right. \\
&\quad\cdot[\Psi(t,x,v_1(t,x))-\Psi(\tau,y,v_2(\tau,y))] \\
&\quad\left.\cdot\eta\lambda_\varepsilon dt\,dx\,d\tau\,dy\right|\geq 0.
\end{aligned}\tag{46}$$

We note that the first two terms of the integrand of (46) have the form

$$F_\varepsilon=F(t,x,\tau,y,v_1(t,x),v_2(\tau,y))\lambda_\varepsilon\left(\vdots\right),\tag{47}$$

where F satisfies the Lipschitz condition in all its variables. Then

$$\begin{aligned}
&\iiiint_{\omega_T\times\omega_T}F_\varepsilon dt\,dx\,d\tau\,dy \\
&\quad=\iiiint_{\omega_T\times\omega_T}F(t,x,\tau,y,v_1(t,x),v_2(\tau,y)) \\
&\quad\cdot\lambda_\varepsilon dt\,dx\,d\tau\,dy \\
&\quad=\iiiint_{\omega_T\times\omega_T}\{F(t,x,\tau,y,v_1(t,x),v_2(\tau,y)) \\
&\quad-F(t,x,t,x,v_1(t,x),v_2(t,x))\}\lambda_\varepsilon dt\,dx\,d\tau\,dy \\
&\quad+\iiiint_{\omega_T\times\omega_T}F(t,x,t,x,v_1(t,x),v_2(t,x)) \\
&\quad\cdot\lambda_\varepsilon dt\,dx\,d\tau\,dy=J_{11}(\varepsilon)+J_{12}.
\end{aligned}\tag{48}$$

Note that $F_\varepsilon=0$ outside the region Π. Applying the estimate $\left|\lambda_\varepsilon\left(\vdots\right)\right|\leq c/\varepsilon^2$ and Lemma 4, we get

$$\begin{aligned}
&|J_{11}(\varepsilon)| \\
&\quad\leq c\left[\varepsilon+\frac{1}{\varepsilon^2}\iiiint_\Pi|v_1(t,x)-v_2(\tau,y)|\,dt\,dx\,d\tau\,dy\right],
\end{aligned}\tag{49}$$

where c is a positive constant independent of ε. Using Lemma 3, we know $J_{11}(\varepsilon)\to 0$ as $\varepsilon\to 0$.

For the term J_{12}, we substitute $t=\alpha$, $(t-\tau)/2=\beta$, $x=\zeta$, $(x-y)/2=\gamma$. Combining with the identity

$$\int_{-\varepsilon}^{\varepsilon}\int_{-\infty}^{+\infty}\lambda_\varepsilon(\beta,\gamma)\,d\beta\,d\gamma=1,\tag{50}$$

we obtain

$$\begin{aligned}
J_{12}&=2^2\iint_{\omega_T\times\omega_T}F(\alpha,\zeta,\alpha,\zeta,\phi(\alpha,\zeta),\psi(\alpha,\zeta)) \\
&\quad\cdot\left(\int_{-\varepsilon}^{\varepsilon}\int_{-\infty}^{+\infty}\lambda_\varepsilon(\beta,\gamma)\,d\beta\,d\gamma\right)d\alpha\,d\zeta=4 \\
&\quad\cdot\iint_{\omega_T}F(t,x,t,x,v_1(t,x),v_2(t,x))\,dt\,dx.
\end{aligned}\tag{51}$$

Thus, we have

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\iiiint_{\omega_T\times\omega_T}F_\varepsilon \\
&\quad=4\iint_{\omega_T}F(t,x,t,x,v_1(t,x),v_2(t,x))\,dt\,dx.
\end{aligned}\tag{52}$$

Similarly, the integrand of the third term in (46) can be represented as

$$\begin{aligned}
\overline{F}_\varepsilon&=\operatorname{sign}(v_1(t,x)-v_2(\tau,y)) \\
&\quad\cdot[\Psi(t,x,v_1(t,x))-\Psi(\tau,y,v_2(\tau,y))]g\lambda_\varepsilon \\
&=\overline{F}(t,x,\tau,y,v_1(t,x),v_2(\tau,y))\lambda_\varepsilon\left(\vdots\right).
\end{aligned}\tag{53}$$

Then

$$\begin{aligned}
&\iiiint_{\omega_T\times\omega_T}\overline{F}_\varepsilon dt\,dx\,d\tau\,dy \\
&\quad=\iiiint_{\omega_T\times\omega_T}\{\overline{F}(t,x,\tau,y,v_1(t,x),v_2(\tau,y)) \\
&\quad-\overline{F}(t,x,t,x,v_1(t,x),v_2(t,x))\lambda_\varepsilon\}\,dt\,dx\,d\tau\,dy \\
&\quad+\iiiint_{\omega_T\times\omega_T}\overline{F}(t,x,t,x,v_1(t,x),v_2(t,x)) \\
&\quad\cdot\lambda_\varepsilon dt\,dx\,d\tau\,dy=J_{21}(\varepsilon)+J_{22}.
\end{aligned}\tag{54}$$

It holds that

$$\begin{aligned}
&|J_{21}(\varepsilon)| \\
&\quad\leq c\left[\varepsilon+\frac{1}{\varepsilon^2}\iiiint_\Pi|v_1(t,x)-v_2(\tau,y)|\,dt\,dx\,d\tau\,dy\right].
\end{aligned}\tag{55}$$

Using Lemma 3, it yields $J_{21}(\varepsilon)\to 0$ as $\varepsilon\to 0$. Repeating the steps as before, we have

$$J_{22}=4\iint_{\omega_T}\overline{F}(t,x,t,x,v_1(t,x),v_2(t,x))\,dt\,dx.\tag{56}$$

From (46) to (56), we get

$$\begin{aligned}&\iint_{\omega_T}\{|v_1(t,x)-v_2(t,x)|\eta_t\\&\quad+\operatorname{sign}(v_1(t,x)-v_2(t,x))\\&\quad\cdot[P(t,x,v_1(t,x))-P(\tau,y,v_2(\tau,y))]\eta_x\}\,dt\,dx\\&\quad+\left|\iint_{\omega_T}\operatorname{sign}(v_1(t,x)-v_2(t,x))\right.\\&\quad\left.\cdot[\Psi(t,x,v_1)-\Psi(t,x,v_2)]\eta\,dt\,dx\right|\geq 0.\end{aligned}\tag{57}$$

We set

$$\begin{aligned}g(t)&=\int_{-\infty}^{+\infty}|v_1(t,x)-v_2(t,x)|\,dx,\\\mu_\varepsilon(\sigma)&=\int_{-\infty}^{\sigma}\delta_\varepsilon(\sigma)\,d\sigma.\end{aligned}\tag{58}$$

Take two numbers $\rho,\tau\in(0,T_1)$ and $\rho<\tau$. In (57), we set

$$\begin{aligned}\eta(t,x)&=[\mu_\varepsilon(t-\rho)-\mu_\varepsilon(t-\tau)]\chi(t,x),\\&\qquad\varepsilon<\min(\rho,T_1-\tau),\end{aligned}\tag{59}$$

in which

$$\chi(t,x)=\chi_\theta(t,x)=1-\mu_\theta(|x|+Nt-R_1+\theta),\tag{60}$$

where θ is a small positive constant and $\chi(t,x)=0$ outside the cone Ω. When $\theta\to 0, R_1\to+\infty$, we observe that $\chi_\theta\to 1$. By the definition of the number N, we have

$$0=\chi_t+N|\chi_x|\geq\chi_t+N\chi_x,\quad(t,x)\in\Omega.\tag{61}$$

Applying (57)–(60), we get

$$\begin{aligned}&\iint_{\omega_{T_1}}\{|v_1(t,x)-v_2(t,x)|[\delta_\varepsilon(t-\rho)-\delta_\varepsilon(t-\tau)]\\&\quad\cdot\chi_\theta(t,x)\}\,dt\,dx+\int_0^{T_1}dt\\&\quad\cdot\int_{-\infty}^{+\infty}\{|\Psi(t,x,v_1)-\Psi(t,x,v_2)|\\&\quad\cdot[\mu_\varepsilon(t-\rho)-\mu_\varepsilon(t-\tau)]\chi_\theta(t,x)\}\,dx\geq 0.\end{aligned}\tag{62}$$

In (62), sending $\theta\to 0, R_1\to+\infty$ and using Lemma 2, we obtain

$$\begin{aligned}&\int_0^{T_1}[\delta_\varepsilon(t-\rho)-\delta_\varepsilon(t-\tau)]h(t)\,dt\\&\quad+c\int_0^{T_1}[\mu_\varepsilon(t-\rho)-\mu_\varepsilon(t-\tau)]g(t)\,dt\geq 0,\end{aligned}\tag{63}$$

where c is independent of ε.

Applying the properties of the function δ_ε for $\varepsilon\leq\min(\rho,T_1-\rho)$, we get

$$\begin{aligned}&\left|\int_0^{T_1}\delta_\varepsilon(t-\rho)g(t)-g(\rho)\,dt\right|\\&\quad=\left|\int_0^{T_1}\delta_\varepsilon(t-\rho)[g(t)-g(\rho)]\,dt\right|\\&\quad\leq\frac{c}{\varepsilon}\int_{\rho-\varepsilon}^{\rho+\varepsilon}|g(t)-g(\rho)|\,dt.\end{aligned}\tag{64}$$

Then

$$\int_0^{T_0}\delta_\varepsilon(t-\rho)g(t)\,dt\longrightarrow g(\rho)\quad\text{as }\varepsilon\longrightarrow 0.\tag{65}$$

Let

$$\begin{aligned}W(\rho)&=\int_0^{T_1}\mu_\varepsilon(t-\rho)g(t)\,dt\\&=\int_0^{T_1}dt\int_{-\infty}^{t-\rho}\delta_\varepsilon(\sigma)g(t)\,d\sigma.\end{aligned}\tag{66}$$

We observe that

$$W'(\rho)=-\int_0^{T_1}\delta_\varepsilon(t-\rho)g(t)\,dt.\tag{67}$$

Letting $\varepsilon\to 0$, it derives that

$$\begin{aligned}&W'(\rho)\longrightarrow -g(\rho),\\&W(\rho)\longrightarrow W(0)-\int_0^{\rho}g(t)\,dt,\\&W(\tau)\longrightarrow W(0)-\int_0^{\tau}g(t)\,dt.\end{aligned}\tag{68}$$

Therefore, we have

$$W(\rho)-W(\tau)\longrightarrow\int_\rho^\tau g(t)\,dt\quad\text{as }\varepsilon\longrightarrow 0.\tag{69}$$

From (63)–(69), we obtain inequality

$$g(\rho)+c\int_\rho^\tau g(t)\,dt\geq g(\tau).\tag{70}$$

Choosing $\rho\to 0,\tau\to t$, we get

$$\begin{aligned}&\int_{-\infty}^{+\infty}|\phi(0,x)-\psi(0,x)|\,dx\\&\quad+c\int_0^t dt\int_{-\infty}^{+\infty}|\phi(t,x)-\psi(t,x)|\,dx\\&\quad\geq\int_{-\infty}^{+\infty}|\phi(t,x)-\psi(t,x)|\,dx.\end{aligned}\tag{71}$$

Applying the Gronwall inequality, we complete the proof of Theorem 6. □

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work is supported by National Natural Science Foundation of China (11471263).

References

[1] A. Degasperis and M. Procesi, "Asymptotic integrability," in *Symmetry and Perturbation Theory, Rome, 1998*, vol. 1, pp. 23–37, World Scientific, Singapore, 1999.

[2] A. Degasperis, D. Holm, and A. Hone, "A new integral equation with peakon solutions," *Theoretical and Mathematical Physics*, vol. 133, pp. 1461–1472, 2002.

[3] G. M. Coclite and K. H. Karlsen, "On the well-posedness of the degasperis-procesi equation," *Journal of Functional Analysis*, vol. 233, no. 1, pp. 60–91, 2006.

[4] G. M. Coclite and K. H. Karlsen, "On the uniqueness of discontinuous solutions to the Degasperis-Procesi equation," *Journal of Differential Equations*, vol. 234, no. 1, pp. 142–160, 2007.

[5] J. Escher, Y. Liu, and Z. Y. Yin, "Global weak solutions and blow-up structure for the Degasperis-Procesi equation," *Journal of Functional Analysis*, vol. 241, no. 2, pp. 457–485, 2006.

[6] S. Y. Lai, H. Y. Yan, H. J. Chen, and Y. Wang, "The stability of local strong solutions for a shallow water equation," *Journal of Inequalities and Applications*, vol. 2014, article 410, 2014.

[7] S. Y. Lai and Y. H. Wu, "Global solutions and blow-up phenomena to a shallow water equation," *Journal of Differential Equations*, vol. 249, no. 3, pp. 693–706, 2010.

[8] J. Lenells, "Traveling wave solutions of the Degasperis-Procesi equation," *Journal of Mathematical Analysis and Applications*, vol. 306, no. 1, pp. 72–82, 2005.

[9] X. L. Ai and G. L. Gui, "Global well-posedness for the Cauchy problem of the viscous Degasperis-Procesi equation," *Journal of Mathematical Analysis and Applications*, vol. 361, no. 2, pp. 457–465, 2010.

[10] H. C. Ma, Y. D. Yu, and D. J. Ge, "New exact traveling wave solutions for the modified form of Degasperis-Procesi equation," *Applied Mathematics and Computation*, vol. 203, no. 2, pp. 792–798, 2008.

[11] Y. G. Fu, Z. R. Liu, and X. Y. Yang, "Orbital stability of peakons for the Degasperis-Procesi equation with strong dispersion," *Nonlinear Analysis. Theory, Methods & Applications*, vol. 73, no. 2, pp. 538–546, 2010.

[12] S. N. Kružkov, "First order quasilinear equations in several independent variables," *Mathematics of the USSR-Sbornik*, vol. 10, no. 2, pp. 217–243, 1970.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 429476, 11 pages
http://dx.doi.org/10.1155/2015/429476

Research Article

Robust Regulation and Tracking Control of a Class of Uncertain 2DOF Underactuated Mechanical Systems

David I. Rosas Almeida, Carlos Gamez, and Raul Rascón

Engineering Faculty, Autonomous University of Baja California, 21900 Mexicali, BC, Mexico

Correspondence should be addressed to David I. Rosas Almeida; drosas@uabc.edu.mx

Received 15 May 2015; Accepted 2 August 2015

Academic Editor: Wenguang Yu

A strategy to design and implement a robust controller for a class of underactuated mechanical systems, with two degrees of freedom, which solves the problems of regulation and trajectory tracking, is proposed. This control strategy considers the partial measurement of the state vector and the presence of parametric uncertainties in the plant; these conditions are common in the implementation of a control system. The strategy is based on the use of robust finite time convergence observers to estimate the unmeasured state variables, unknown disturbances, and other signals needed for the control system implementation. The performance of the control strategy is illustrated numerically and experimentally.

1. Introduction

Antecedents and Motivation. Control of underactuated mechanical systems, systems with fewer number of control inputs than their degrees of freedom, has received much attention in the last decades. This is because of the theoretical challenges as well as practical applicability; robots, aerospace vehicles, underwater vehicles, and surface vessels are some examples of underactuated mechanical systems. Some important papers which address this control problem for different situations are [1–9]. While many interesting techniques and results have been presented for this class of systems, the control of them still remains an open problem. Important issues are as follows: how control models can be formulated for such systems and how closed-loop control problems can be solved and implemented. These issues are addressed in this paper for a particular class of uncertain underactuated mechanical systems. These problems have been addressed by many authors and important solutions have been proposed, some of which are as follows.

In [10] a sliding mode control method for a class of second-order underactuated mechanical systems is proposed; the controller has the double-layer structure. Firstly, the system states are divided into several different subsystems. For each of these subsystems, a first-layer sliding plane is constructed; from that, a second-layer sliding plane is constructed. By analyzing the features of the model of the plant, they derive the sliding control law. Here, the proposed controller only solves the regulation problem; furthermore, the implementation of the controller requires the measurement of full state vector; this condition is not satisfied in practice. For a similar class of systems, in [11], the Olfati transformation is applied first to represent the system into a special cascade form. Since, in general cases, some of the terms in the new space might become too complex to drive, they are regarded as uncertainties. A backstepping-like adaptive control based on function approximation technique is designed so that the system in the new space can be stabilized with uniformly ultimately bounded performance. This paper assumes knowledge of all system parameters and the measurement of all state variables and the perturbation terms that appear in the approach are well known, but experimental results are not presented.

Other works deal with particular systems, for example, [1, 3, 12, 13], but many of them only deal with the regulation problem and present performance results through numerical simulations. Reference [14] addresses the observer-based multivariable control of a class of nonlinear, underactuated Lagrangian systems with application to trajectory tracking and sway control of a 3D overhead gantry crane subject to

Coulomb friction. A second-order sliding mode observer is used for the estimation of velocities. Based on these estimates, the sliding function of a second-order sliding mode controller for trajectory tracking and antiswing control is proposed. This is a very important paper because it considers a very common situation in practice, the lack of measurement of the velocities, but it only show numerical results.

An important work is presented in [15] where, for a class of second-order underactuated mechanical systems, a robust finite time control strategy is proposed. The robust finite time controller drives the tracking error to be zero at the fixed final time. By utilizing a Lyapunov stability theorem, the controller can achieve finite time tracking of desired reference signals for underactuated systems, which are subject to both external disturbances and system uncertainties. However, the complete measurement of the state vector is assumed and only stabilization problem is solved. Moreover, illustration controller performance is through numerical simulations.

Main Contribution. We propose a strategy to design and implement a robust controller for a class of underactuated mechanical systems, with two degrees of freedom, which solves the problems of regulation and trajectory tracking. This control strategy considers the partial measurement of the state vector and the presence of parametric uncertainties in the plant; these conditions are common in the implementation of a control system.

The strategy is based on the use of robust finite time convergence observers to estimate the unmeasured state variables, unknown disturbances, and other signals needed for the control system implementation. The performance of the control strategy is illustrated numerically and experimentally.

Paper Structure. This paper is organized as follows: Section 2 provides the control problem, the model of the plant, and the control objective. In Section 3, we propose the solution to the problem; to implement such solution is necessary to know the velocities, the exact value of the disturbances, and the availability of auxiliary signals and their derivatives, which are unknown. One way to implement this control signal is presented in Section 4, where with the help of robust observers with finite time convergence we estimate all the terms needed for implementation. Section 5 shows the performance of the controller through numerical simulations and experimental results. Finally, in Section 6, we present some general conclusions.

2. Problem Statement

Consider a 2DOF underactuated mechanical system whose dynamics are given by

$$\begin{aligned} \ddot{q}_1 &= f_1(q_1, \dot{q}_1, q_2, \dot{q}_2) + g_1(q_1, q_2)u + \gamma_1(\cdot), \\ \ddot{q}_2 &= f_2(q_2, \dot{q}_2) + g_2 v(q_1) + \gamma_2(\cdot), \\ y_1 &= q_1, \\ y_2 &= q_2, \end{aligned} \tag{1}$$

where q_1, q_2, $\dot{q}_1$, and $\dot{q}_2$ are the generalized positions and velocities, respectively. $f_1(\cdot)$, $f_2(\cdot)$, and $g_1(\cdot)$ are smooth functions; $g_1(\cdot) \neq 0$ for all q_1 and q_2. $g_2 \neq 0$ is a constant, $v(q_1)$ is an invertible function for all q_1 in the domain of operation of the system, and u is the control input. $\gamma_1(\cdot)$ and $\gamma_2(\cdot)$ are smooth terms due to parameter variations; based on Lagrangian model (1), these terms may depend on q_1, q_2, $\dot{q}_1$, $\dot{q}_2$, $\ddot{q}_1$, and $\ddot{q}_2$, but if these variables are bounded, $\gamma_1(\cdot)$ and $\gamma_2(\cdot)$ also are bounded [16]. Additionally, we considered that there is no measurement of the velocities $\dot{q}_1$ and $\dot{q}_2$.

Some well known mechanisms that belong to this class of underactuated systems are the mass-spring-damper, magnetic suspension, and the ball and beam systems.

A state space representation of system (1) is

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= f_1(x_1, x_2, x_3, x_4) + g_1(x_1, x_2)u + \gamma_1(\cdot), \\ \dot{x}_3 &= x_4, \\ \dot{x}_4 &= f_2(x_3, x_4) + g_2 v(x_1) + \gamma_2(\cdot), \\ y_1 &= x_1, \\ y_2 &= x_3. \end{aligned} \tag{2}$$

The control problem, for system (2), is to design a control input u such that the underactuated position x_3 tracks a reference signal $y_r(t)$ in asymptotic form; in other words

$$\lim_{t \to \infty} |y_2 - y_r(t)| = 0, \tag{3}$$

where $y_r(t)$ is a $\mathscr{C}^k$ function, for a sufficiently large k.

To solve the control problem we define the error variables $e_1 = x_3 - y_r(t)$ and $e_2 = x_4 - \dot{y}_r(t)$, whose dynamics are given by

$$\dot{e}_1 = e_2, \tag{4}$$

$$\dot{e}_2 = f_2(e_1 + y_r, e_2 + \dot{y}_r) + g_2 v(x_1) + \gamma_2(\cdot) - \ddot{y}_r(t), \tag{5}$$

$$\dot{x}_1 = x_2, \tag{6}$$

$$\begin{aligned} \dot{x}_2 = {} & f_1(x_1, x_2, e_1 + y_r, e_2 + \dot{y}_r) + g_1(x_1, x_2)u \\ & + \gamma_1(\cdot), \end{aligned} \tag{7}$$

$$y_e = e_1, \tag{8}$$

$$y_1 = x_1. \tag{9}$$

Now we can say that the control problem is to design a control input u such that the origin of the error variables of subsystems (4) and (5) will be an asymptotic stable equilibrium point, while the variables x_1 and x_2 stay bounded.

3. Control Strategy

In this section we present a strategy to solve the control problem considering that every disturbance terms and velocities are known; the next section will show its implementation.

System ((4)–(9)) is formed by two subsystems; the unactuated part is

$$\begin{aligned} \dot{e}_1 &= e_2, \\ \dot{e}_2 &= f_2\left(e_1 + y_r, e_2 + \dot{y}_r\right) + g_2 v\left(x_1\right) + \gamma_2(\cdot) - \ddot{y}_r(t), \\ y_e &= e_1, \end{aligned} \tag{10}$$

and the actuated part is

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= f_1\left(x_1, x_2, e_1 + y_r, e_2 + \dot{y}_r\right) + g_1\left(x_1, x_2\right) u \\ &\quad + \gamma_1(\cdot), \\ y_1 &= x_1. \end{aligned} \tag{11}$$

The function $v(x_1)$ in (10) can be seen as a control input for this subsystem; therefore we rename $v(x_1)$ as u_e or $x_1 = v^{-1}(u_e)$ in (10):

$$\begin{aligned} \dot{e}_1 &= e_2, \\ \dot{e}_2 &= f_2\left(e_1 + y_r, e_2 + \dot{y}_r\right) + \gamma_2(\cdot) - \ddot{y}_r(t) + g_2 u_e, \\ y_e &= e_1. \end{aligned} \tag{12}$$

An ideal control u_e that stabilizes the origin of system (12) is

$$\begin{aligned} u_e = \frac{1}{g_2}&\left(-f_2\left(e_1 + y_r, e_2 + \dot{y}_r\right) - \gamma_2(\cdot) + \ddot{y}_r(t) - k_1 e_1\right. \\ &\left. - k_2 e_2\right); \end{aligned} \tag{13}$$

substituting it in (12) results in

$$\begin{aligned} \dot{e}_1 &= e_2, \\ \dot{e}_2 &= -k_1 e_1 - k_2 e_2, \\ y_e &= e_1. \end{aligned} \tag{14}$$

If k_1 and k_2 are positive constants the origin of system (14) is an exponentially stable equilibrium point.

Now $v^{-1}(u_e)$ must be the reference signal for the position x_1 in (11). Define $x_{\mathrm{re}} = v^{-1}(u_e)$ and new error variables $\varepsilon_1 = x_1 - x_{\mathrm{re}}$ and $\varepsilon_2 = x_2 - \dot{x}_{\mathrm{re}}$ whose dynamics are given by

$$\begin{aligned} \dot{\varepsilon}_1 &= \varepsilon_2, \\ \dot{\varepsilon}_2 &= f_1\left(x_1, x_2, e_1 + y_r, e_2 + \dot{y}_r\right) + g_1\left(x_1, x_2\right) u + \gamma_1(\cdot) \\ &\quad - \ddot{x}_{\mathrm{re}}. \end{aligned} \tag{15}$$

A control u that stabilizes the origin of system (15) is

$$\begin{aligned} u = \frac{1}{g_1\left(x_1, x_2\right)}&\left(-f_1\left(x_1, x_2, e_1 + y_r, e_2 + \dot{y}_r\right) - \gamma_1(\cdot)\right. \\ &\left. + \ddot{x}_{\mathrm{re}} - \alpha_1 \varepsilon_1 - \alpha_2 \varepsilon_2 - \alpha_3 \operatorname{sign}\left(\varepsilon_1\right)\right); \end{aligned} \tag{16}$$

substituting it in (15) we have

$$\begin{aligned} \dot{\varepsilon}_1 &= \varepsilon_2, \\ \dot{\varepsilon}_2 &= -\alpha_1 \varepsilon_1 - \alpha_2 \varepsilon_2 - \alpha_3 \operatorname{sign}\left(\varepsilon_1\right). \end{aligned} \tag{17}$$

If the constants α_1, α_2, and α_3 are positive the origin of system (17) is an exponentially stable equilibrium point.

It is important to note that, because this section considered that we have the measurement of all terms, it is not necessary to incorporate the term discontinuous in the control (16); however this term is very useful when the implementation is done because it will give robustness to closed-loop system.

3.1. Stability Analysis. To prove the stability of the closed-loop system consider the error system

$$\begin{aligned} \dot{e}_1 &= e_2, \\ \dot{e}_2 &= f_2\left(e_1 + y_r, e_2 + \dot{y}_r\right) + g_2 v\left(x_1\right) + \gamma_2(\cdot) - \ddot{y}_r(t), \\ \dot{\varepsilon}_1 &= \varepsilon_2, \\ \dot{\varepsilon}_2 &= -\alpha_1 \varepsilon_1 - \alpha_2 \varepsilon_2 - \alpha_3 \operatorname{sign}\left(\varepsilon_1\right). \end{aligned} \tag{18}$$

The last two equations of (18) form a subsystem uncoupled of the e_1 and e_2, regardless of the value of u_e. Then, if α_1, α_2, and α_3 are positive constants, the origin of this subsystem is exponentially stable [17].

Because x_1 converges exponentially to u_e, x_1 can be expressed in the form

$$\begin{aligned} x_1 = v^{-1}\Bigg(\frac{1}{g_2}&\left(-f_2\left(e_1 + y_r, e_2 + \dot{y}_r\right) - \gamma_2(\cdot) + \ddot{y}_r(t)\right. \\ &\left. - k_1 e_1 - k_2 e_2\right)\Bigg) + \Delta(\cdot), \end{aligned} \tag{19}$$

where $\Delta(\cdot)$ is the difference between x_1 and x_{re} and satisfies

$$\left|\Delta(\cdot)\right| \leq \rho_0 e^{-\sigma_0 t}, \tag{20}$$

for some positive constants ρ_0 and σ_0. Substituting (19) in (18) we have

$$\begin{aligned} \dot{e}_1 &= e_2, \\ \dot{e}_2 &= f_2\left(e_1 + y_r, e_2 + \dot{y}_r\right) + \gamma_2(\cdot) - \ddot{y}_r(t) \\ &\quad + g_2 v\Bigg(v^{-1}\Bigg(\frac{1}{g_2}\left(-f_2\left(e_1 + y_r, e_2 + \dot{y}_r\right) - \gamma_2(\cdot)\right. \\ &\quad \left. + \ddot{y}_r(t) - k_1 e_1 - k_2 e_2\right)\Bigg) + \Delta(\cdot)\Bigg), \\ \dot{\varepsilon}_1 &= \varepsilon_2, \\ \dot{\varepsilon}_2 &= -\alpha_1 \varepsilon_1 - \alpha_2 \varepsilon_2 - \alpha_3 \operatorname{sign}\left(\varepsilon_1\right). \end{aligned} \tag{21}$$

This system can be rewritten in the form

$$\begin{aligned}\dot{e}_1 &= e_2,\\ \dot{e}_2 &= -k_1 e_1 - k_2 e_2 + \mu(\cdot),\\ \dot{\varepsilon}_1 &= \varepsilon_2,\\ \dot{\varepsilon}_2 &= -\alpha_1 \varepsilon_1 - \alpha_2 \varepsilon_2 - \alpha_3 \operatorname{sign}(\varepsilon_1),\end{aligned} \tag{22}$$

where $\mu(\cdot)$ is a term produced by $\Delta(\cdot)$ that satisfies

$$|\mu(\cdot)| \le \rho_1 e^{-\sigma_1 t}, \tag{23}$$

for some positive constants ρ_1 and σ_1. System (22) is a piecewise linear system with a vanishing disturbance; therefore, there exist a set of constants k_1, k_2, α_1, α_2, and α_3 such that the origin will be an asymptotically stable equilibrium point in a sufficient large region $\Omega \subset \Re^4$ [17, 18].

4. Controller Implementation

The control given in (13) and (16) cannot be implemented directly because e_2, ε_2, $\gamma_1(\cdot)$, $\gamma_2(\cdot)$, and $\ddot{x}_{re}$ are not available. In this section we present a method to implement this control input. It is based on discontinuous observers with finite time convergence.

4.1. Estimation of e_2 and $\gamma_2(\cdot)$. To estimate e_2 and $\gamma_2(\cdot)$ we propose a finite time observer for the underactuated part:

$$\begin{aligned}\dot{e}_1 &= e_2,\\ \dot{e}_2 &= f_2(e_1 + y_r, e_2 + \dot{y}_r) + g_2 v(x_1) + \gamma_2(\cdot) - \ddot{y}_r(t),\\ y_e &= e_1.\end{aligned} \tag{24}$$

The following observer is based on Levant's exact deriver [19]:

$$\begin{aligned}\dot{\hat{e}}_1 &= z_1 + c_{1,1}|e_1 - \hat{e}_1|^{1/2} \operatorname{sign}(e_1 - \hat{e}_1),\\ \dot{z}_1 &= c_{2,1} \operatorname{sign}(e_1 - \hat{e}_1),\\ \hat{y}_1 &= \hat{e}_1,\\ \dot{\hat{e}}_2 &= g_2 v(x_1) - \ddot{y}_r(t) + z_2\\ &\quad + c_{1,2}|z_1 - \hat{e}_2|^{1/2} \operatorname{sign}(z_1 - \hat{e}_2),\\ \dot{z}_2 &= c_{2,2} \operatorname{sign}(z_1 - \hat{e}_2),\\ \hat{y}_2 &= \hat{e}_2.\end{aligned} \tag{25}$$

To show the stability of the observer define the errors $\epsilon_1 = e_1 - \hat{e}_1$ and $\epsilon_2 = e_2 - \hat{e}_2$, whose dynamics are given by

$$\begin{aligned}\dot{\epsilon}_1 &= e_2 - z_1 - c_{1,1}|\epsilon_1|^{1/2} \operatorname{sign}(\epsilon_1),\\ \dot{z}_1 &= c_{2,1} \operatorname{sign}(\epsilon_1),\\ \dot{\epsilon}_2 &= f_2(e_1 + y_r, e_2 + \dot{y}_r) + \gamma_2(\cdot) - z_2\\ &\quad - c_{1,2}|z_1 - \hat{e}_2|^{1/2} \operatorname{sign}(z_1 - \hat{e}_2),\\ \dot{z}_2 &= c_{2,2} \operatorname{sign}(z_1 - \hat{e}_2).\end{aligned} \tag{26}$$

Making a change of variables in the first two equations

$$\begin{aligned}v_1 &= \epsilon_1,\\ v_2 &= e_2 - z_1,\end{aligned} \tag{27}$$

we obtain the subsystem

$$\begin{aligned}\dot{v}_1 &= v_2 - c_{1,1}|v_1|^{1/2} \operatorname{sign}(v_1),\\ \dot{v}_2 &= \dot{e}_2 - c_{2,1} \operatorname{sign}(v_1).\end{aligned} \tag{28}$$

If $|\dot{e}_2| \le \delta_1$, where δ_1 is a known constant, there exist constants $c_{1,1}$ and $c_{2,1}$ such that the trajectories converge in finite time to ($v_1 = 0$, $v_2 = 0$), [20]; therefore

$$z_1 = e_2 \tag{29}$$

in finite time.

The criteria to choosing the constants $c_{1,1}$ and $c_{2,1}$ are

$$\begin{aligned}c_{1,1} &> \sqrt{\frac{2}{c_{2,1} + \delta_1}} \frac{(c_{2,1} + \delta_1)(1 + p)}{(1 - p)},\\ c_{2,1} &> \delta_1,\end{aligned} \tag{30}$$

where $0 < p < 1$.

For the last two equations in (26),

$$\begin{aligned}\dot{\epsilon}_2 &= f_2(e_1 + y_r, e_2 + \dot{y}_r) + \gamma_2(\cdot) - z_2\\ &\quad - c_{1,2}|\epsilon_2 - v_2|^{1/2} \operatorname{sign}(\epsilon_2 - v_2),\\ \dot{z}_2 &= c_{2,2} \operatorname{sign}(\epsilon_2 - v_2).\end{aligned} \tag{31}$$

Making a change of variables

$$\begin{aligned}v_3 &= \epsilon_2 - v_2,\\ v_4 &= f_2(e_1 + y_r, e_2 + \dot{y}_r) + \gamma_2(\cdot) - z_2 - \dot{v}_2,\end{aligned} \tag{32}$$

the dynamics of these variables are given by

$$\begin{aligned}\dot{v}_3 &= v_4 - c_{1,2}|v_3|^{1/2} \operatorname{sign}(v_3),\\ \dot{v}_4 &= \dot{f}_2(e_1 + y_r, e_2 + \dot{y}_r) + \dot{\gamma}_2(\cdot) - \ddot{v}_2 - c_{2,2} \operatorname{sign}(v_3).\end{aligned} \tag{33}$$

If $|\dot{f}_2(x_3, x_4) + \dot{\gamma}_2(\cdot) - \ddot{v}_2| \le \delta_2$ there exist constants $c_{1,2}$ and $c_{2,2}$ such that the trajectories converge in finite time to ($v_3 = 0$, $v_4 = 0$); therefore

$$\begin{aligned}0 &= \epsilon_2 - v_2,\\ 0 &= f_2(e_1 + y_r, e_2 + \dot{y}_r) + \gamma_2(\cdot) - z_2 - \dot{v}_2.\end{aligned} \tag{34}$$

Considering that for (28) $v_2 = 0$ in finite time, after this time, we estimate the velocity error e_2 and disturbance terms $f_2(e_1 + y_r, e_2 + \dot{y}_r) + \gamma_2(\cdot)$:

$$\begin{aligned}\hat{e}_2 &= e_2,\\ z_2 &= f_2(e_1 + y_r, e_2 + \dot{y}_r) + \gamma_2(\cdot).\end{aligned} \tag{35}$$

As we can see $\gamma_2(\cdot)$ is estimated through z_2.

4.2. Estimation of ε_2, $\gamma_1(\cdot)$, and $\ddot{x}_{re}$. Now we design an observer for system (15):

$$\begin{aligned}\dot{\varepsilon}_1 &= \varepsilon_2,\\ \dot{\varepsilon}_2 &= f_1(x_1, x_2, e_1 + y_r, e_2 + \dot{y}_r) + g_1(x_1, x_3)u\\ &\quad + \gamma_1(\cdot) - \ddot{x}_{\mathrm{re}},\\ y_\varepsilon &= \varepsilon_1.\end{aligned} \tag{36}$$

The observer is

$$\begin{aligned}\dot{\hat{\varepsilon}}_1 &= w_1 + a_{1,1}\left|\varepsilon_1 - \hat{\varepsilon}_1\right|^{1/2}\operatorname{sign}(\varepsilon_1 - \hat{\varepsilon}_1),\\ \dot{w}_1 &= a_{2,1}\operatorname{sign}(\varepsilon_1 - \hat{\varepsilon}_1),\\ \hat{y}_1 &= \hat{\varepsilon}_1,\\ \dot{\hat{\varepsilon}}_2 &= g_1(x_1, x_3)u + w_2\\ &\quad + a_{1,2}\left|w_1 - \hat{\varepsilon}_2\right|^{1/2}\operatorname{sign}(w_1 - \hat{\varepsilon}_2),\\ \dot{w}_2 &= a_{2,2}\operatorname{sign}(w_1 - \hat{\varepsilon}_2),\\ \hat{y}_2 &= \hat{\varepsilon}_2.\end{aligned} \tag{37}$$

To show the stability of the observer define the errors $\epsilon_3 = \varepsilon_1 - \hat{\varepsilon}_1$ and $\epsilon_4 = \varepsilon_2 - \hat{\varepsilon}_2$, whose dynamics are given by

$$\dot{\epsilon}_3 = \varepsilon_2 - w_1 - a_{1,1}\left|\epsilon_3\right|^{1/2}\operatorname{sign}(\epsilon_3), \tag{38}$$

$$\dot{w}_1 = a_{2,1}\operatorname{sign}(\epsilon_3), \tag{39}$$

$$\begin{aligned}\dot{\epsilon}_4 &= f_1(x_1, x_2, e_1 + y_r, e_2 + \dot{y}_r) + \gamma_1(\cdot) - \ddot{u}_e - w_2\\ &\quad - a_{1,2}\left|w_1 - \hat{\varepsilon}_2\right|^{1/2}\operatorname{sign}(w_1 - \hat{\varepsilon}_2),\end{aligned} \tag{40}$$

$$\dot{w}_2 = a_{2,2}\operatorname{sign}(w_1 - \hat{\varepsilon}_2). \tag{41}$$

Making a change of variables for the first two equations ((38)-(39)),

$$\begin{aligned}v_5 &= \epsilon_3,\\ v_6 &= \varepsilon_2 - w_1,\end{aligned} \tag{42}$$

we obtain the subsystem

$$\begin{aligned}\dot{v}_5 &= v_6 - a_{1,1}\left|v_5\right|^{1/2}\operatorname{sign}(v_5),\\ \dot{v}_6 &= \dot{\varepsilon}_2 - a_{2,1}\operatorname{sign}(v_5).\end{aligned} \tag{43}$$

According to (30), if $|\dot{\varepsilon}_2| \le \delta_3$ there exist constants $a_{1,1}$ and $a_{2,1}$ such that the trajectories converge in finite time to ($v_5 = 0$, $v_6 = 0$) [20]; therefore

$$w_1 = \varepsilon_2, \tag{44}$$

in finite time.

For the last two equations in ((40)-(41)),

$$\begin{aligned}\dot{\epsilon}_4 &= f_1(x_1, x_2, e_1 + y_r, e_2 + \dot{y}_r) + \gamma_1(\cdot) - \ddot{x}_{\mathrm{re}} - w_2\\ &\quad - a_{1,2}\left|\epsilon_4 - v_6\right|^{1/2}\operatorname{sign}(\epsilon_4 - v_6),\\ \dot{w}_2 &= a_{2,2}\operatorname{sign}(\epsilon_4 - v_6).\end{aligned} \tag{45}$$

Making a change of variables

$$\begin{aligned}v_7 &= \epsilon_4 - v_6,\\ v_8 &= f_1(x_1, x_2, e_1 + y_r, e_2 + \dot{y}_r) + \gamma_1(\cdot) - \ddot{x}_{\mathrm{re}} - \dot{v}_6\\ &\quad - w_2,\end{aligned} \tag{46}$$

the dynamics of these variables are given by

$$\begin{aligned}\dot{v}_7 &= v_8 - a_{1,2}\left|v_7\right|^{1/2}\operatorname{sign}(v_7),\\ \dot{v}_8 &= \dot{f}_1(x_1, x_2, x_3, x_4) + \dot{\gamma}_1(\cdot) - \dddot{x}_{\mathrm{re}} - \ddot{v}_6\\ &\quad - a_{2,2}\operatorname{sign}(v_7).\end{aligned} \tag{47}$$

If $|\dot{f}_1(x_1, x_2, e_1 + y_r, e_2 + \dot{y}_r) + \dot{\gamma}_1(\cdot) - \dddot{x}_{\mathrm{re}} - \ddot{v}_6| \le \delta_4$ there exist constants $a_{1,2}$ and $a_{2,2}$ such that the trajectories converge in finite time to ($v_7 = 0$, $v_8 = 0$) [20]; therefore

$$\begin{aligned}0 &= \epsilon_4 - v_6,\\ 0 &= f_1(x_1, x_2, e_1 + y_r, e_2 + \dot{y}_r) + \gamma_1(\cdot) - \ddot{x}_{\mathrm{re}} - \dot{v}_6\\ &\quad - w_2.\end{aligned} \tag{48}$$

Considering that for (43) $v_6 = 0$ in finite time, after a finite time, we have

$$\begin{aligned}\hat{\varepsilon}_2 &= \varepsilon_2,\\ w_2 &= f_1(x_1, x_2, e_1 + y_r, e_2 + \dot{y}_r) + \gamma_1(\cdot) - \ddot{x}_{\mathrm{re}}.\end{aligned} \tag{49}$$

Now, the control inputs (13) and (16) may be implemented in the following form:

$$\begin{aligned}x_{\mathrm{re}} &\approx \frac{1}{g_2}\left(-z_2 + \ddot{y}_r(t) - k_1 e_1 - k_2 \hat{\varepsilon}_2\right),\\ u &\approx \frac{1}{g_1(x_1, x_3)}\left(-w_2 - \alpha_1 \varepsilon_1 - \alpha_2 \hat{\varepsilon}_2 - \alpha_3 \operatorname{sign}(\varepsilon_1)\right).\end{aligned} \tag{50}$$

The implementation of the controller must be in several stages. First we have to apply a signal u, in open loop, such that the behavior of the system will be bounded; in this way the observers can estimate the state and disturbances in finite time. After this time, the control signals (50) can be implemented and then close the control loop.

5. Control System Performance

This section shows the performance of the control system through numerical simulations and experimental results; the control systems are a ball and beam system and a spring-mass-damper mechanism.

5.1. Control of a Ball and Beam System. Consider the ball and beam system shown in Figure 1; its model is given by

$$\begin{aligned}\left(J + mx^2\right)\ddot{\alpha} + 2mx\dot{\alpha}\dot{x} - (mgx)\cos(\alpha) + \delta_1\dot{\alpha} &= u,\\ \frac{7}{5}\ddot{x} - x\dot{\alpha}^2 - g\sin(\alpha) + \delta_2\dot{x} &= 0,\end{aligned} \tag{51}$$

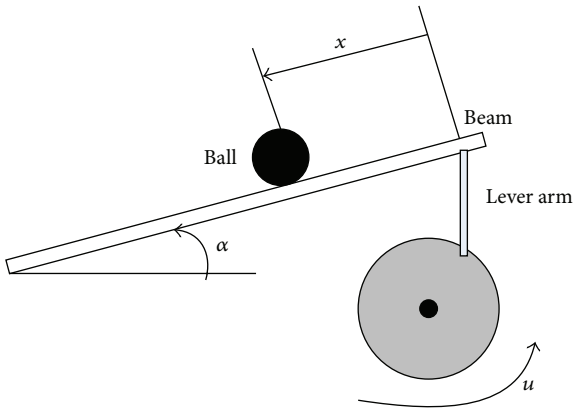


Figure 1: Ball and beam mechanical system.

where x is the position of the ball, α is the angle of the frame, $J = 0.032\,\text{kg}\cdot\text{m}^2$ is the moment of inertia of the beam, $m = 0.06\,\text{kg}$ is the mass of the ball, δ_1 and δ_2 are viscous friction coefficients, and $g = 9.8\,\text{m/seg}^2$ is the gravitational force. Defining the state variables as $x_1 = \alpha$, $x_2 = \dot{\alpha}$, $x_3 = x$, and $x_4 = \dot{x}$ and substituting the values of the constants we have the model

$$\begin{aligned}\ddot{\alpha} &= \frac{mx\dot{\alpha}\dot{x}}{(J+mx^2)} + \frac{(mgx)\cos(\alpha)}{(J+mx^2)} - \frac{\delta_1}{(J+mx^2)}\dot{\alpha} \\ &\quad + \frac{u}{(J+mx^2)}, \\ \ddot{x} &= \frac{5}{7}x\dot{\alpha}^2 - \frac{5}{7}\delta_2\dot{x} + \frac{5}{7}g\sin(\alpha).\end{aligned} \tag{52}$$

A state variable representation is as follows:

$$\begin{aligned}\dot{x}_1 &= x_2, \\ \dot{x}_2 &= \frac{mx_3x_2x_4}{(J+mx_3^2)} + \frac{(mgx_3)\cos(x_1)}{(J+mx_3^2)} - \frac{\delta_1}{(J+mx^2)}x_2 \\ &\quad + \frac{u}{(J+mx_3^2)}, \\ \dot{x}_3 &= x_4, \\ \dot{x}_4 &= \frac{5}{7}x_3x_2^2 - \frac{5}{7}\delta_2x_4 + \frac{5}{7}g\sin(x_1), \\ y_1 &= x_1, \\ y_2 &= x_3.\end{aligned} \tag{53}$$

In this example, without loss of generality, the model is free from uncertainties and external disturbances. It is important to note that the variable x_1 is the argument of the sine function, so the control will have a bounded amplitude.

The control objective is that the ball position x_3 tracks the reference signal $y_r(t)$. Define the error variables $e_1 = x_3 - y_r(t)$ and $e_2 = x_4 - \dot{y}_r(t)$ and the auxiliary control $u_e = \sin(x_1)$ to obtain the following system:

$$\begin{aligned}\dot{x}_1 &= x_2, \\ \dot{x}_2 &= \frac{mx_3x_2x_4}{(J+mx_3^2)} + \frac{(mgx_3)\cos(x_1)}{(J+mx_3^2)} - \frac{\delta_1}{(J+mx^2)}x_2 \\ &\quad + \frac{u}{(J+mx_3^2)}, \\ \dot{e}_1 &= e_2, \\ \dot{e}_2 &= \frac{5}{7}x_3x_2^2 - \frac{5}{7}\delta_2x_4 - \ddot{y}_r(t) + \frac{5}{7}gu_e, \\ y_1 &= x_1, \\ y_e &= e_1,\end{aligned} \tag{54}$$

where u_e only may take values in the $[-1, 1]$ interval.

The ideal controller u_e to stabilize the origin of system (54) is

$$u_e = \frac{7}{5g}\left(-\frac{5}{7}x_3x_2^2 + \frac{5}{7}\delta_2x_4 + \ddot{y}_r(t) - k_1e_1 - k_2e_2\right), \tag{55}$$

where $k_1 = 30$ and $k_2 = 10$. Thus, the reference signal for x_1 is

$$\begin{aligned}x_{\text{re}} = \arcsin\Bigg(&\frac{7}{5g}\left(-\frac{5}{7}x_3x_2^2 + \frac{5}{7}\delta_2x_4 + \ddot{y}_r(t) - k_1e_1\right. \\ &\left.- k_2e_2\right)\Bigg).\end{aligned} \tag{56}$$

Define new error variables $\varepsilon_1 = x_1 - x_{\text{re}}$ and $\varepsilon_2 = x_2 - \dot{x}_{\text{re}}$ whose dynamics are given by

$$\begin{aligned}\dot{\varepsilon}_1 &= \varepsilon_2, \\ \dot{\varepsilon}_2 &= \frac{mx_3x_2x_4}{(J+mx_3^2)} + \frac{(mgx_3)\cos(x_1)}{(J+mx_3^2)} - \frac{\delta_1}{(J+mx^2)}x_2 \\ &\quad - \ddot{x}_{\text{re}} + \frac{u}{(J+mx_3^2)}.\end{aligned} \tag{57}$$

Then, an ideal control to stabilize the origin of system (57) is

$$\begin{aligned}u &= (J+mx_3^2)\left(-\frac{mx_3x_2x_4}{(J+mx_3^2)} - \frac{(mgx_3)\cos(x_1)}{(J+mx_3^2)}\right. \\ &\quad \left. + \frac{\delta_1}{(J+mx^2)}x_2 + \ddot{x}_{\text{re}} + u_o\right), \\ u_o &= -\sigma_1\varepsilon_1 - \sigma_2\varepsilon_2 - \sigma_3\,\text{sign}(\varepsilon_1),\end{aligned} \tag{58}$$

where $\sigma_1 = 40$, $\sigma_2 = 10$, and $\sigma_3 = 0.7$. Figure 2 shows the results when the reference y_r takes different constant values; this is the case of regulation. Steady state error is practically zero and the control signal takes values suitable for a possible implementation.

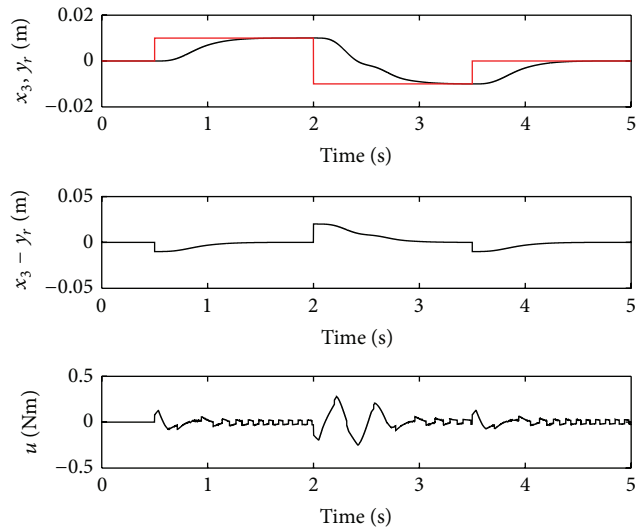


FIGURE 2: Simulation results with the ball and beam system; performance for the regulation problem.

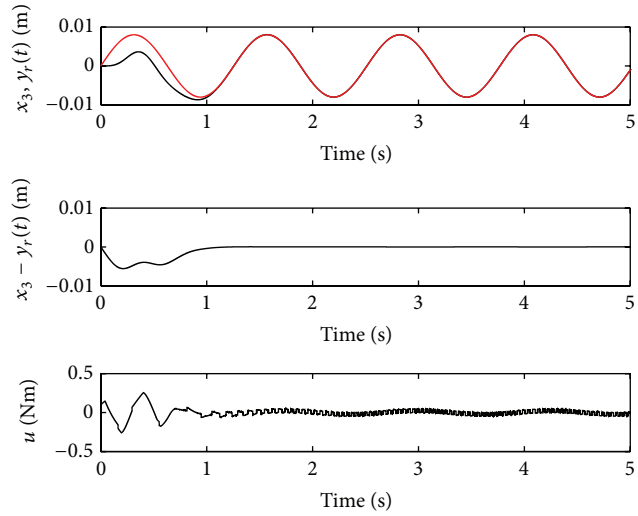


FIGURE 3: Simulation results with the ball and beam system; performance for the tracking problem.

The same situation occurs when the reference is a time-varying signal. Figure 3 shows that the output signal x_3 of the nonactuated link converges to a time-varying signal with an almost zero steady-state error and a control signal with adequate performance for experimental implementation.

5.2. Control of a Mass-Spring-Damper System. Consider the 2DOF underactuated mass-spring-damper mechanical system shown in Figure 4, with the model

$$\begin{aligned} m_1\ddot{x} &= -k_1 x - \delta_1 \dot{x} + (z - x) k_2 + (\dot{z} - \dot{x}) \delta_2 + k_m u \\ &\quad + \gamma_1(\cdot), \\ m_2\ddot{z} &= -(z - x) k_2 - (\dot{z} - \dot{x}) \delta_2 + \gamma_2(\cdot), \end{aligned} \tag{59}$$

where x, $\dot{x}$, and $\ddot{x}$ are the position, velocity, and acceleration of the first mass, z, $\dot{z}$, and $\ddot{z}$ are the position, velocity, and acceleration of the second mass, u is the control input, and $\gamma_1(\cdot)$ and $\gamma_2(\cdot)$ are disturbances that include terms produced by parameter uncertainties.

The nominal parameters are $k_1 = k_2 = 189.65$ N/m, $\delta_1 = 10.54$ kg/sec, $\delta_2 = 1.19$ kg/sec, $m_1 = 0.77$ kg, $m_2 = 0.60$ kg, and $k_m = 2.85$ N/V; these are the nominal parameter values for the mass-spring-damper system manufactured by Educational Control Products Inc. A state representation of system (59) is

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -491.74x_1 - 15.23x_2 + 10245.87x_3 + 1.55x_4 \\ &\quad + 3.69u + \gamma_1(\cdot), \end{aligned}$$

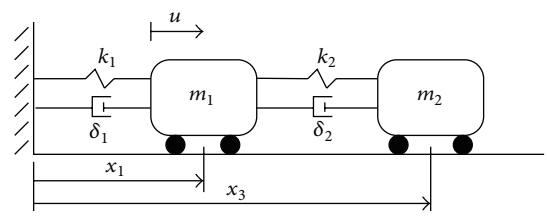


FIGURE 4: Underactuated mass-spring-damper mechanical systems.

$$\begin{aligned} \dot{x}_3 &= x_4, \\ \dot{x}_4 &= 312.72x_1 + 1.97x_2 - 312.72x_3 - 1.97x_4 + \gamma_2(\cdot), \\ y_1 &= x_1, \\ y_2 &= x_3. \end{aligned} \tag{60}$$

The control objective is

$$\lim_{t \to \infty} |x_3 - y_r(t)| = 0, \tag{61}$$

with a bounded behavior in x_1, x_2, and x_4.

Define the error $e_1 = x_3 - y_r(t)$ whose dynamics are given by

$$\begin{aligned} \dot{e}_1 &= e_2, \\ \dot{e}_2 &= -312.72e_1 - 1.97e_2 - 312.72y_r(t) - 1.97\dot{y}_r(t) \\ &\quad - \ddot{y}_r(t) + 1.97x_2 + \gamma_2(\cdot) + 312.72x_1, \\ y_e &= e_1, \\ \dot{x}_1 &= x_2, \end{aligned}$$

$$\begin{aligned}\dot{x}_2 &= -491.74x_1 - 15.23x_2 + 10245.87x_3 + 1.55x_4 \\ &\quad + 3.69\tau + \gamma_1(\cdot), \\ y_1 &= x_1.\end{aligned} \tag{62}$$

In this case the signal $u_e = x_{\text{re}}$ and is given by

$$\begin{aligned}x_{\text{re}} &= 3.19 \\ &\quad \times 10^{-3}\left(-c_1e_1 - c_2e_2 - 1.97x_2 - \gamma_2(\cdot) + \Theta\right),\end{aligned} \tag{63}$$

where $c_1 = c_2 = 10$ and $\Theta = 312.72y_r(t) + 1.97\dot{y}_r(t) + \ddot{y}_r(t)$.

To design the control u define the errors

$$\begin{aligned}\varepsilon_1 &= x_1 - x_{\text{re}}, \\ \varepsilon_2 &= x_2 - \dot{x}_{\text{re}},\end{aligned} \tag{64}$$

with dynamics given by

$$\begin{aligned}\dot{\varepsilon}_1 &= \varepsilon_2, \\ \dot{\varepsilon}_2 &= -491.74x_1 - 15.23x_2 + 10245.87x_3 + 1.55x_4 \\ &\quad + 3.69u + \gamma_1(\cdot) - \ddot{x}_{\text{re}},\end{aligned} \tag{65}$$

and the control input for this subsystem is

$$\begin{aligned}u &= 0.270\left(491.74x_1 + 15.23x_2 - 10245.87x_3\right. \\ &\quad \left. - 1.55x_4 - \gamma_1(\cdot) + \ddot{x}_{\text{re}} + u_o\right), \\ u_o &= -\alpha_1\varepsilon_1 - \alpha_2\varepsilon_2 - \alpha_3\,\text{sign}(\varepsilon_1).\end{aligned} \tag{66}$$

The observer used to estimate the unknown signals e_2 and $\gamma_2(\cdot)$ is

$$\begin{aligned}\dot{\hat{e}}_1 &= z_1 + c_{1,1}\left|e_1 - \hat{e}_1\right|^{1/2}\text{sign}\left(e_1 - \hat{e}_1\right), \\ \dot{z}_1 &= c_{2,1}\,\text{sign}\left(e_1 - \hat{e}_1\right), \\ \hat{y}_1 &= \hat{e}_1, \\ \dot{\hat{e}}_2 &= -312.72e_1 - 1.97e_2 - 312.72y_r(t) - 1.97\dot{y}_r(t) \\ &\quad - \ddot{y}_r(t) + 312.72x_1 + z_2 \\ &\quad + c_{1,2}\left|z_1 - \hat{e}_2\right|^{1/2}\text{sign}\left(z_1 - \hat{e}_2\right), \\ \dot{z}_2 &= c_{2,2}\,\text{sign}\left(z_1 - \hat{e}_2\right), \\ \hat{y}_2 &= \hat{e}_2,\end{aligned} \tag{67}$$

where $z_1 \approx e_2$ and $z_2 \approx \gamma_2(\cdot)$.

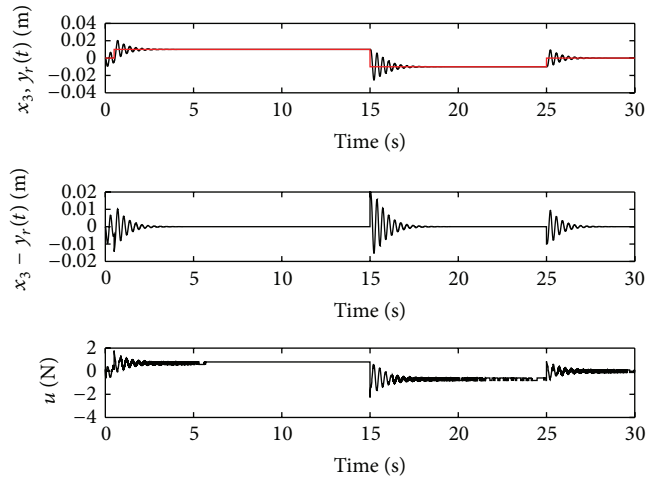


Figure 5: Simulation results with the mass-spring-damper system; performance for the regulation problem.

The observer to estimate ε_2, $\gamma_1(\cdot)$, and $\ddot{x}_{\text{re}}$ is

$$\begin{aligned}\dot{\hat{\varepsilon}}_1 &= w_1 + a_{1,1}\left|\varepsilon_1 - \hat{\varepsilon}_1\right|^{1/2}\text{sign}\left(\varepsilon_1 - \hat{\varepsilon}_1\right), \\ \dot{w}_1 &= a_{2,1}\,\text{sign}\left(\varepsilon_1 - \hat{\varepsilon}_1\right), \\ \hat{y}_1 &= \hat{\varepsilon}_1, \\ \dot{\hat{\varepsilon}}_2 &= -491.74x_1 + 10245.87x_3 + 3.69u + w_2 \\ &\quad + a_{1,2}\left|w_1 - \hat{\varepsilon}_2\right|^{1/2}\text{sign}\left(w_1 - \hat{\varepsilon}_2\right), \\ \dot{w}_2 &= a_{2,2}\,\text{sign}\left(w_1 - \hat{\varepsilon}_2\right), \\ \hat{y}_2 &= \hat{\varepsilon}_2,\end{aligned} \tag{68}$$

where $w_1 \approx \varepsilon_2$ and $w_2 \approx -491.74x_1 - 15.23x_2 + 10245.87x_3 + 1.55x_4 + \gamma_1(\cdot) - \ddot{x}_{\text{re}}$.

The control inputs (13) and (16) may be implemented in the following form:

$$\begin{aligned}x_{\text{re}} &\approx \frac{1}{g_2}\left(-z_2 + \ddot{y}_r(t) - k_1e_1 - k_2\hat{e}_2\right), \\ u &\approx \frac{1}{g_1(x_1, x_3)}\left(-w_2 - \alpha_1\varepsilon_1 - \alpha_2\hat{\varepsilon}_2 - \alpha_3\,\text{sign}(\varepsilon_1)\right),\end{aligned} \tag{69}$$

where $k_1 = 10$, $k_2 = 1$, $\alpha_1 = 20$, $\alpha_2 = 20$, and $\alpha_3 = 0.4$.

5.2.1. Numerical Results. Figure 5 shows the results when the reference y_r, red line, takes different constant values; this is the case of regulation. As can be seen, the output signal x_3 (black line) reaches asymptotically the reference after a short transient. The steady state error $x_3 - y_r(t)$ is practically zero and the control signal u takes values suitable for a possible implementation.

The same situation occurs when the reference is a time-varying signal. Figure 6 shows that the output signal x_3 (black line) of the unactuated link converges to a time-varying signal

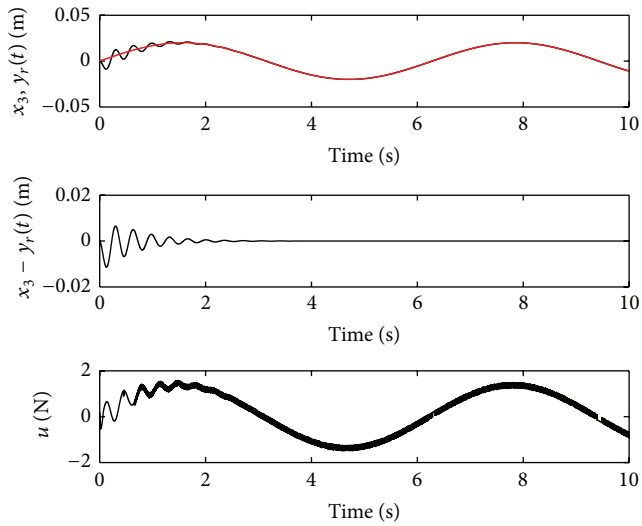


Figure 6: Simulation results with the mass-spring-damper system; performance for the tracking problem.

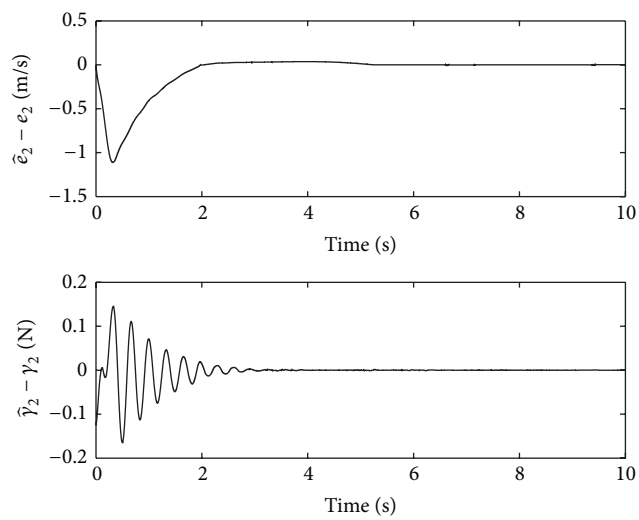


Figure 7: Estimation errors of e_2 and $\gamma_2(\cdot)$.

$y_r(t)$ (red line) with an almost zero steady-state error, $x_3 - y_r(t)$, and a control signal u with good characteristics for experimental implementation.

For this last case, we analyze the behavior of the state observers. Figure 7 shows the behavior of the errors $\widehat{e}_2 - e_2$ and $\widehat{\gamma}_2(\cdot) - \gamma_2(\cdot)$; these errors converge to zero in few seconds.

For the observer that estimates ε_2 and the term that includes $\gamma_1(\cdot)$ and $\ddot{x}_{re}$, it is not possible to compare the actual values with the estimate values. Therefore we check the behavior of the errors $\varepsilon_1 - \widehat{\varepsilon}_1$ and $\varepsilon_2 - \widehat{\varepsilon}_2$; as these errors go to zero, as we can see in Figure 8, the estimation of ε_2 and the term that includes $\gamma_1(\cdot)$ and $\ddot{x}_{re}$ is correct.

5.2.2. Experimental Results. The proposed controller is applied to a spring mass damping system manufactured by Educational Control Products Inc., shown in Figure 9. In this

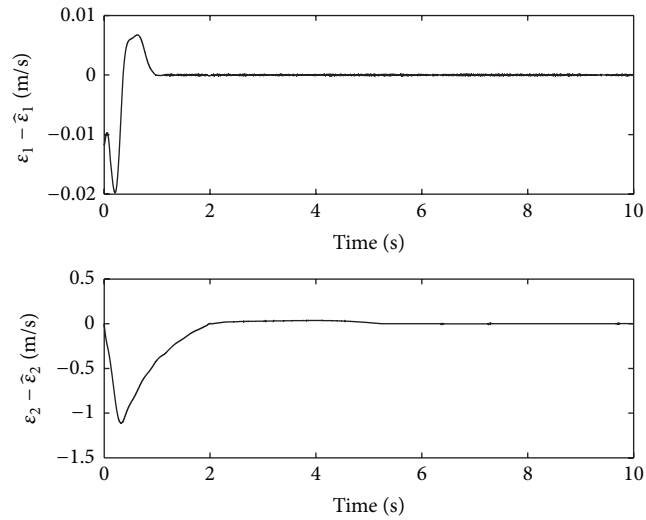


Figure 8: Behavior of the internal errors of the observer that estimates ε_2, $\gamma_1(\cdot)$, and $\ddot{x}_{re}$.

Figure 9: Mass damper spring system used in the experiments.

experiment it is assumed that the plant has the same parameters as those considered in the numerical simulation of the previous section and did not conduct a rigorous procedure to estimate the parameters of the real plant, creating a significant challenge to the controller. This situation was resolved by tuning the parameters of each observer, the internal control signal u_e, and the total control u; where $k_1 = 20$, $k_2 = 2$, $\alpha_1 = 20$, $\alpha_2 = 5$, and $\alpha_3 = 0.4$.

Experimental results are shown with a reference signal y_r with constant values at different times; that is, the control objective is regulation. The results are shown in Figure 10, where we can see that the transient takes about one second; the amplitude of the steady-state error, $x_3 - y_r$, has a maximum of 8×10^{-6} meters and the control signal u takes values which are in the permissible range of the control system, ±3 volts.

In the second experiment we apply a time varying signal; a sine function, that is, the control target, is tracking. The results are shown in Figure 11, where we can see that the transient takes about 5.8 seconds with an initial error of about 0.01 meters, the steady-state error amplitude, $x_3 - y_r(t)$, has a maximum of $\pm 3 \times 10^{-4}$ meters, and the control signal takes values which are in the permissible range control system, as in the previous case, between ±3 volts.

6. Conclusions

The control strategy proposed formally guarantees the control objective, either regulation or trajectory tracking, and at

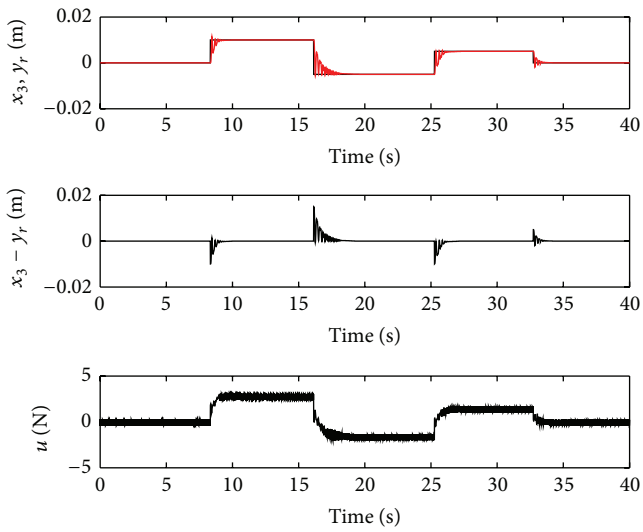


Figure 10: Experimental results and the regulation case.

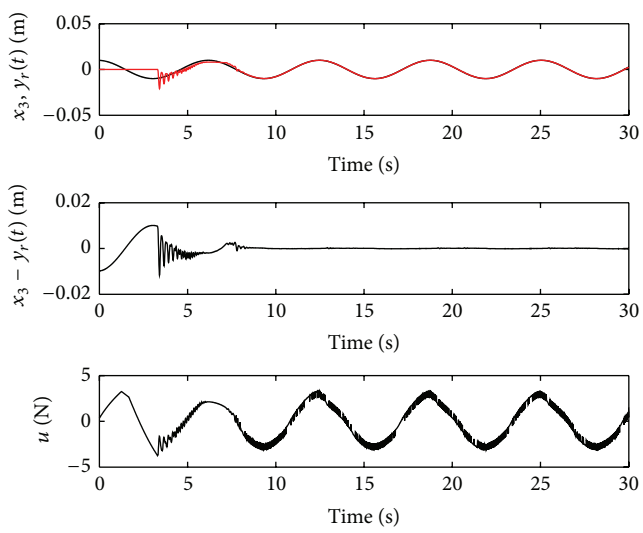


Figure 11: Experimental results and tracking case.

the same time establishes a strategy for its implementation considering partial measurement of the state variables and parametric uncertainties. Although stability is not global, the subspace that can ensure stability can be made as large as needed in practice. Some of its limitations are the number of parameters to adjust, both in the observers and in the controller, and the need to use a real-time platform to implement the controller to ensure a sample time less than or equal to one millisecond, and thus the actual sliding mode, produced by discontinuous terms, enough approaches the ideal sliding mode.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] N. Adhikary and C. Mahanta, "Integral backstepping sliding mode control for underactuated systems: swing-up and stabilization of the Cart-Pendulum System," *ISA Transactions*, vol. 52, no. 6, pp. 870–880, 2013.

[2] S. Andary, A. Chemori, M. Benoit, and J. Sallantin, "A dual model-free control of underactuated mechanical systems, application to the inertia wheel inverted pendulum," in *Proceedings of the American Control Conference (ACC '12)*, pp. 1029–1034, June 2012.

[3] M. Bettayeb, C. Boussalem, R. Mansouri, and U. M. Al-Saggaf, "Stabilization of an inverted pendulum-cart system by fractional PI-state feedback," *ISA Transactions*, vol. 53, no. 2, pp. 508–516, 2014.

[4] L. Xu, Q. Hu, and G. Ma, "Output feedback stabilization control for underactuated mechanical systems," in *Proceedings of the 31st Chinese Control Conference (CCC '12)*, pp. 4267–4272, IEEE, July 2012.

[5] F. Mnif, "VSS control for a class of underactuated mechanical systems," *International Journal of computational Cognition*, vol. 3, no. 2, 2005.

[6] P. Morin and C. Samson, "Control of underactuated mechanical systems by the transverse function approach," in *Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference (CDC-ECC '05)*, pp. 7508–7513, IEEE, December 2005.

[7] M. S. Park, D. Chwa, and S. K. Hong, "Decoupling control of a class of underactuated mechanical systems based on sliding mode control," in *Proceedings of the International Joint Conference (SICE-ICASE '06)*, pp. 806–810, IEEE, 2006.

[8] M. Reyhanoglu, A. van der Schaft, N. H. McClamroch, and I. Kolmanovsky, "Dynamics and control of a class of underactuated mechanical systems," *IEEE Transactions on Automatic Control*, vol. 44, no. 9, pp. 1663–1671, 1999.

[9] V. Sankaranarayanan and A. D. Mahindrakar, "Control of a class of underactuated mechanical systems using sliding modes," *IEEE Transactions on Robotics*, vol. 25, no. 2, pp. 459–467, 2009.

[10] W. Wang, J. Yi, D. Zhao, and X. Liu, "Double layer sliding mode control for second-order underactuated mechanical systems," in *Proceedings of the IEEE IRS/RSJ International Conference on Intelligent Robots and Systems (IROS '05)*, pp. 295–300, IEEE, August 2005.

[11] Y.-F. Chen and A.-C. Huang, "Controller design for a class of underactuated mechanical systems," *IET Control Theory & Applications*, vol. 6, no. 1, pp. 103–110, 2012.

[12] S. Rudra, R. K. Barai, M. Maitra et al., "Global stabilization of a flat underactuated inertia wheel: a block backstepping approach," in *Proceedings of the 3rd International Conference on Computer Communication and Informatics (ICCCI '13)*, pp. 1–4, IEEE, January 2013.

[13] J.-X. Xu, Z.-Q. Guo, and T. H. Lee, "Sliding mode controller design for underactuated systems," in *Proceedings of the 12th International Workshop on Variable Structure Systems (VSS '12)*, pp. 385–390, January 2012.

[14] R. M. T. Raja Ismail, D. T. Nguyen, and Q. P. Ha, "Observer-based trajectory tracking for a class of underactuated Lagrangian systems using higher-order sliding modes," in *Proceedings of the IEEE International Conference on Automation Science and Engineering: Green Automation Toward a Sustainable Society (CASE '12)*, pp. 1204–1209, August 2012.

[15] C.-C. Cheng, K.-S. Yang, and J.-H. Yang, "Robust finite time controller design for second order nonlinear underactuated mechanical systems," *Transactions of the Canadian Society for Mechanical Engineering*, vol. 37, no. 3, pp. 549–557, 2013.

[16] P. K. Khosla and T. Kanade, "Parameter identification of robot dynamics," in *Proceedings of the 24th IEEE Conference on Decision and Control*, pp. 1754–1760, IEEE, Fort Lauderdale, Fla, USA, December 1985.

[17] D. I. R. Almeida, J. Alvarez, and L. Fridman, "Robust observation and identification of nDOF Lagrangian systems," *International Journal of Robust and Nonlinear Control*, vol. 17, no. 9, pp. 842–861, 2007.

[18] H. K. Khalil, *Nonlinear Systems*, vol. 3, Prentice Hall, Upper Saddle River, NJ, USA, 2002.

[19] A. Levant, "Robust exact differentiation via sliding mode technique," *Automatica*, vol. 34, no. 3, pp. 379–384, 1998.

[20] J. Davila, L. Fridman, and A. Levant, "Second-order sliding-mode observer for mechanical systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 11, pp. 1785–1789, 2005.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 365206, 9 pages
http://dx.doi.org/10.1155/2015/365206

Research Article

Three-Stage Tracking Control for the LED Wafer Transporting Robot

Zuoxun Wang[1] and Zhiguo Yan[1,2]

[1] *School of Electrical Engineering and Automation, Qilu University of Technology, Jinan 250353, China*
[2] *Key Laboratory of Pulp and Paper Science & Technology of Ministry of Education of China, Qilu University of Technology, Jinan 250353, China*

Correspondence should be addressed to Zuoxun Wang; wangzuoxun@126.com

Received 16 May 2015; Accepted 17 June 2015

Academic Editor: Xinguang Zhang

In order to ensure the steady ability of the LED wafer transporting robot, a high order polynomial interpolation method is proposed to plan the motion process of the LED wafer transporting robot. According to the LED wafer transporting robot which is fast and has no vibration, fifth-order polynomial is applied to complete the robot's motion planning. A new subsection search method is proposed to optimize the transporting robot's acceleration. Optimal planning curve is achieved by the subsection searching. Extended Kalman filter algorithm and PID algorithm are employed to follow the tracks of planned path. MATLAB simulation and experiment confirm the validity and efficiency of the proposed method.

1. Introduction

The robot motion planning is one of the most important problems for the robot control. Motion planning includes path planning and motion control. Path planning is to search an optimal trajectory of the path from the beginning point to the finishing point in the robot motion space. Reference [1] uses an ant colony optimization algorithm to realize path planning. Reference [2] proposes coordinated trajectory planning methods of two typical applications, which can assure the applications' stabilization. Reference [3] proposes autonomous motion control approaches to control dual-arm space robot for target capturing. Reference [4] proposes quantum-behaved particle swarm optimization (QPSO) algorithm to plan the robot path, and author uses the probability theory to study the relationship with the parameters and convergence of mobile robot path planning, and at last author proposed an improved trajectory planning method. Reference [5] points out that the basic problem of the path planning is the common model expression and path optimization strategy. Common model expression methods include visibility graph, free space method, and grid method. Optimal path search problem is then converted to search the shortest route from a beginning point to the target point via the visible lines. By far, most robots employ trapezoidal speed curves for their motion planning.

Some problems were discovered by research. The classical planning methods have some distinguishing feature. First, when the acceleration or the velocity is fixed, the acceleration and the velocity must be set to a low value to ensure that they do not exceed the constraints during the whole process which is impossible to optimize acceleration and velocity at some points in whole moving process. Second, sudden change of acceleration can cause system oscillation. System oscillation can be reduced by planning the motion acceleration. When the robot moves according to a control method based on its motion model, sudden change of the acceleration will certainly cause the system oscillation. The principal purpose of motion planning is to discover a reasonable polynomial function or other functions to conduct the interpolation, so that the robot motion can be smooth and stable with vibration within the acceptable range. And the consumed time from the beginning point of the motion to the finishing point is as few as possible.

In order to solve the problems on how to realize large range, high speed, and high accuracy trajectory tracking, robot dynamic real-time control method is raised which has two problems. The first is how to keep the system stable. Advanced control method should be studied and applied to robot real-time control, so that the tracking error can diminish to possible range as quickly as possible. The second is how to diminish disturbance and how to reduce the influence of the disturbance to the accuracy of the tracking. If the precise motion model of the robot can be achieved and the disturbance signal can be eliminated, then controller designed with linear control theory can solve the two problems. But the precise and complete motion model of the robot almost cannot be constructed, because of the error in measurement and modeling, the changing of the load, and disturbance from the environment [5–7]. So when the robot motion model is constructed, some reasonable approximations had to be made and some unimportant uncertainties might be ignored. In the field of industrial control, the PID control algorithm is one of the most important control algorithms. It plays a vital role in the industrial production process. As far as control field is concerned, PID control method has the dominated position in the field for many years. Although much progress has been made in model based mathematical control science, PID control method has the significant impact on industry. In recent years, some scholars begin to attempt some new control method to replace the PID method [8–10]. After many years' unremitting efforts, using the Kalman filter tracking control is a trend in the aviation field. But Kalman filter's application is less in any other areas [11–14].

In the process of transporting the LED wafer, we need to transport it as fast as possible. Moreover, the vibration of the fast transporting process of the robot must be as small as possible. In this paper, two fifth-order polynomial interpolation functions and a first-order polynomial interpolation function are used to plan the motion trajectory of the fast transporting robot. In addition, a new method is proposed to minimize the maximum acceleration value in the acceleration stage and deceleration stage to reduce the oscillation. PID algorithm and extended Kalman filter algorithm are proposed to track control the movement of the fast transporting robot.

2. Motion Planning Theory

The fundamental task of the robot motion trajectory planning is to select reasonable polynomial function or other linear functions to accomplish interpolation operation task [15–18]. It can make robot movement smooth, steady ability, and keep robot movement error within certain range. In the process of the robot movement, the robot position y_0 at the beginning point is known, and the robot position y_e at the finishing point can be achieved by using the inverse kinematics. Thus, the description of the motion trajectory can be represented by a smooth interpolation function, which can describe robot position $y(x)$ from the beginning point to the finishing point. At the time x_0, y_0 is the beginning point of robot position $y(x)$. At the time x_e, y_e is the finishing point of robot position $y(x)$. In order to realize the smooth and steady movement of transporting robot, trajectory $y(x)$ at least needs to meet four limit conditions:

$$\begin{aligned} y(0) &= y_0, \\ y(x_e) &= y_e, \\ y'(0) &= 0, \\ y'(x_e) &= 0. \end{aligned} \tag{1}$$

The above four conditions can define a unique third-order polynomial equation:

$$y(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3. \tag{2}$$

Equation (2) is position of the robot. The first derivative of (2) is the speed of the robot:

$$y'(t) = c_1 + 2c_2 x + 3c_3 x^2. \tag{3}$$

Equation (1) was substituted into (2) and (3); the following equations can be obtained:

$$\begin{aligned} y(0) &= c_0 = y_0 \\ y(x_e) &= c_0 + c_1 x_e + c_2 x_e^2 + c_3 x_e^3 = y_e \\ y'(0) &= c_1 = 0 \\ y'(x_e) &= c_1 + 2c_2 x_e + 3c_3 x_e^2 = 0. \end{aligned} \tag{4}$$

Equation (4) can be written into matrix form:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & x_e & x_e^2 & x_e^3 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 2x_e & 3x_e^2 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} y_0 \\ y_e \\ 0 \\ 0 \end{pmatrix}. \tag{5}$$

The following result can be achieved by calculating (5):

$$\begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} y_0 \\ 0 \\ \dfrac{3}{x_e^2}(y_e - y_0) \\ -\dfrac{2}{x_e^3}(y_e - y_0) \end{pmatrix}. \tag{6}$$

A unique third-order polynomial equation can be determined by (6). Therefore, if the beginning position, beginning speed, finishing position, and finishing speed are known, using third-order polynomial interpolation method can acquire a complete motion trajectory equation. When the robot system's acceleration has to be limited, a fifth-order polynomial interpolation method will be needed to plan the motion trajectory of the robot movement process. Equation of the robot's position is expressed in the following form:

$$y(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + c_4 x^4 + c_5 x^5. \tag{7}$$

First-order derivative of (7) is the robot's speed of the motion:

$$y'(x) = c_1 + 2c_2x + 3c_3x^2 + 4c_4x^3 + 5c_5x^4. \quad (8)$$

Second-order derivative of (7) is the robot's acceleration of the motion:

$$y''(x) = 2c_2 + 6c_3x + 12c_4x^2 + 20c_5x^3. \quad (9)$$

The moving process of the fast transporting robot requires small vibration. According to those characteristics, the whole motion trajectory is divided into three stages: accelerating stage, uniform stage, and decelerating stage. The acceleration and time are all important in the accelerating stage and decelerating stage, so fifth-order polynomial interpolation is employed to plan accelerating stage and decelerating stage. One order polynomial interpolation is used to plan uniform stage. The accelerating stage uses the following equation to plan:

$$y(x) = c_{00} + c_{01}x + c_{02}x^2 + c_{03}x^3 + c_{04}x^4 + c_{05}x^5. \quad (10)$$

The uniform stage uses the following equation to plan:

$$y(x) = c_{10} + c_{11}x. \quad (11)$$

The decelerating stage uses the following equation to plan:

$$y(x) = c_{20} + c_{21}x + c_{22}x^2 + c_{23}x^3 + c_{24}x^4 + c_{25}x^5. \quad (12)$$

Four moments are critical for the whole trajectory of motion planning, which are the beginning time of the accelerating stage x_0, the finishing time of the accelerating stage x_1, the beginning time of the decelerating stage x_2, and the finishing time of the decelerating stage x_3, and the corresponding positions of the robot are y_0, y_1, y_2, and y_3. In order to optimize the maximum acceleration value both in the accelerating stage and in the decelerating stage, x_1, y_1, x_2, and y_2 must be optimized. Hence, subsection searching method is used to find the optimum x_1, y_1, x_2, and y_2.

Remark 1. For the motion equation, x is variable of equation and c is coefficient of equation. The beginning point and end point of x are selected to acquire the value of coefficient c.

3. Acceleration Optimization

The movement trajectory can be planned by interpolation through some point in the path. The whole trajectory can be divided into a number of segments by those special points. If robot stayed at some point of the path for a while, in which beginning velocity and the finishing velocity are zero, therefore polynomial interpolation method can be directly used. If it does not stop at a point in the process of movement, inverse kinematics solution can be used to determine polynomial interpolation function and connect every point of path smoothly. Fifth-order polynomial interpolation method is used to plan the trajectory of the movement process of the fast transporting robot during the accelerating stage. Its beginning position $y(0)$, beginning velocity $y'(0)$, and beginning acceleration $y''(0)$ are as follows:

$$\begin{aligned} y(0) &= y_0 = 0, \\ y'(0) &= y'_0 = 0, \\ y''(0) &= y''_0 = 0. \end{aligned} \quad (13)$$

At the end of the acceleration stage x_1, the position $y(x_1)$, velocity $y'(x_1)$, and acceleration $y''(x_1)$ are as follows:

$$\begin{aligned} y(x_1) &= y_1, \\ y'(x_1) &= y'_1 = a_{11}, \\ y''(x_1) &= y''_1 = 0. \end{aligned} \quad (14)$$

Equations (13) and (14) were substituted into (10); the following equation can be obtained:

$$\begin{aligned} y(0) &= c_{00} = 0 \\ y'(0) &= c_{01} = 0 \\ y''(0) &= 2c_{02} = 0 \\ y(x_1) &= c_{00} + c_{01}x_1 + c_{02}x_1^2 + c_{03}x_1^3 + c_{04}x_1^4 + c_{05}x_1^5 \\ &= y_1 \\ y'(x_1) &= c_{01} + 2c_{02}x_1 + 3c_{03}x_1^2 + 4c_{04}x_1^3 + 5c_{05}x_1^4 \\ &= c_{11} \\ y''(x_1) &= 2c_{02} + 6c_{03}x_1 + 12c_{04}x_1^2 + 20c_{05}x_1^3 = 0. \end{aligned} \quad (15)$$

Remark 2. In (15), x_1 is known; c_{02}, c_{03}, c_{04}, and c_{05} are unknown. c_{02}, c_{03}, c_{04}, and c_{05} can be obtained by x_1 and y_1.

Equation (15) can be rewritten as

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 1 & x_1 & x_1^2 & x_1^3 & x_1^4 & x_1^5 \\ 0 & 1 & 2x_1 & 3x_1^2 & 4x_1^3 & 5x_1^4 \\ 0 & 0 & 2 & 6x_1 & 12x_1^2 & 20x_1^3 \end{pmatrix} \begin{pmatrix} c_{00} \\ c_{01} \\ c_{02} \\ c_{03} \\ c_{04} \\ c_{05} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ y_1 \\ c_{11} \\ 0 \end{pmatrix}. \quad (16)$$

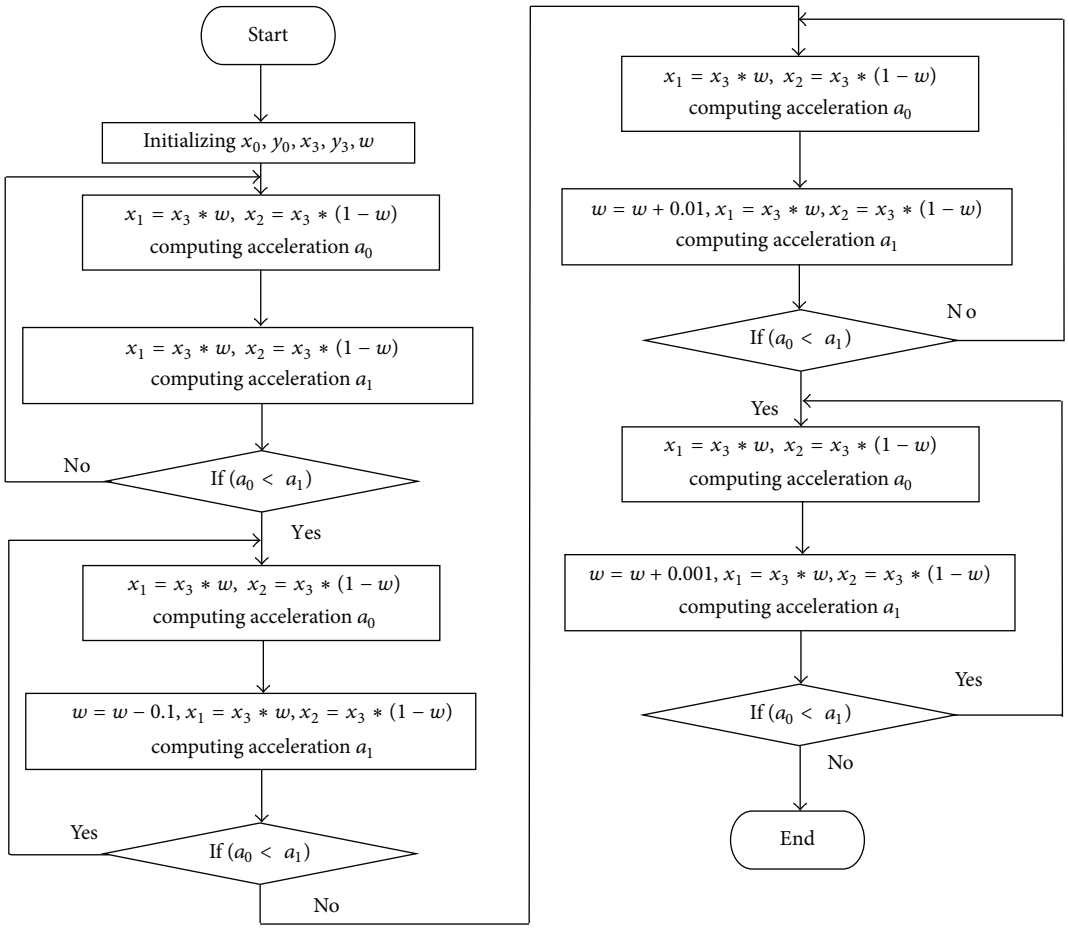


Figure 1: Subsection searching optimizing acceleration.

In (16) c_{11} is the robot velocity of the uniform stage, which can be achieved by substituting x_1, y_1, x_2, and y_2 into (11):

$$\begin{aligned} y(x_1) &= c_{10} + c_{11}x_1 = y_1 \\ y(x_2) &= c_{10} + c_{11}x_2 = y_2. \end{aligned} \tag{17}$$

The acceleration equation of the accelerating stage is

$$y''(x) = 2c_{02} + 6c_{03}x + 12c_{04}x^2 + 20c_{05}x^3. \tag{18}$$

Subsection searching method is used to find the optimal c_{02}, c_{03}, c_{04}, and c_{05}, which can minimize the maximum acceleration value of the nonlinear high order polynomial of the accelerating stage. Subsection searching method is proposed based on branch and bound method of optimization theory. Branch and bound method needs two stages of operations: The first is branch, which divides the solutions into several nonintersect solution sets, according to certain rules. The second is bound, which selects an appropriate algorithm to compute the bound of the subsection which will be conducted again and again; thus the solution set will become smaller and smaller, and at last, an accurate solution will be achieved.

Figure 1 is the process of searching x_1 using subsection searching method. The proposed subsection searching method in this paper searching process includes the following stages. The first is to choose a random point in the solution set as the starting searching point. The second is to begin subsection. In the figure, a_0 is last computing maximum acceleration value based on given subset and target function and a_1 is this time computing maximum acceleration value. Subsection direction depends on the results of comparison a_0 and a_1 value. Search speed depends on w. In order to improve the computing speed, w can be chosen to increase or decrease. The step length can choose 0.1, 0.01, or 0.001. After repeated computing and comparison many times, optimal x_1 can be obtained. For obtaining optimal y_1, the same method can be applied and then optimal x_2 and y_2 can be obtained. Substituting x_1, y_1, x_2, and y_2 into (16) and (17), c_{00}, c_{01}, c_{02}, c_{03}, c_{04}, and c_{05} can be obtained.

Remark 3. By searching the proportion of the time of accelerating stage, decelerating stage and uniform stage motion curve can be optimized.

4. System Modeling and Control

4.1. Modeling. Servomotor is used to control the motion of the fast wafer transporting robot. Servomotor's job is to transfer the input electric power into the robot system's

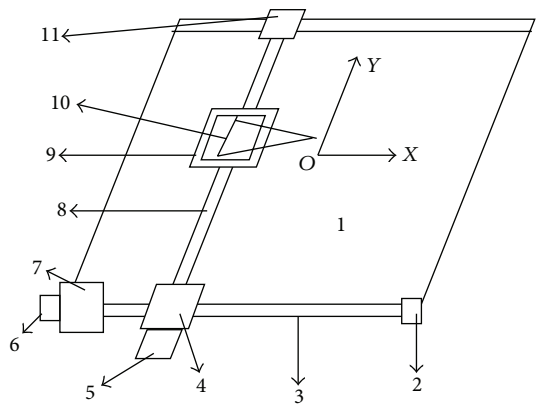


Figure 2: The mechanical structure of the transporting robot.

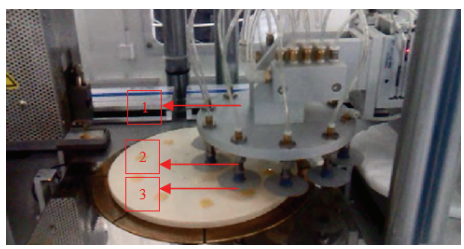


Figure 3: The picture of the transporting robot.

mechanical energy [19, 20]. The mechanical structure of the transporting robot is shown in Figure 2.

In Figure 2, 1 is installer. 2 is bearing A. 3 is lead screw A. 4 is servomotor B. 5 is encoder B. 6 is encoder A. 7 is servomotor A. 8 is lead screw B. 9 is workbench. 10 is transporting arm. 11 is bearing B.

In Figure 3, 1 is transporting arm. 2 is vacuum pad. 3 is wafer. Vacuum pads grab the wafer and transport them from source position to target position. In the transporting process, movement is smooth and has no vibration.

In servomotor, the rotor voltage $e(x)$ and the rotor current $I(x)$ are induced in the rotor circuit. Then the rotor current and stator magnetic flux interact to produce electromagnetic torque $F(x)$. Its equation is as follows:

$$e(x) = L\frac{dI(x)}{dx} + RI(x) + E, \tag{19}$$

where E is counter electromotive force (EMF), $E = A_e\omega(x)$, A_e is EMF constant, L is the inductance parameter, and R is the rotor circuit resistance value.

Electromagnetic torque equation is

$$F(x) = A_m I(x), \tag{20}$$

where A_m is servomotor torque coefficient and $F(x)$ is the electromagnetic torque produced by servomotor.

Servomotor torque balance equation is

$$B_m\frac{d\omega(x)}{dx} + f_m\omega(x) = F(x) - F_c(x), \tag{21}$$

where f_m is motor shaft sticky friction coefficient, B_m is motor shaft rotary inertia, and $F_c(x)$ is total load torque. Remove the middle variable of (19), (20), and (21); the following motor differential equation (22) can be obtained

$$\begin{aligned} LB_m\frac{d^2\omega(x)}{dx^2} &+ (Lf_m + RB_m)\frac{d\omega(x)}{dx} \\ &+ (Rf_m + A_mA_e)\omega(x) \\ &= A_me(x) - L\frac{dF_c(x)}{dx} - RF_c(x). \end{aligned} \tag{22}$$

In (22), the inductance L is very small, which can be ignored, so (22) is simplified as follows:

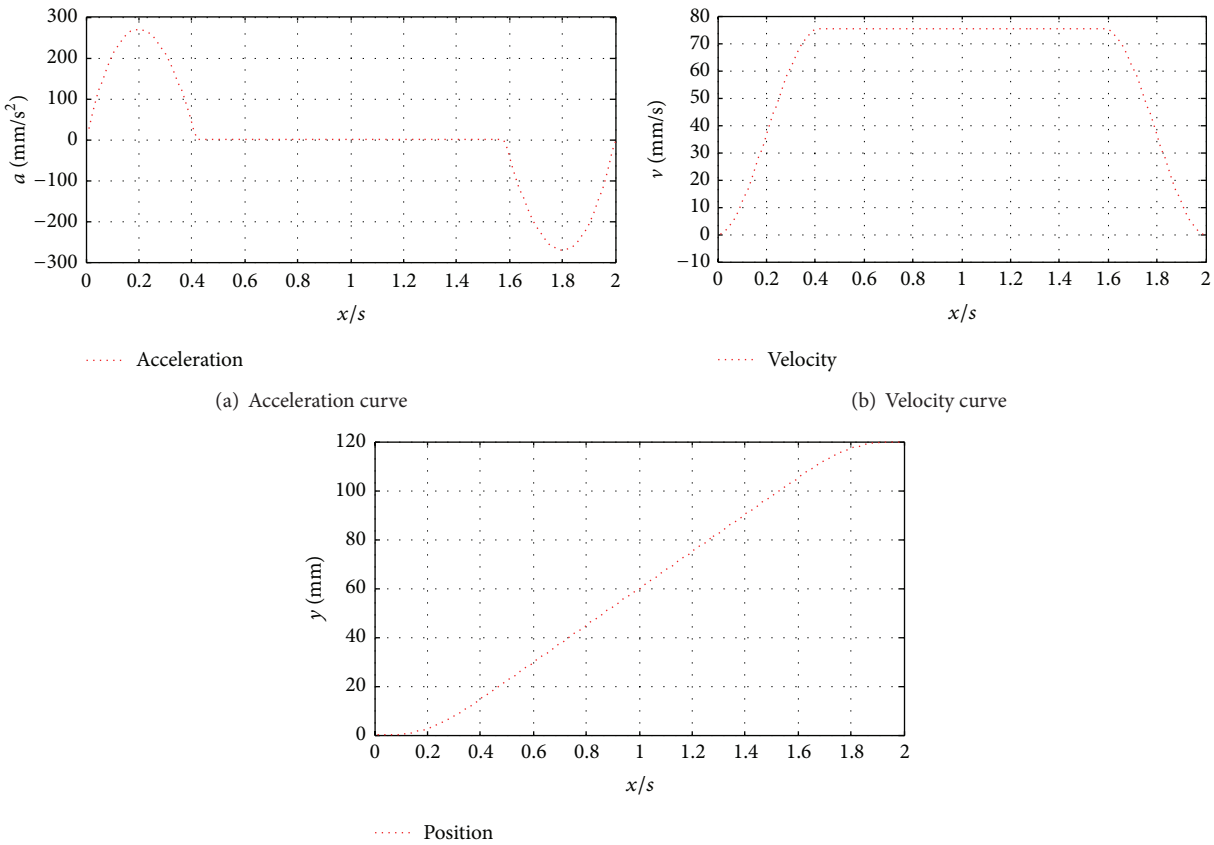
$$H_m\frac{d\omega(x)}{dx} + \omega(x) = N_1e(x) - N_2F_c(x), \tag{23}$$

where

$$\begin{aligned} H_m &= \frac{RB_m}{Rf_m + A_mA_e}, \\ N_1 &= \frac{A_m}{Rf_m + A_mA_e}, \\ N_2 &= \frac{R}{Rf_m + A_mA_e}. \end{aligned} \tag{24}$$

If $F_c(x) = 0$, (23) becomes

$$H_m\frac{d\omega(x)}{dx} + \omega(x) = N_1e(x). \tag{25}$$



(a) Acceleration curve

(b) Velocity curve

(c) Position curve

Figure 4: The curve of before optimization ($w = 0.21$).

Applying Laplace transform for (25), the following equation can be obtained:

$$G(s) = \frac{\Omega(s)}{E(s)} = \frac{N_1}{H_m s + 1}. \tag{26}$$

Furthermore, transfer function can be obtained from voltage $u(x)$ and angular displacement p:

$$G(s) = \frac{P(s)}{E(s)} = \frac{N_1}{s(H_m s + 1)}. \tag{27}$$

4.2. Kalman Filter Tracking Control and PID Control. The extended Kalman filter is used to track the position and velocity of system. Transporting robot moved according to the given direction and speed. Sensors are used to measure the distance and azimuth. Considering the noise of the system motion process, at time n, the system speed component is as follows:

$$v[n] = v[n-1] + u[n]. \tag{28}$$

In the equation, $u[n]$ is noise disturbance. According to the motion equations, the following location equations in N time are as follows:

$$r[n] = r[n-1] + v[n-1]\Delta. \tag{29}$$

In the equation, Δ is the interval between samples. In the discrete model of the equations of motion, the system will move according to the speed of a moment ago and then suddenly change at the next moment. Now, the signal vector by the choice is made of the position and velocity components. The equation is as follows:

$$s[n] = \begin{bmatrix} r[n] \\ v[n] \end{bmatrix}. \tag{30}$$

Equations (28), (29), and (30) are replaced by

$$\begin{bmatrix} r[n] \\ v[n] \end{bmatrix} = \begin{bmatrix} 1 & \Delta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} r[n-1] \\ v[n-1] \end{bmatrix} + \begin{bmatrix} 0 \\ u[n] \end{bmatrix}. \tag{31}$$

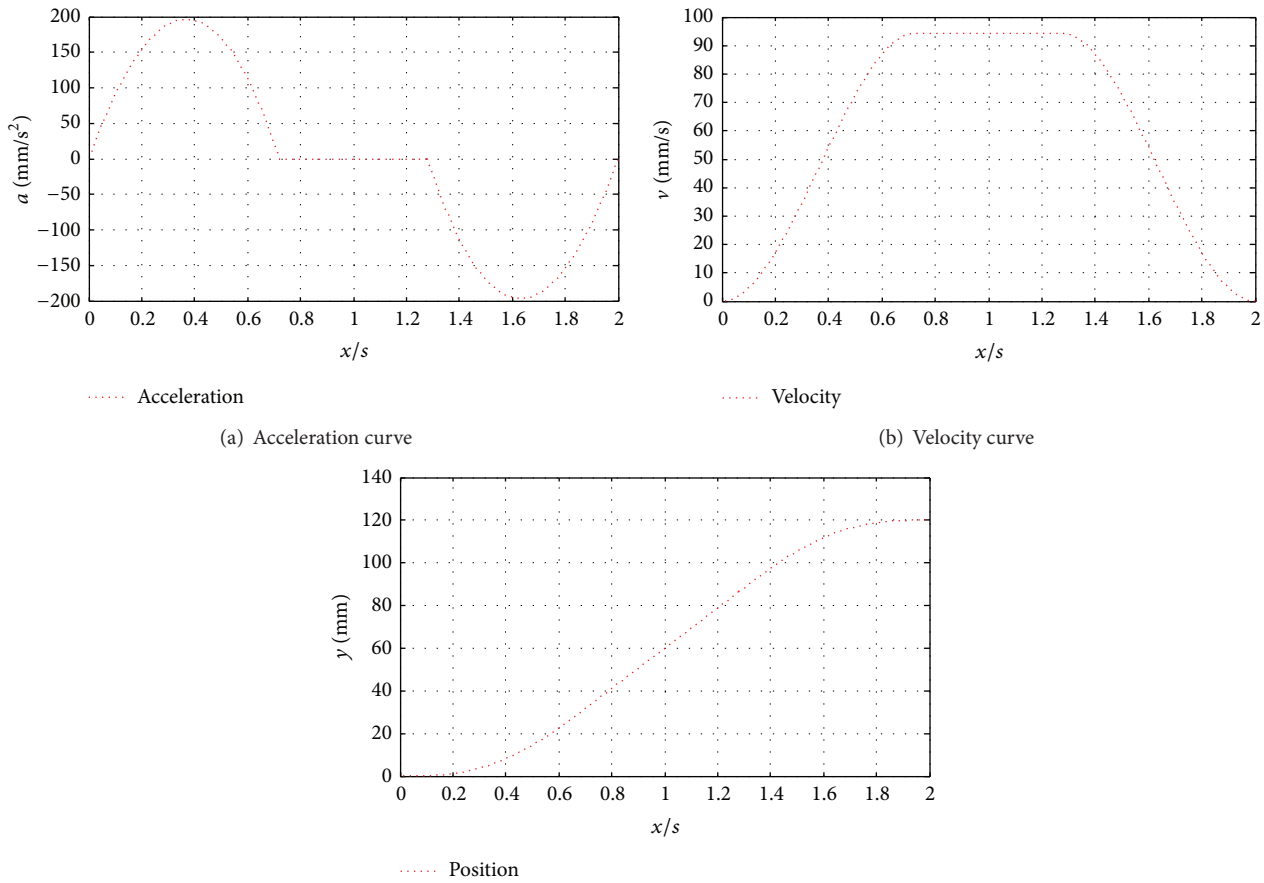
The observation equation of the system is

$$x[n] = H(s[n]) + w[n]. \tag{32}$$

In the equation, H is a function. In order to estimate the signal vector, the extended Kalman filter is applied; we now need to determine H:

$$H[n] = \left.\frac{\partial H}{\partial s[n]}\right|_{s[n]=\hat{s}[nn-1]}. \tag{33}$$

The Jacobian matrix can be obtained through seeking for the derivative of the observation equation.



(a) Acceleration curve

(b) Velocity curve

(c) Position curve

Figure 5: The curve after optimization ($w = 0.305$).

Remark 4. Through Jacobian matrix, Kalman filter can estimate H; therefore, it can track the motion trajectory, and it can obtain the better tracking effect.

PID controller is as follows:

$$u(x) = K_P\left[E(x) + \frac{1}{T_I}\int_0^x E(z)\,dz + T_D\frac{dE(x)}{dx}\right], \quad (34)$$

where $E(x) = R(x) - O(x)$, $R(t)$ is the input, and $O(x)$ is the output. K_P, T_I, and T_D are PID parameters. $u(x)$ is control variable. Furthermore, (34) can be written as follows:

$$u(x) = K_P E(x) + K_I\int_0^x E(z)\,dz + K_D\frac{dE(x)}{dt}, \quad (35)$$

where $K_I = K_P/T_I$ and $K_D = K_P T_D$.

The discrete equation is as follows:

$$\begin{aligned} u(n) = {} & K_P[E(n) - E(n-1)] + K_I E(n) \\ & + K_D[E(n) - 2E(n-1) + E(n-2)]. \end{aligned} \quad (36)$$

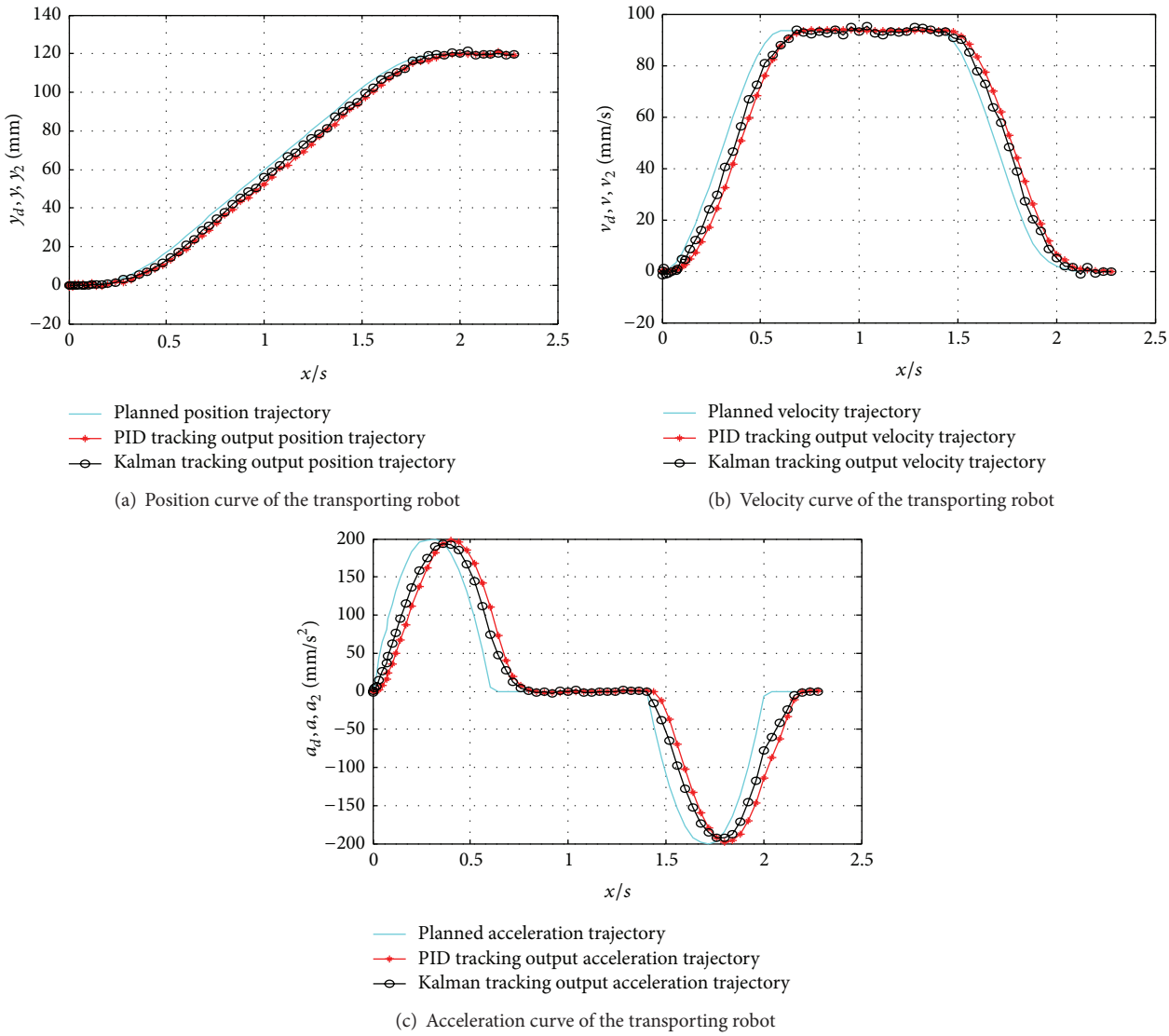
5. Simulation and Experiment

The fast transporting process of the LED wafer transporting robot was simulated by applying the proposed method. Firstly, the motion trajectory is divided into three segments, which are acceleration stage, uniform stage, and deceleration stage. The parameters are set as follows: $x_3 = 2$ s, $y_3 = 120$ mm, and $w = 0.21$. Then, we begin to search the extreme value point using the multisegments searching method. After this point is found, two fifth-order polynomials and a first-order polynomial are used to plan system's trajectory. At last, we obtain the motion trajectory with the minimum acceleration when $w = 0.305$.

Figures 4 and 5 show the simulation results of the motion planning using MATLAB. As shown in these figures, all the curves of position, velocity, and acceleration have no singular point. The changes of the acceleration can be seen from Figures 4 and 5. Maximum acceleration is 270 mm/s^2 before optimization, and then it becomes 200 mm/s^2 after optimization. Maximum acceleration reduces 26%. Thus, this planning trajectory is more reasonable.

The mechanical parameters of the transporting robot are shown in Table 1. According to these parameters, we can obtain $N_1 = 1.95$ and $H_m = 0.0236$. The transfer function is

$$G(s) = \frac{1.95}{0.0236s + 1} = \frac{82.63}{s + 42.37}. \quad (37)$$



(a) Position curve of the transporting robot

(b) Velocity curve of the transporting robot

(c) Acceleration curve of the transporting robot

Figure 6: Tracking control experiment curve of the LED wafer transporting robot.

Table 1: The mechanical parameters of the motor.

Rated torque	Torque constant	Rotation inertia	Rated acceleration	Damping coefficient
2.03 N·m	0.926 N·m/A	53.321×10^{-4} kgm^2	36700 rad/s^2	9.361×10^{-5} Nms/rad

The movement curves of the LED wafer transporting robot are shown in Figure 6, where the Kalman filter and PID controller are used to track the movement curve. In Figure 6(a), y_d is planned position curve, y is used PID controller to track output position curve, and y_2 is used Kalman filter to track position curve. In Figure 6(b), v_d is planned velocity curve, v is used PID controller to track output velocity curve, and v_2 is used Kalman filter to track output velocity curve. In Figure 6(c), a_d is planned acceleration curve, a is PID used controller to track output acceleration curve, and a_2 is used Kalman filter to track output acceleration curve. From the above curves, we can see that the motion of the transporting robot agrees more similarly with the planned trajectory. Position curve, velocity curve, and acceleration curve are all very smooth, and acceleration is very small. The tracking result of the Kalman filter is better than the PID controller.

6. Conclusion

The LED wafer transporting robot should open as fast as possible with small vibration. This paper proposes the fifth-order polynomial to plan the motion trajectory for the fast transporting process of the transporting robot. In order to minimize the maximum acceleration value during the motion, subsection searching method is designed to select the

fifth-order polynomial coefficient. The optimized fifth-order polynomial is simulated in MATLAB. At last, PID method and extended Kalman filter are used to track the planned curve. The simulation and experiment result show that the motion trajectory, velocity, and acceleration are smooth in the whole process. And acceleration is small, which satisfies the design requirements. Simulation and experiment show that the proposed subsection searching method for planning motion trajectory for the LED wafer transporting robot is very effective. The extended Kalman filter is applied to track motion trajectory of planning. Simulation and experiment show that the track effect of the extended Kalman filter is better than PID.

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 Science and Technology Innovation Project of Jinan under Grant no. 201309, National Natural Science Foundation of China under Grant no. 61403221, and Open Foundation of Key Laboratory of Pulp and Paper Science and Technology of Ministry of Education of China under Grant nos. KF201419 and 08031347.

References

[1] M. Brand, M. Masuda, N. Wehner, and X.-H. Yu, "Ant colony optimization algorithm for robot path planning," in *Proceedings of the International Conference on Computer Design and Applications (ICCDA '10)*, vol. 3, pp. V3436–V3440, June 2010.

[2] W. F. Xu, X. Q. Wang, Q. Xue, and B. Liang, "Study on trajectory planning of dual-arm space robot keeping the base stabilized," *Acta Automatica Sinica*, vol. 39, no. 1, pp. 69–80, 2013.

[3] W. Xu, Y. Liu, and Y. Xu, "The coordinated motion planning of a dual-arm space robot for target capturing," *Robotica*, vol. 30, no. 5, pp. 755–771, 2012.

[4] R. F. Li, M. C. Dokgo, L. Hu, and C. H. Han, "Mobile robot trajectory planning based on QPSO algorithm and experiment," *Control and Decision*, vol. 29, no. 12, pp. 2151–2157, 2013.

[5] D.-Q. Zhu and M.-Z. Yan, "Survey on technology of mobile robot path planning," *Control and Decision*, vol. 25, no. 7, pp. 961–967, 2010.

[6] Z. Yao and K. Gupta, "Path planning with general end-effector constraints," *Robotics and Autonomous Systems*, vol. 55, no. 4, pp. 316–327, 2007.

[7] H. H. Tian and Y. X. Su, "Nonlinear decentralized repetitive control for global asymptotic tracking of robot manipulators," *Acta Automatica Sinica*, vol. 37, no. 10, pp. 1264–1271, 2011.

[8] J. Q. Han, "From PID to active disturbance rejection control," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 3, pp. 900–906, 2009.

[9] J. K. Liu, *Advanced PID Control Based on MATLAB*, Publishing House of Electronics Industry, Beijing, China, 2011.

[10] L. Huang, M. L. Xi, and J. Sun, "Method of trajectory tracking control for mobile robots with improved QPSO algorithm," *Computer Engineering and Applications*, vol. 48, no. 34, pp. 230–236, 2012.

[11] F. Karami, J. Poshtan, and M. Poshtan, "Detection of broken rotor bars in induction motors using nonlinear Kalman filters," *ISA Transactions*, vol. 49, no. 2, pp. 189–195, 2010.

[12] P. H. Leong, S. Arulampalam, T. A. Lamahewa, and T. D. Abhayapala, "A Gaussian-sum based cubature Kalman filter for bearings-only tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 2, pp. 1161–1176, 2013.

[13] F. Sun and L.-J. Tang, "Estimation precision comparison of Cubature Kalman filter and Unscented Kalman filter," *Control and Decision*, vol. 28, no. 2, pp. 303–308, 2013.

[14] O. Straka, J. Duník, and M. Šimandl, "Gaussian sum unscented Kalman filter with adaptive scaling parameters," in *Proceedings of the 14th International Conference on Information Fusion (Fusion '11)*, pp. 1–8, Chicago, Ill, USA, July 2011.

[15] Z. Shi, Y.-Z. Wang, and Q.-L. Hu, "A polynomial interpolation based particle swarm optimization algorithm for trajectory planning of free-floating space robot," *Journal of Astronautics*, vol. 32, no. 7, pp. 1516–1521, 2011.

[16] L. Liu, C. Y. Chen, and X. H. Zhao, "Smooth trajectory planning for parallel manipulator with joint friction torque," *Journal of Mechanical Engineering*, vol. 50, no. 10, pp. 9–17, 2014.

[17] Y. Xiao, Z. Du, and W. Dong, "Smooth and near time-optimal trajectory planning of industrial robots for online applications," *Industrial Robot*, vol. 39, no. 2, pp. 169–177, 2012.

[18] J. Y. Shao, C. Q. Zhang, Y. Chen, and K. Chen, "Trajectory planning for redundant robots for internal surface spraying," *Journal of Tsinghua University (Science and Technology)*, vol. 54, no. 6, pp. 799–804, 2014.

[19] R. Kikuuwe, S. Yasukouchi, H. Fujimoto, and M. Yamamoto, "Proxy-based sliding mode control: a safer extension of PID position control," *IEEE Transactions on Robotics*, vol. 26, no. 4, pp. 670–683, 2010.

[20] R. Saravanan, S. Ramabalan, and C. Balamurugan, "Evolutionary multi-criteria trajectory modeling of industrial robots in the presence of obstacles," *Engineering Applications of Artificial Intelligence*, vol. 22, no. 2, pp. 329–342, 2009.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 502475, 12 pages
http://dx.doi.org/10.1155/2015/502475

Research Article

Stability Analysis for Autonomous Dynamical Switched Systems through Nonconventional Lyapunov Functions

V. Nosov, J. A. Meda-Campaña, J. C. Gomez-Mancilla, J. O. Escobedo-Alva, and R. G. Hernández-García

Sección de Estudios de Posgrado e Investigación, Escuela Superior de Ingeniería Mecánica y Eléctrica, Instituto Politécnico Nacional, 07738 México, DF, Mexico

Correspondence should be addressed to J. O. Escobedo-Alva; aerojonathan@yahoo.com.mx

Received 25 July 2015; Accepted 13 September 2015

Academic Editor: Rongwei Guo

The stability of autonomous dynamical switched systems is analyzed by means of multiple Lyapunov functions. The stability theorems given in this paper have finite number of conditions to check. It is shown that linear functions can be used as Lyapunov functions. An example of an exponentially asymptotically stable switched system formed by four unstable systems is also given.

1. Introduction

Switched systems are present in different areas of science and technology as aeronautical and automotive control, telecommunications, traffic control, chemical process, and so forth [1–5].

The switched system is a special class of hybrid or variable structure systems [1, 3, 6–10]. Similar to variable structure systems, the dynamics of switched systems is described by different differential equations in different space regions and the change of dynamics occurs when the trajectories pass through the boundaries between two regions. Variable structure systems may have special type of solutions, the so-called sliding mode solutions. The theory and different applications of sliding mode solutions in control are investigated in many books and papers (see, e.g., [9]). In contrast to sliding mode theory, the case when the variable structure system has no sliding mode solution is not sufficiently studied. To separate these two cases, we call variable structure systems without sliding mode solutions as switched systems. In the theory of hybrid systems the change of systems dynamics may occur by action of automata or by other reasons. In this paper, the stability of switched systems is studied. The problem of stability of switched systems is not simple. First of all, we give some examples illustrating different aspects of this problem and showing that stability of all subsystems is not sufficient to ensure the stability of the whole switched system. Namely, in one example the whole switched system formed by stable subsystems is unstable or asymptotically stable depending on the structure of the switched systems, that is, depending on the regions where these subsystems are acting. If some of subsystems are unstable and others are asymptotically stable, the switched system may be unstable or asymptotically stable or it may have periodic solutions. Also, in some cases the stability of the whole switched system depends neither on the structure of the switched system nor on the order of the switching.

The papers [11, 12] were dedicated essentially to the case of linear systems. In paper [13] the stability of switched and hybrid systems is investigated. However, for each trajectory, the theorem on stability from [13] imposes certain conditions on Lyapunov functions at the moment of passing through the switching lines. These conditions are possible to check only if the trajectory is known. Furthermore, typical trajectory of switched systems has an infinite number of switchings and it would be necessary to check an infinite number of conditions. Our previous works [14, 15] investigate some stability problems in switched systems and also investigate possibilities of appearance of chaotic solutions in such systems.

In present paper, some stability theorems using multiple Lyapunov functions are established. In contrast to [13] our conditions are imposed on values of Lyapunov functions

in corresponding space regions and on switching lines, allowing in this way obtaining different stability conditions without knowing trajectories, while only a finite number of conditions depending only on Lyapunov functions are verified. The Lyapunov functions used in these theorems may be different from usual Lyapunov functions defined in the whole space, which allows extending the class of functions used as Lyapunov functions. For example, it is possible to use linear Lyapunov functions to investigate the stability of switched systems. The sum of quadratic and linear functions may be used also as Lyapunov functions. Some examples of such functions are also given. The using of unusual Lyapunov functions simplifies the search for convenient Lyapunov function. We present also one example (Example 15) where all subsystems are unstable by considering the whole space, but the switched system formed by these subsystems is exponentially asymptotically stable.

Our theorems which give conditions of asymptotic stability or instability of whole switched system in case when all subsystems are only stable are especially interesting. These results may be considered as a special type of parametric excitation or parametric stabilization of stable systems.

2. Description of Autonomous Two-Dimensional Switched Systems

Suppose the phase space $\mathbb{R}^2$, $X = (x, y)^T \in \mathbb{R}^2$, with norm $\|X\|$, is divided into a finite number p of open 1-connected regions Q_i, $1 \le i \le p$, with smooth boundaries such that the origin $\{0\}$ of rectangular coordinate systems pertains to closure of any region Q_i. The boundary between two regions Q_i and Q_{i+1} is noted by $L_{i,i+1}$ and is called switching line. These switching lines are supposed to be smooth and let the normal $N_{i,i+1}(X)$ to the switching lines existing in each point $X \in L_{i,i+1}$ have direction from Q_i to Q_{i+1}. Suppose each region Q_i has points X with norms such that $\|X\| > H$, where H is any number. The topological properties of plane conduce to the conclusion that each region Q_i has only two boundaries which go from origin $\{0\}$ to infinity without intersections.

In each region Q_i the dynamics of switched systems is described by a proper autonomous equation E_i with Lipschitz continuous function $f_i(X) = f_i(x, y)$:

$$E_i\colon \dot{X} = f_i(X), \quad X \in Q_i,\ 1 \le i \le p, \tag{1}$$

with initial condition

$$X(t_0) = X_0. \tag{2}$$

The trajectory of switched system (1)-(2) may pass from region Q_i to region Q_{i+1} only crossing the switching line $L_{i,i+1}$. Suppose also the nonexistence of sliding modes in switched system (1)-(2). The sufficient condition for absence of sliding modes is the following *transversality condition*: the normal $N_{i,i+1}(X)$ which goes from Q_i to Q_{i+1} and the trajectory velocities $f_i(X)$ and $f_{i+1}(X)$ from one to another side of switching line form acute angles; that is,

$$\begin{aligned} &\operatorname{sign}\left(\left\langle N_{i,i+1}(X), f_i(X)\right\rangle\right) \\ &\quad = \operatorname{sign}\left(\left\langle N_{i,i+1}(X), f_{i+1}(X)\right\rangle\right) = 1, \quad X \in L_{i,i+1}, \end{aligned} \tag{3}$$

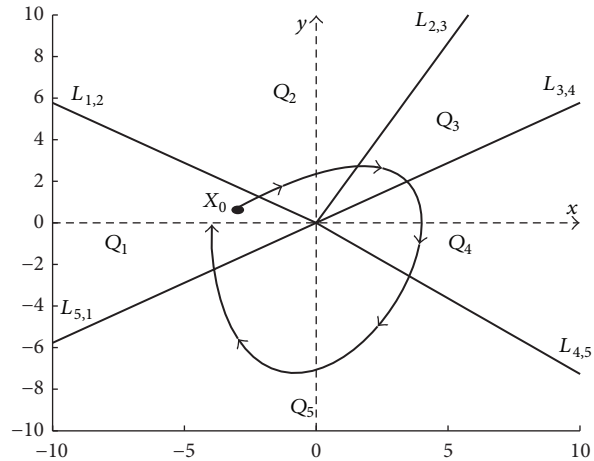


Figure 1: Autonomous two-dimensional switched system.

for every point on switching line $L_{i,i+1}$ and all lines $L_{i,i+1}$, $i = 1, 2, 3, \ldots, p$ (Figure 1). Transversality condition (3) guarantees the passage only from region Q_i to Q_{i+1} and not in opposite direction.

If at instant $t_{i,\text{fin}}$ the trajectory of switched system arrives from region Q_i on the switching line $L_{i,i+1}$, the trajectory passes across the switching line $L_{i,i+1}$ and at the following instants $t > t_{i,\text{fin}}$ the system dynamic is described by equation E_{i+1} acting in region Q_{i+1} with initial condition $X_{t_{i+1,\text{init}}}$ for the new equation E_{i+1} coinciding with the final condition $X_{t_{i,\text{fin}}}$ of the previous equation E_i on the switching line $L_{i,i+1}$, for $i = 1, 2, 3, \ldots$; that is, $X_{t_{i+1,\text{init}}} = X_{t_{i,\text{fin}}}$. Under this condition, all trajectories of switched system will be continuous for all $t > t_0$. The regions Q_i together with equations E_i completely define the switched system. The set of regions Q_i defines the geometrical structure while the set of equations E_i defines the dynamical structure of switched system.

Any switched system for a chosen initial condition $X(t_0) = X_0$ generates a sequence of continuous dynamical subsystems $S_{1(X_0)}, S_{2(X_0)}, \ldots, S_{i(X_0)}, \ldots$ acting in regions $Q_{1(X_0)} = Q_k$, $1 \le k \le p$, $Q_{2(X_0)} = Q_{k+1}, \ldots, Q_{i(X_0)}, \ldots$ and switching on lines $L_{1(X_0),2(X_0)} = L_{k,k+1}, L_{2(X_0),3(X_0)}, \ldots, L_{i(X_0),(i+1)(X_0)}, \ldots$; that is,

$$\operatorname{SW}(X_0) = \left\{S_{1(X_0)}, S_{2(X_0)}, \ldots, S_{i(X_0)}, \ldots\right\}. \tag{4}$$

The initial condition X_0 defines the first element in (4). Depending on X_0, the first subsystem $S_{1(X_0)}$ is acting from initial time $t_0 = t_{1(X_0),\text{init}} = t_{k,\text{init}}$ to the first switching instant $t_{1(X_0),\text{fin}} = t_{k,\text{fin}}$ in region Q_k such that $X(t_0) \in Q_k$. The second subsystem $S_{2(X_0)}$ is acting in Q_{k+1} from time $t_{2,\text{init}} = t_{1,\text{fin}}$ to the second switching instant $t_{2,\text{fin}} = t_{3,\text{init}}$ and so on. In other words, each subsystem S_i is defined by three elements $S_i = (Q_i, E_i, X(t_{i,\text{init}}))$:

(a) region Q_i where this subsystem is acting,

(b) differential equation E_i acting in Q_i,

(c) initial values $t_{i,\text{init}}$ and $X(t_{i,\text{init}})$.

This definition is slightly different from definition given in [1, 3, 13]. The sequence (4) may contain finite or infinite number of dynamical subsystems S_i although the total number of different regions Q_i is finite because trajectory can return to region Q_i after the whole rotation around origin, and therefore $Q_{p+i} = Q_i$ for all i. If condition (3) holds for all switching lines, then the sequence (4) may contain an infinite number of subsystems $S_1, S_2, \ldots, S_n, \ldots$. In case of infinite number of subsystems S_i, in sequence (4) each region Q_i has two boundary switching lines: by one of them the trajectories enter from the precedent region Q_{i-1}, and by the other, after a finite time, the trajectories go out from region Q_i to region Q_{i+1}. After the whole rotation around origin the switched system returns to initial region Q_i and $L_{p,p+1} = L_{p,1}$.

This property has, as a consequence, the following fundamental property of autonomous two-dimensional switched systems: the order of terms in sequence (4) does not depend on initial conditions.

In multidimensional phase space $\mathbb{R}^n$, $n > 2$, each region may have more than two switching surfaces. For that reason, there does not exist an analogy of announced fundamental property in multidimensional space. So, in $\mathbb{R}^n$ the order of terms in (4) may depend on initial conditions. This complicates the stability investigation of switched systems in multidimensional space $\mathbb{R}^n$, $n > 2$.

The number of dynamical subsystems $S_{i(X_0)}$ is finite, $1 \le i \le m$, if the switched system stays in the final region Q_N for all time after last switching $t_{m-1,\text{fin}}$, $t > t_{m-1,\text{fin}}$. The switched system may have more than one final region and the final region may depend on initial conditions X_0. The sufficient condition for finiteness of sequence (4) is the existence of at least one line $L_{m,m+1}$ such that from one side of the line $L_{m,m+1}$ the normal $N_{m,m+1}(X)$ and the trajectory velocities $f_m(X)$ and $f_{m+1}(X)$ form an obtuse and acute angle, respectively:

$$\begin{aligned} \operatorname{sign}\left(\left\langle N_{m,m+1}(X), f_m(X)\right\rangle\right) &= -1, \\ \operatorname{sign}\left(\left\langle N_{m,m+1}(X), f_{m+1}(X)\right\rangle\right) &= 1, \end{aligned} \tag{5}$$

for $X \in L_{m,m+1}$.

In this case, the subsystem S_m will be the last subsystem in the sequence (4). If there exist more than one switching line satisfying condition (5) then the switched system may have more than one final region and the final region may depend on initial conditions X_0. Let us consider two examples.

Example 1. Suppose the switched system is described by two pendulum equations E_1 and E_2 with different natural frequencies acting in regions Q_1 and Q_2 of phase plane $\mathbb{R}^2$:

$$\begin{aligned} E_1\colon \dot{x} &= y, \\ \dot{y} &= -x, \\ &\forall X \in Q_1 = \mathbb{R}^2 - Q_2, \end{aligned} \tag{6}$$

while equation E_2 is

$$\begin{aligned} E_2\colon \dot{x} &= y, \\ \dot{y} &= -\omega^2 x, \\ &\forall X \in Q_2 = \{x < 0, y > 0\}, \end{aligned} \tag{7}$$

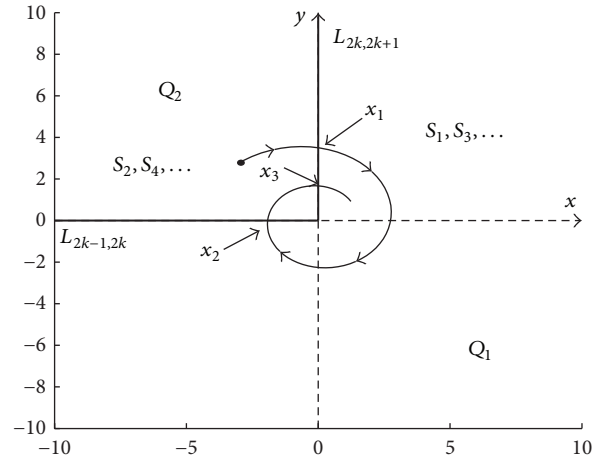


Figure 2: Switched system with infinite sequence (4) of subsystems S_i.

Every trajectory of switched system has an infinite number of switchings and it is represented by an infinite number of continuous dynamical subsystems $S_1, S_2, \ldots, S_i, \ldots$.

The subsystems $S_1, S_3, \ldots, S_{2j+1}, \ldots$ are described by the same equation E_1 in the form of (6) acting in region Q_1 and the subsystems $S_2, S_4, \ldots, S_{2j}, \ldots$ are described by equation E_2 in the form of (7) acting in region Q_2. But the initial conditions for $S_1, S_3, \ldots, S_{2j+1}, \ldots$ and $S_2, S_4, \ldots, S_{2j}, \ldots$ are different: X_1 for S_1, X_3 for S_3, and so on. Therefore all subsystems $S_1, S_2, \ldots, S_i, \ldots$ are different (see Figure 2).

Example 2. Consider now the switched system with E_1 acting in the region $Q_1 = \{x < 0\} \cap \{y > 0\} \cap \{y < (-2 + \sqrt{3})x\}$, E_2 acting in the lower half-space $Q_2 = \{y < 0\}$, and E_3 acting in $Q_3 = \mathbb{R}^2 - Q_1 - Q_2$ with

$$E_1\colon \ddot{x} + 4x + x = 0, \quad \forall X \in Q_1, \tag{8}$$

$$E_2\colon \ddot{x} + \omega^2 x = 0, \quad \forall X \in Q_2, \tag{9}$$

$$E_3\colon \ddot{x} + x = 0, \quad \forall X \in Q_3. \tag{10}$$

The equation E_1 is overdamped and its general solution is

$$\begin{aligned} x(t) = A \cdot \exp\left(\left(-2 + \sqrt{3}\right)t\right) + B \\ \cdot \exp\left(\left(-2 - \sqrt{3}\right)t\right), \quad A, B = \text{const}. \end{aligned} \tag{11}$$

The trajectories of the whole switched system are presented in Figure 3 with $\omega = 3$. The region denoted as Q_1 in Figure 3 is a final region of the switched system (8)–(10). Depending on initial conditions, the trajectory may has two, one, or zero commutations. If X_0 belongs to region Q_3 then the trajectory has two commutations, if X_0 belongs to region Q_2 then the trajectory has one commutation, and if X_0 belongs to region Q_1 then the trajectory has no commutations.

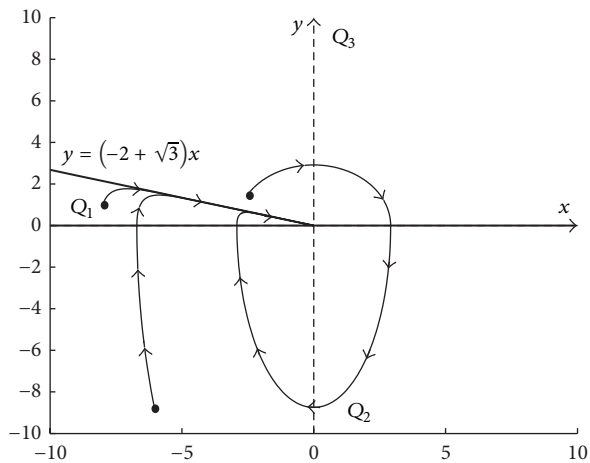


Figure 3: Switched system with finite sequence of subsystems (8)-(9).

3. Stability of Switched Systems

Suppose

$$f_i(0) = 0, \quad 1 \le i \le p. \tag{12}$$

Under condition (12) switched system (1)-(2) or (4) has the trivial solution $X(t) \equiv 0$.

In the following, the stability analysis of trivial solution (or origin) of switched system (1)-(2) will be carried out.

Let us introduce some definitions of Lyapunov stability.

Definition 3. The trivial solution $X(t) \equiv 0$ (or origin) of switched system (1)-(2) is said to be stable if for any $\epsilon > 0$ there exists a $\delta = \delta(\epsilon)$ such that the inequality $\|X(t, t_0, X_0)\| < \epsilon$ is satisfied for any time $t > t_0$ whenever $\|X_0\| < \delta(\epsilon)$.

Definition 4. The trivial solution $X(t) \equiv 0$ (or origin) of system (1)-(2) is said to be asymptotically stable if

(a) it is stable;

(b) there exists $\Delta > 0$ such that $\|X(t, t_0, X_0)\| \to 0, t \to \infty$ for $\|X_0\| < \Delta$.

Clearly, the stability (asymptotic stability) is uniform with respect to t_0 because the switched system is a stationary system with all elements independent of t.

Obviously, in case of finite sequence (4) the stability or instability of the whole switched system origin depends only on stability of instability of origin for final subsystem S_m.

In case of infinite sequence (4) the instability of one equation E_i does not conduce automatically to the instability of the whole switched system. Also, the stability of all equations E_i is not sufficient to conclude that the whole switched system is stable [15].

Let us give other corresponding examples.

Example 5. Consider the switched system with two equations E_1 and E_2 acting in two half-planes $Q_1 = \{\dot{x} = y > 0\}$ and $Q_2 = \{\dot{x} = y < 0\}$, respectively, with

$$E_1\colon \ddot{x} - 0.2\dot{x} + 1.01x = 0, \quad (x, \dot{x} = y) \in Q_1, \tag{13}$$

$$E_2\colon \ddot{x} + 0.4\dot{x} + 1.04x = 0, \quad (x, \dot{x} = y) \in Q_2. \tag{14}$$

Switching lines $L_{1,2}$ and $L_{2,1}$ in this case are $L_{1,2} = \{x > 0, y = 0\}$ and $L_{2,1} = \{x < 0, y = 0\}$. Equation E_1 is unstable and E_2 is asymptotically stable. The asymptotic stability of equation E_2 can be called stronger than the instability of E_1. It is easy to verify that for any $X_1 = \{x_1 > 0, y_1 = 0\}$ the subsequent points, $X_2 = \{x_2 < 0, y_2 = 0\}$, $X_3 = \{x_3 > 0, y_3 = 0\}, \ldots$, of intersection with the axis x are equal to

$$\begin{aligned} x_2 &= -x_1 \exp(-0.2\pi), \\ x_3 &= -x_2 \exp(0.1\pi) = x_1 \exp(-0.1\pi) < x_1. \end{aligned} \tag{15}$$

Therefore, the trivial solution of switched system (13)-(14) is asymptotically stable.

Replace (14) in region Q_2 by

$$E_3\colon \ddot{x} + 0.2\dot{x} + 1.01x = 0, \quad (x, \dot{x} = y) \in Q_2. \tag{16}$$

Now, the same calculations show that $x_3 = x_1$ for all x_1. Therefore all solutions of switched system formed by unstable equation (13) and by asymptotically stable equation (16) are periodic.

Example 6. Consider switched system of Example 1:

$$E_1\colon \ddot{x} + \omega^2 x = 0, \quad (x, \dot{x} = y) \in Q_1 = \{x < 0, \dot{x} > 0\}, \tag{17}$$

$$E_2\colon \ddot{x} + x = 0, \quad (x, \dot{x} = y) \in Q_2 = \mathbb{R}^2 - Q_1. \tag{18}$$

Both equations E_1 and E_2 are stable. Also, it is easy to calculate for any $X_1 = \{x_1 = 0, y_1 > 0\}$ the values of $X_2 = \{x_2 = -y_1 < 0, y_2 = 0\}$, $X_3 = \{x_3 = 0, y_3 > 0\}$ and that

$$y_3 = \omega y_1. \tag{19}$$

Therefore, if $\omega < 1$ then the switched system origin is asymptotically stable, but if $\omega > 1$ the switched system origin is unstable. Suppose now Q_1 coincides with the first quadrant; that is, $Q_1 = \{x > 0, y > 0\}$; then analogous calculations show that the switched system origin is asymptotically stable if $\omega > 1$, but if $\omega < 1$ the switched system origin is unstable.

Remark 7. These two examples show that stability of switched systems depends not only on the stability of equations E_i but also on all other elements which define the switched system; that is, it depends also on the regions where different equations E_i are acting and on the order of the switching.

Thus, the following definitions are justified.

Definition 8. The origin of switched system (4) formed by an infinite sequence of subsystems $S_{1(X_0)}, S_{2(X_0)}, \ldots, S_{i(X_0)}, \ldots$ is stable (asymptotically stable) independently of geometrical structure of switched system if the origin is stable (asymptotically stable) for any choice of equations $E_1, \ldots, E_p$ acting in arbitrary chosen regions $Q_1, \ldots, Q_p$.

Definition 9. The origin of switched system (4) formed by a sequence of subsystems $S_1, S_2, \ldots, S_i, \ldots$ is stable (asymptotically stable) for given geometrical and dynamical structure of switched system if the origin is stable (asymptotically stable) for a given choice of equations $E_1, \ldots, E_p$ acting in corresponding regions $Q_1, \ldots, Q_p$.

4. General Theorems

The previous examples show that the stability properties of switched system origin need special investigation and they do not follow directly from simple stability conditions imposed on subsystems, but also they depend on how the switching occurs within the whole system.

The only interesting case when sequence (4) which determines the switched system is infinite is considered below.

Suppose equations E_i are stable (asymptotically stable) and also suppose the existence of some smooth positive definite functions $V_i(X)$ defined for $X \in Q_i$. Each function $V_i(X)$ is called Lyapunov function in region Q_i, if it is continuously differentiable in Q_i and fulfilling $V_i(X) > 0$, $X \in Q_i$, $X \neq 0$, and $V_i(0) = 0$ [16, 17].

Denote by $\omega_k(u)$ scalar continuous nondecreasing functions (also called wedges) defined and positive for $u > 0$ such that $\omega_k(0) = 0$.

The symbols $\dot{V}_i(X(t))$ denote the derivatives of the functions $V_i(X(t))$ along the trajectory of equation E_i:

$$\dot{V}_i(X(t)) = \langle \nabla(V(X)), f_i(X) \rangle, \quad X \in Q_i. \tag{20}$$

The existence of a common Lyapunov function for all equations E_i simplifies the stability analysis of switched system origin. For this reason, Theorem 10 has been explicitly formulated but not proven, because this theorem is a direct consequence of the more general Theorem 12.

Theorem 10. *Suppose there exists a common Lyapunov function $V(X)$ defined in the whole space $\mathbb{R}^2$ satisfying the following:*

(a) $\omega_1(\|X\|) \leq V(X) \leq \omega_2(\|X\|)$, $X \in \mathbb{R}^2$,

(b) *for any equation E_i the following conditions are fulfilled*

$$\dot{V}(X(t)) = \langle \nabla(V(X)), f_i(X) \rangle \leq 0, \quad \left(\dot{V}_i(X) \leq -\omega_3(\|X\|) < 0\right), \tag{21}$$

for $X \in Q_i$, $1 \leq i \leq p$.

Then, the trivial solution $X(t) \equiv 0$ of switched system (1)-(2) is stable (asymptotically stable) independently of geometrical structure of switched system.

Example 11. Consider three equations:

$$\begin{aligned} E_1&: \ddot{x} + x = 0, \\ E_2&: \ddot{x} + x + x^2\dot{x}^3 = 0, \\ E_3&: \ddot{x} + \sin(\dot{x}) + x = 0. \end{aligned} \tag{22}$$

As a common Lyapunov function $V(X)$ we can take the function

$$V(x, \dot{x}) = x^2 + \dot{x}^2. \tag{23}$$

The derivatives of these functions along the trajectories of equations E_1, E_2, and E_3 are not positive:

$$\begin{aligned} \dot{V}_{(ec21)}(x, \dot{x}) &= 2x\dot{x} - 2x\dot{x} = 0, \\ \dot{V}_{(ec22)}(x, \dot{x}) &= 2x\dot{x} - 2\dot{x}\left(-x - x^2\dot{x}^3\right) = -2x^2\dot{x}^4 \leq 0, \\ V_{(ec23)}(x, \dot{x}) &= 2x\dot{x} - 2\dot{x}(-x - \sin(\dot{x})) = -2\dot{x}\sin(\dot{x}) \\ &\leq 0, \\ &\qquad |\dot{x}| < 1. \end{aligned} \tag{24}$$

Therefore, the trivial solution of the switched system formed by equations E_1, E_2, and E_3 is stable independently of geometrical structure of switched system determined by a choice of arbitrary regions Q_1, Q_2, and Q_3 such that $Q_1 \cup Q_2 \cup Q_3 = \mathbb{R}^2$.

Theorem 12. *Suppose conditions (3) hold for all switching lines $L_{i,i+1}$ and suppose that for each equation E_i acting in $Q_i \in \mathbb{R}^2$ there exists a Lyapunov function $V_i(X)$ defined in Q_i such that*

(a) $\omega_{1,i}(\|X\|) \leq V_i(X) \leq \omega_{2,i}(\|X\|)$, $X \in Q_i$, $i = 1, \ldots, p$,

(b) $\dot{V}_i(X(t)) = \langle \nabla(V_i(X)), f_i(X) \rangle \leq 0$, $(\dot{V}_i(X(t)) \leq -\omega_{3,i}(\|X(t)\|) < 0)$, $X \in Q_i$, $i = 1, \ldots, p$,

(c) *on all switching lines $L_{i,i+1}$ where trajectories pass from region Q_i to Q_{i+1}, the following inequalities hold: $V_i(X) \geq V_{i+1}(X)$, $X \in L_{i,i+1}$.*

Then the trivial solution $X(t) \equiv 0$ of switched system (4) is stable (asymptotically stable) for given switched system.

Proof. Let us proof stability of trivial solution $X(t) \equiv 0$. Denote

$$\begin{aligned} \omega_1(u) &= \min_i(\omega_{1,i}(u)), \quad u > 0, \\ \omega_2(u) &= \max_i(\omega_{2,i}(u)), \quad u > 0, \\ \omega_3(u) &= \min_i(\omega_{3,i}(u)), \quad u > 0, \\ \omega_1(0) &= \omega_2(0) = \omega_3(0) = 0. \end{aligned} \tag{25}$$

The wedge functions $\omega_i(u)$, $i = 1, 2, 3$, are scalar continuous nondecreasing functions positive for $u > 0$ and satisfy (25). For a given $\epsilon > 0$ we define a number $\delta = \delta(\epsilon)$ such that $\omega_2(\delta) \leq \omega_1(\epsilon)$. Evidently, such δ does exist. Using conditions (b) and (c) we obtain that the sequence of Lyapunov functions

$V_k(X(t,t_0,X_0))$ for arbitrary k and $t_{k,\text{init}} < t < t_{k,\text{fin}}$ is nonincreasing on the trajectories of switched system:

$$\begin{aligned}\omega_1(\|X(t,t_0,X_0)\|) &\le \omega_{1,k}(\|X(t,t_0,X_0)\|)\\ &\le V_k(X(t,t_0,X_0))\\ &\le V_k(X(t_{k,\text{init}},t_0,X_0))\\ &\le V_{k-1}(X(t_{k-1,\text{fin}},t_0,X_0))\\ &\le V_{k-1}(X(t_{k-1,\text{init}},t_0,X_0)) \le \cdots\\ &\le V_{1(X_0)}(X_0).\end{aligned} \tag{26}$$

Using now condition (a) of the theorem, we obtain for arbitrary k and t, $t_{k,\text{init}} < t < t_{k,\text{fin}}$ the following inequalities:

$$\begin{aligned}\omega_1(\|X(t,t_0,X_0)\|) &\le V_{1(X_0)}(X_0) \le \omega_{2,1}(\|X_0\|)\\ &\le \omega_2(\|X_0\|) \le \omega_2(\delta) \le \omega_1(\epsilon),\end{aligned} \tag{27}$$

Therefore,

$$\|X(t,t_0,X_0)\| \le \epsilon, \quad t \ge t_0, \ \|X_0\| \le \delta. \tag{28}$$

Inequality (28) is equivalent to stability of trivial solution of switched system (1)-(2).

To proof asymptotic stability it is necessary to demonstrate that for any $\gamma > 0$ there exist numbers $\Delta > 0$ and $T(\gamma)$ such that $\|X(t,t_0,X_0)\| \le \gamma$ for $t \ge t_0 + T(\gamma)$ and $\|X_0\| \le \Delta$. Take any $\gamma > 0$ and determine $\delta > 0$ which corresponds to γ in demonstration of stability; that is, $\omega_2(\delta) \le \omega_1(\gamma)$. Take also $T(\gamma) = 2w_2(\Delta)/\omega_3(\delta) > 0$.

Let us demonstrate that on the interval $[t_0, t_0+T(\gamma)]$ there exists an instant t_1 such that $\|X(t_1,t_0,X_0)\| \le \delta$.

If this is not so, that is, $\|X(t,t_0,X_0)\| > \delta$ for all $t \in [t_0, t_0+T(\gamma)]$, then using continuity of $X(t,t_0,X_0)$ and conditions (b) and (c) of Theorem 12 we have

$$\begin{aligned}&V_k(X(t_0+T(\gamma),t_0,X_0))\\ &\quad\le V_{1(X_0)}(X_0) + \int_{t_0}^{t_{1,\text{fin}}} \dot V_{1(X_0)}(X(t,t_0,X_0))\,dt\\ &\qquad + \int_{t_{1,\text{fin}}}^{t_{2,\text{fin}}} \dot V_{2(X_0)}(X(t,t_0,X_0))\,dt + \cdots\\ &\qquad + \int_{t_{k,\text{init}}}^{t_0+T(\gamma)} \dot V_k(X(t,t_0,X_0))\,dt\\ &\quad\le V_{1(X_0)}(X_0) - \int_{t_0}^{t_{1,\text{fin}}} \omega_{3,1}(\|X(t,t_0,X_0)\|)\,dt\\ &\qquad - \int_{t_{1,\text{fin}}}^{t_{2,\text{fin}}} \omega_{3,2}(\|X(t,t_0,X_0)\|)\,dt - \cdots\\ &\qquad - \int_{t_{k,\text{init}}}^{t_0+T(\gamma)} \omega_{3,k}(\|X(t,t_0,X_0)\|)\,dt\\ &\quad\le V_{1(X_0)}(X_0) - \int_{t_{k,\text{init}}}^{t_0+T(\gamma)} \omega_3(\|X(t,t_0,X_0)\|)\,dt\\ &\quad\le V_{1(X_0)}(X_0) - T(\gamma)\,\omega_3(\delta)\\ &\quad\le \omega_2(\Delta) - T(\gamma)\,\omega_3(\delta) = -\omega_2(\Delta) < 0.\end{aligned} \tag{29}$$

Inequality (29) contradicts the positiveness of Lyapunov function V_k. Therefore, there exists an instant $T(\gamma)$ and t_1, $t_0 \le t_1 \le t_0 + T(\gamma)$ such that $\|X(t_1,t_0,X_0)\| \le \delta$ for $\|X_0\| \le \Delta$. From condition $w_2(\delta) \le \omega_1(\gamma)$ and the stability of the trivial solution it follows that $\|X(t,t_0,X_0)\| \le \gamma$ for all $t \ge t_1$, $t_1 \le t_0 + T(\gamma)$. Therefore $\|X(t,t_0,X_0)\| \le \gamma$ for all $\|X_0\| \le \Delta$ and for all $t \ge t_0 + T(\gamma)$. The asymptotic stability of the origin is proven. □

Example 13. Consider once again first switched system of Example 6 where Q_1 coincides with second quadrant, $Q_1 = \{x < 0, y > 0\}$. Lyapunov functions for equations E_1, E_2 are

$$\begin{aligned}V_1(x,y) &= \omega^2x^2 + y^2,\\ V_2(x,y) &= x^2 + y^2,\end{aligned} \tag{30}$$

respectively.

On switching line $L_{1,2}$ where $x = 0$ and $y > 0$, we have $V_1(x,y) = y^2 = V_2(x,y)$. On switching line $L_{2,1}$ where $y = 0$ we have $V_1(x,y) = \omega^2x^2$ and $V_2(x,y) = x^2$. Therefore, condition (c) of Theorem 12 holds on both lines $L_{1,2}$ and $L_{2,1}$ if $\omega^2 < 1$. Under this condition, the origin of the switched system described by (17)-(18) of Example 6 is stable.

Consider now the second switched system of Example 6, where Q_1 coincides with the first quadrant $Q_1 = \{x \ge 0, y \ge 0\}$. In this case, on line $L_{2,1}$, where $x = 0$ we have $V_1(x,y) = y^2 = V_2(x,y)$ and on line $L_{1,2}$ where $y = 0$ we have $V_1(x,y) = \omega^2x^2$ and $V_2(x,y) = x^2$. Therefore condition (c) of Theorem 12 holds if $\omega^2 > 1$. In this case, the origin of the second switched system of Example 6 is stable under this situation. This conclusion coincides with results of Example 6 obtained by direct analytical computations. Analytical calculations show asymptotic stability of switched system origin (17)-(18), while by using Theorem 12 we can only establish stability but no asymptotic stability.

Replacing (18) in region Q_2 by new equation

$$E_2\colon \ddot x + x + x^2\dot x^3 = 0, \quad X \in Q_2, \tag{31}$$

and considering the same Lyapunov function $V_2(x,y)$, we can also establish stability of the trivial solution of this new switched system (17), (31), but analytical calculations cannot be used in this case.

Example 14. Consider the switched system with two equations E_1 and E_2 acting in two half-planes $Q_1 = \{y > 0\}$ and $Q_2 = \{y < 0\}$:

$$\begin{aligned}E_1\colon \dot x &= y + x^2,\\ \dot y &= -2xy - 2x^3,\\ &\qquad \{x,y\} \in Q_1,\end{aligned}$$

$$E_2\colon \dot{x} = y,\quad \dot{y} = -x,\qquad \{x, y\} \in Q_2. \tag{32}$$

The switching lines are $L_{1,2} = \{x > 0, y = 0\}$ and $L_{2,1} = \{x < 0, y = 0\}$. Normal $N_{1,2}$ going from Q_1 to Q_2 is $N_{1,2} = \left\{\begin{smallmatrix} 0 \\ -1 \end{smallmatrix}\right\}$, and condition (3) on line $L_{1,2}$ has the form

$$\begin{aligned} \langle N_{1,2}, f_1\rangle &= 2x^3 > 0, \\ \langle N_{1,2}, f_2\rangle &= x > 0. \end{aligned} \tag{33}$$

On the other hand, normal $N_{2,1}$ going from Q_2 to Q_1 is $N_{2,1} = \left\{\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right\}$, and condition (3) on line $L_{2,1}$ has the form

$$\begin{aligned} \langle N_{2,1}, f_1\rangle &= -2x^3 > 0, \\ \langle N_{2,1}, f_2\rangle &= -x > 0. \end{aligned} \tag{34}$$

As Lyapunov functions for equations E_1 and E_2 in regions $y > 0$ and $y < 0$ take

$$V_1(x, y) = x^2 + y > 0, \quad y > 0, \tag{35}$$

$$V_2(x, y) = x^2 + y^2, \quad y < 0. \tag{36}$$

The strange function (35) satisfies condition (a) of Theorem 12 for unusual norm in $\mathbb{R}^2$ of the form $\|X\| = x^2 + |y|$. The switching lines are $L_{1,2} = \{y = 0, x > 0\}$ and $L_{2,1} = \{y = 0, x < 0\}$ and on these switching lines $V_1(x, y) = x^2 = V_2(x, y)$. The derivatives of functions (35) and (36) in regions $y > 0$ and $y < 0$ are equal to

$$\begin{aligned} \dot{V}_1(x, y) &= 2x\dot{x} + \dot{y} = 2x\left(y + x^2\right) + \left(-2xy - 2x^3\right) \\ &= 0, \quad y > 0; \\ \dot{V}_2(x, y) &= 2x\dot{x} + 2y\dot{y} = 2xy - 2xy = 0, \quad y < 0. \end{aligned} \tag{37}$$

All conditions of Theorem 12 hold and therefore, the trivial solution of switched system (32) is stable.

Example 15. Consider a switched system with four equations acting in four quadrants of the plane as shown in Figure 4:

$$\begin{aligned} E_1\colon\ & \dot{x} = 3x + 2y,\quad \dot{y} = -4x - 3y, \qquad (x, y) \in Q_1 = \{x > 0, y > 0\}, \\ E_2\colon\ & \dot{x} = -3x + 5y,\quad \dot{y} = -2x + 4y, \qquad (x, y) \in Q_2 = \{x > 0, y < 0\}, \\ E_3\colon\ & \dot{x} = 3x + y,\quad \dot{y} = -4x - 2y, \qquad (x, y) \in Q_3 = \{x < 0, y < 0\}, \\ E_4\colon\ & \dot{x} = -4x + 3y,\quad \dot{y} = -2x + 2y, \qquad (x, y) \in Q_4 = \{x < 0, y > 0\}, \end{aligned} \tag{38}$$

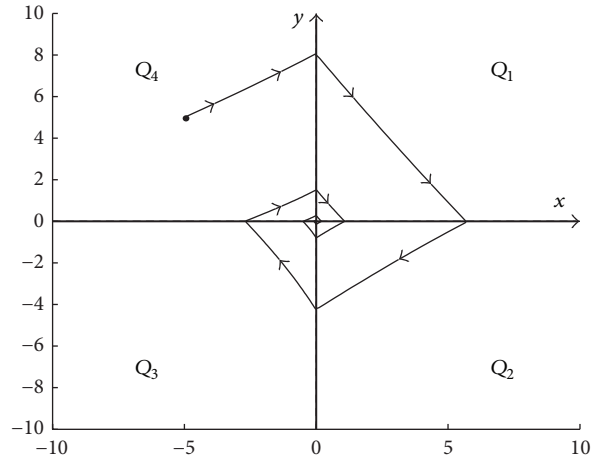


Figure 4: Trajectory of switched system (38).

where $Q_1 \cup Q_2 \cup Q_3 \cup Q_4 = \mathbb{R}^2$.

It is easy to observe that equations E_1, E_2, E_3, and E_4 are unstable in the whole plane $\mathbb{R}^2$. To analyze the stability of the switched system origin consider the following four linear Lyapunov functions defined in corresponding regions:

$$\begin{aligned} V_1(x, y) &= x + y, \qquad V_1(x, y) > 0,\ (x, y) \in Q_1 = \{x > 0, y > 0\}, \\ V_2(x, y) &= x - y, \qquad V_2(x, y) > 0,\ (x, y) \in Q_2 = \{x > 0, y < 0\}, \\ V_3(x, y) &= -x - y, \qquad V_3(x, y) > 0,\ (x, y) \in Q_3 = \{x < 0, y < 0\}, \\ V_4(x, y) &= -x + y, \qquad V_4(x, y) > 0,\ (x, y) \in Q_4 = \{x < 0, y > 0\}. \end{aligned} \tag{39}$$

Clearly, functions (39) satisfy conditions (a) of Theorem 12 if the norm in $\mathbb{R}^2$ is equal to $\|X\|_1 = |x| + |y|$. The derivatives of functions (39) are

$$\begin{aligned} \dot{V}_1(x, y) &= \dot{x} + \dot{y} = 3x + 2y - 4x - 3y = -x - y \\ &= -V_1, \quad (x, y) \in Q_1, \end{aligned}$$

$$\begin{aligned}
\dot{V}_2(x,y) &= \dot{x} - \dot{y} = -3x + 5y + 2x - 4y = -x + y \\
&= -V_2, \quad (x,y) \in Q_2, \\
\dot{V}_3(x,y) &= -\dot{x} - \dot{y} = -3x - y + 4x + 2y = x + y \\
&= -V_3, \quad (x,y) \in Q_3, \\
\dot{V}_4(x,y) &= -\dot{x} + \dot{y} = 4x - 3y - 2x + 2y = 2x - y \\
&< -V_4, \quad (x,y) \in Q_4.
\end{aligned} \tag{40}$$

The derivatives of Lyapunov functions V_1, V_2, V_3, and V_4 are negative definite with respect to norm $\|X\|_1$ in the corresponding regions Q_i.

On switching lines $L_{1,2}, \ldots, L_{4,1}$, where $x = 0$ or $y = 0$, the corresponding Lyapunov functions are

$$\begin{aligned}
&L_{1,2}: V_1(x,y) = V_2(x,y) = x, \\
&\qquad N_{1,2} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \\
&\qquad \langle N_{1,2}, f_1 \rangle = 4x > 0, \\
&\qquad \langle N_{1,2}, f_2 \rangle = 2x > 0, \\
&\qquad\qquad y = 0, \\
&L_{2,3}: V_2(x,y) = V_3(x,y) = -y, \\
&\qquad N_{2,3} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \\
&\qquad \langle N_{2,3}, f_2 \rangle = -5y > 0, \\
&\qquad \langle N_{2,3}, f_3 \rangle = -y > 0, \\
&\qquad\qquad x = 0, \\
&L_{3,4}: V_3(x,y) = V_4(x,y) = -x, \\
&\qquad N_{3,4} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \\
&\qquad \langle N_{3,4}, f_3 \rangle = -4x > 0, \\
&\qquad \langle N_{3,4}, f_4 \rangle = -2x > 0, \\
&\qquad\qquad y = 0, \\
&L_{4,1}: V_4(x,y) = V_1(x,y) = y, \\
&\qquad N_{4,1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \\
&\qquad \langle N_{4,1}, f_1 \rangle = 3y > 0, \\
&\qquad \langle N_{4,1}, f_1 \rangle = 2y > 0, \\
&\qquad\qquad x = 0.
\end{aligned} \tag{41}$$

Furthermore, all conditions of Theorem 12 are fulfilled; therefore, the trivial solution of the switched system is asymptotically stable (Figure 4). Moreover, as $\dot{V}_i \le -V_i$ the trivial solution is exponentially asymptotically stable.

Example 16. Consider now the switched systems formed by the following nonlinear subsystems:

$$\begin{aligned}
&E_1: \dot{x} = 3x + 2y - xy, \\
&\qquad \dot{y} = -4x - 3y - x^2 y^2, \\
&\qquad\qquad (x,y) \in Q_1, \\
&E_2: \dot{x} = -3x + 5y + xy, \\
&\qquad \dot{y} = -2x + 4y + x^2 y, \\
&\qquad\qquad (x,y) \in Q_2, \\
&E_3: \dot{x} = 3x + y, \\
&\qquad \dot{y} = -4x - 2y + xy, \\
&\qquad\qquad (x,y) \in Q_3, \\
&E_4: \dot{x} = -4x + 3y - xy, \\
&\qquad \dot{y} = -2x + 2y + x^3 y, \\
&\qquad\qquad (x,y) \in Q_4,
\end{aligned} \tag{42}$$

with Q_1, Q_2, Q_3, and Q_4 as in Example 15. As before, functions (39) fulfill conditions of Theorem 12 if the norm in $\mathbb{R}^2$ is equal to $\|X\|_1 = |x| + |y|$. The derivatives of functions (39) are

$$\begin{aligned}
\dot{V}_1(x,y) &= \dot{x} + \dot{y} = -x - y - xy - x^2 y \le -V_1(x,y) \\
&< 0, \quad (x,y) \in Q_1, \\
\dot{V}_2(x,y) &= \dot{x} - \dot{y} = -x + y + xy - x^2 y \le -V_2(x,y) \\
&< 0, \quad (x,y) \in Q_2, \\
\dot{V}_3(x,y) &= -\dot{x} - \dot{y} = x + y - xy \le -V_3(x,y) < 0, \\
&\qquad (x,y) \in Q_3, \\
\dot{V}_4(x,y) &= -\dot{x} + \dot{y} = 2x - y + xy + x^3 y \le -V_4(x,y) \\
&< 0, \quad (x,y) \in Q_4.
\end{aligned} \tag{43}$$

The derivatives of Lyapunov functions V_1, V_2, V_3, and V_4 are negative definite with respect to norm $\|X\|_1$ in the corresponding regions Q_i, and on switching lines $L_{1,2}, \ldots, L_{4,1}$, where $x = 0$ or $y = 0$, the corresponding Lyapunov functions coincide with (41). Therefore, the trivial solution of the switched system (42) is exponentially asymptotically stable. The simulation results are shown in Figure 5.

Theorem 17. *Suppose that sequence (4) is infinite and that for each equation E_i, $i = 1, \ldots, p$, there exists a Lyapunov function $V_i(X)$ such that*

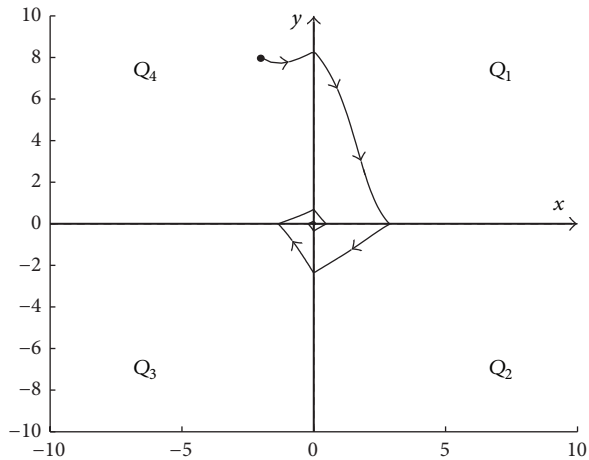


FIGURE 5: Trajectory of switched system (42).

(a) $\omega_{1,i}(\|X\|) < V_i(X) < \omega_{2,i}(\|X\|)$, $X \in Q_i$, $i = 1, \ldots, p$,

(b) $\dot{V}_i(X(t)) = \langle \nabla(V_i), f_i(X) \rangle \le 0$, $X \in Q_i$, $i = 1, \ldots, p$,

(c) *on all switched lines $L_{i,i+1}$, $1 < i < p$, where the trajectories pass from region Q_i to Q_{i+1}, the following inequalities hold:*

$$V_i(X) \ge V_{i+1}(X), \quad X \in L_{i,i+1}, \ 1 < i < p. \tag{44}$$

Furthermore, there exists at least one line $L_{k,k+1}$, $1 < k < p$ such that

$$V_k(X) \ge V_{k+1}(X) + \omega_{3,k}(\|X\|), \quad X \in L_{k,k+1}. \tag{45}$$

Then, the trivial solution $X(t) \equiv 0$ of the switched system (1)-(2) or (4) is asymptotically stable for a given switched system.

Proof. The trivial solution is stable because conditions of Theorem 12 are satisfied. It means that any Lyapunov function $V_k(X(t))$ in region Q_k as function of t does not increase. Moreover, the sequence of successive Lyapunov functions as function of t does not increase either. More exactly, for any numbers j, k and any instants t_j, t_k such that $X(t_j) \in Q_j$, $X(t_k) \in Q_k$, $t_j < t_k$, the following inequality holds: $V_k(X(t_k)) \le V_j(X(t_j))$.

Suppose the trivial solution $X(t) \equiv 0$ of switched system (4) is not asymptotically stable. It means that there exists the solution $X(t, t_0, X_0)$ with sufficiently small X_0, $\|X_0\| < \delta$, such that it is bounded $\|X(t, t_0, X_0)\| < H$ and does not tend to zero as $t \to \infty$. In this case there exists $\gamma > 0$ and a sequence $t_j \to \infty$ such that $\|X(t_j, t_0, X_0)\| > \gamma$, $\gamma < H$. The solution $X(t, t_0, X_0)$ must satisfy condition

$$\|X(t, t_0, X_0)\| \ge \mu, \quad t \ge t_1, \tag{46}$$

where μ is a number corresponding to γ in definition of stability for solution $X(t) \equiv 0$; that is, $\omega_2(\mu) \le \omega_1(\gamma)$. If (46) does not hold then there exists an instant $T > t_1$ such that

$$\|X(T, t_0, X_0)\| < \mu. \tag{47}$$

Taking now T as a new initial moment and using stability of $X(t, t_0, X_0) = X(t, T, X(T, t_0, X_0))$ this solution must satisfy condition

$$\|X(t, t_0, X_0)\| < \gamma, \quad t > T. \tag{48}$$

Condition (48) contradicts condition (46) and this means that condition (46) holds. It follows from (46) that for an infinite number of moments $t_{i(k)} > T$, $i(k) \to \infty$, when solution $X(t, t_0, X_0)$ crosses line $L_{k,k+1}$ we have $H > \|X(t_{i(k)}, t_0, X_0)\| \ge \mu > 0$, $\omega_{3,k}(\|X(t_{i(k)}, t_0, X_0)\|) \ge \omega_{3,k}(\mu) > 0$, and

$$\begin{aligned}
\omega_2(H) &\ge V_k(X(t_{i(k)}, t_0, X_0)) \\
&\ge V_{k+1}(X(t_{i(k)}, t_0, X_0)) \\
&\quad + \omega_{3,k}(\|X(t_{i(k)}, t_0, X_0)\|) \\
&\ge V_k(X(t_{i(k)-1}, t_0, X_0)) \\
&\quad + \omega_{3,k}(\|X(t_{i(k)-1}, t_0, X_0)\|) \\
&\quad + \omega_{3,k}(\|X(t_{i(k)}, t_0, X_0)\|) \\
&\ge V_{k+1}(X(t_{i(k)-1}, t_0, X_0)) \\
&\quad + \omega_{3,k}(\|X(t_{i(k)-1}, t_0, X_0)\|) \\
&\quad + \omega_{3,k}(\|X(t_{i(k)}, t_0, X_0)\|) \ge \cdots \\
&\ge V_{k+1}(X(t_1, t_0, X_0)) \\
&\quad + \omega_{3,k}(\|X(t_1, t_0, X_0)\|) + \cdots \\
&\quad + \omega_{3,k}(\|X(t_{i(k)-1}, t_0, X_0)\|) \\
&\quad + \omega_{3,k}(\|X(t_{i(k)}, t_0, X_0)\|) \ge i(k)\,\omega_{3,k}(\mu) \\
&\longrightarrow \infty, \quad i(k) \longrightarrow \infty.
\end{aligned} \tag{49}$$

Inequality (49) contradicts condition $\omega_2(H) < \infty$ and our supposition that $X(t, t_0, X_0)$ does not tend to zero conducts to a contradiction. Furthermore, the asymptotic stability of trivial solution $X(t) \equiv 0$ of switched system (4) is proven. □

Theorem 18. *Suppose that sequence (4) is infinite and that for all switching lines $L_{i,i+1}$ and that for each E_i, $i = 1, \ldots, p$, there exists a Lyapunov function $V_i(X)$ such that*

(a) $\omega_{1,i}(\|X\|) < V_i(X) < \omega_{2,i}(\|X\|)$, $X \in Q_i$, $i = 1, \ldots, p$,

(b) $\dot{V}_i(X(t)) = \langle \nabla(V_i), f_i(X) \rangle \ge 0$, $X \in Q_i$, $i = 1, \ldots, p$,

(c) *on all switched lines $L_{i,i+1}$, $1 \le i < p$, where the trajectories pass from region Q_i to Q_{i+1}, the following inequalities hold:*

$$V_i(X) \le V_{i+1}(X), \quad X \in L_{i,i+1}, \ 1 \le i < p, \tag{50}$$

furthermore, there exists at least one line $L_{k,k+1}$, $1 \le k < p$, such that

$$V_k(X) \le V_{k+1}(X) + \omega_{3,k}(\|X\|), \quad X \in L_{k,k+1}. \tag{51}$$

Then, the trivial solution $X(t) \equiv 0$ of the switched system (4) is unstable for a given switched system.

Proof. Theorem 18 is a clear modification of Theorem 17. Therefore its proof is omitted. □

Example 19. Consider the switched system formed by two nonlinear pendulums:

$$E_1\colon \ddot{x} + \sin x = 0, \quad (x, \dot{x} = y) \in Q_1 = \{x < 0, y > 0\}, \tag{52}$$
$$E_2\colon \ddot{x} + \omega^2 \sin x = 0, \quad (x, \dot{x} = y) \in \mathbb{R}^2 - Q_1.$$

Take as Lyapunov functions for (52) the following functions:

$$V_1(x, y) = (1 - \cos x) + \frac{1}{2}y^2, \quad (x, \dot{x} = y) \in Q_1 = \{x < 0, y > 0\},$$
$$V_2(x, y) = \omega^2 (1 - \cos x) + \frac{1}{2}y^2, \quad (x, \dot{x} = y) \in \mathbb{R}^2 - Q_1. \tag{53}$$

The derivatives of functions (53) along the trajectories of (52) are equal to zero: $\dot{V}_1 = \dot{V}_2 = 0$. On switching line $L_{1,2}$ where $x = 0$, $y > 0$ we have $V_1(x, y) = (1/2)y^2 = V_2(x, y)$. On switching line $L_{2,1}$ where $x < 0$ and $y = 0$ we have $V_1(x, y) = (1 - \cos x)$ and $V_2(x, y) = \omega^2(1 - \cos x)$. Therefore, condition (c) of Theorem 18 holds on both lines $L_{1,2}$ and $L_{2,1}$ if

$$\omega^2 < 1. \tag{54}$$

Condition (54) is a condition of instability of the trivial solution of switched system (52). Also, this condition may be considered as a condition of parametric discontinuous excitations for the nonlinear pendulum.

5. Autonomous Multidimensional Switched Systems

Consider multidimensional switched system acting in $\mathbb{R}^n$ which is divided into a finite number p of open 1-connected regions Q_k, $1 \le k \le p$, $Q_1 + Q_2 + \cdots + Q_p = \mathbb{R}^n$, such that the origin $\{0\}$ of rectangular coordinates systems pertains to closure of any region Q_k. The boundaries of all regions Q_k are supposed to be smooth. Suppose, also, that each region Q_l has points X fulfilling condition $\|X\| > H$, where H is any number. An equation E_k is defined in each region Q_k satisfying conditions of existence and uniqueness of solutions. In multidimensional case, the switched system may have much more complicated behavior because region Q_k may have more than two boundaries separating it from other regions. To eliminate topological complications, consider only the case when all boundaries (switching surfaces) $B_{k,k+l}$ separating regions Q_k and Q_{k+l} are planes in $\mathbb{R}^n$. Suppose the trajectories of switched system pass through the boundary $B_{k,k+l}$ in all points only in direction from Q_k to Q_{k+l} and on all lines $L_{k,k+l}$ where Q_k touches other regions all trajectories pass from Q_k only to one determined region Q_{k+j}. Suppose also the existence of some Lyapunov functions $V_k(X)$ for all equation E_k, $1 \le k \le p$. In this case theorems similar to Theorems 10, 12, 17, and 18 may be established. The definitions of stability for switched systems in $\mathbb{R}^n$ are the same as for the two-dimensional case. Now, consider the case of a switched system with infinite number of switchings. Let us formulate one of the theorems on stability.

Theorem 20. *Suppose that for each equation E_k acting in $Q_k \in \mathbb{R}^n$ there exists a Lyapunov function $V_k(X)$ defined in Q_k such that*

(a) $\omega_{1,k}(\|X\|) < V_k(X) < \omega_{2,k}(\|X\|)$, $X \in Q_k$, $k = 1, \ldots, p$,

(b) $\dot{V}_k(X(t)) = \langle \nabla(V_k), f_k(X) \rangle < 0$, $X \in Q_k$, $k = 1, \ldots, p$,

(c) *on all switched surfaces $B_{k,k+j}$ and lines $L_{k,k+j}$ where the trajectories pass from region Q_k to Q_{k+j}, the following inequalities hold:*

$$V_k(X) > V_{k+j}(X), \quad X \in B_{k,k+j},$$
$$V_k(X) > V_{k+l}(X), \quad X \in L_{k,k+l}. \tag{55}$$

Then, the trivial solution $X(t) \equiv 0$ of the switched system acting in $\mathbb{R}^n$ is stable (asymptotically stable) for a given switched structure.

Proof. The demonstration of this theorem is similar to proof of Theorem 12. Therefore it is omitted. □

Example 21. In space $\mathbb{R}^3$ consider two regions $Q_1 = \{z > 0\}$ and $Q_2 = \{z < 0\}$. Suppose equations E_1 and E_2 are defined in Q_1 and Q_2:

$$E_1\colon \dot{x} = -x + y + y^2, \quad \dot{y} = -x - y, \quad \dot{z} = -2z - zx^2 - 2y^2x, \quad (x, y, z) \in Q_1,$$
$$E_2\colon \dot{x} = -x + y - y^2, \quad \dot{y} = -x - y - xy, \quad \dot{z} = -2z - 4y^2x, \quad (x, y, z) \in Q_2. \tag{56}$$

Normals $N_{1,2}$ and $N_{2,1}$ to the surface $B = z = 0$ separating Q_1 and Q_2 are $N_{1,2}^T = [0, 0, -1]^T$ and $N_{2,1}^T = [0, 0, 1]^T$. Therefore, condition (3) holds on surface $B = z = 0$:

$$\begin{aligned} &\langle N_{1,2}, f_{E_1} \rangle = 2y^2 x > 0, \\ &\langle N_{1,2}, f_{E_2} \rangle = 4y^2 x > 0, \\ &\qquad\qquad\qquad\qquad x > 0, \\ &\langle N_{2,1}, f_{E_1} \rangle = -2y^2 x > 0, \\ &\langle N_{2,1}, f_{E_2} \rangle = -4y^2 x > 0, \\ &\qquad\qquad\qquad\qquad x < 0. \end{aligned} \tag{57}$$

Therefore, surfaces $B_{1,2}$ and $B_{2,1}$ where the trajectories pass from E_1 to E_2 or vice versa are $B_{1,2} = \{z = 0, x > 0, y = \text{arbitrary}\}$ and $B_{2,1} = \{z = 0, x < 0, y = \text{arbitrary}\}$. On line $\{z = 0, x = 0, y = \text{arbitrary}\}$ the derivatives $\dot{x}$ from one and another side of this line are positive: $\dot{x} = y^2 > 0$. Therefore, the trajectory which arrives on this line passes throughout it and then enters into the region Q_2. Consider two Lyapunov functions V_1 and V_2 defined in Q_1 and Q_2, respectively:

$$\begin{aligned} V_1(x, y, z) &= x^2 + y^2 + z, \quad (x, y, z) \in Q_1, \\ V_2(x, y, z) &= x^2 + y^2 - z, \quad (x, y, z) \in Q_2. \end{aligned} \tag{58}$$

Derivatives $\dot{V}_{1,E_1}$ and $\dot{V}_{2,E_2}$ computed on the trajectories of equations E_1 and E_2, respectively, are

$$\begin{aligned} \dot{V}_{1,E_1} &= 2x\dot{x} + 2y\dot{y} + \dot{z} \\ &= 2x\left(-x + y + y^2\right) + 2y(-x - y) \\ &\quad + \left(-2z - zx^2 - 2y^2 x\right) \\ &= -2x^2 - 2y^2 - 2z - zx^2 < -2V_1, \\ &\qquad\qquad\qquad\qquad (x, y, z) \in Q_1, \\ \dot{V}_{2,E_2} &= 2x\dot{x} + 2y\dot{y} - \dot{z} \\ &= 2x\left(-x + y - y^2\right) + 2y(-x - y - xy) \\ &\quad + \left(-2z - 4y^2 x\right) = -2x^2 - 2y^2 + 2z = -2V_2, \\ &\qquad\qquad\qquad\qquad (x, y, z) \in Q_2. \end{aligned} \tag{59}$$

All conditions of Theorem 20 are fulfilled and therefore the considered switching system (56) is asymptotically stable and, moreover, exponentially asymptotically stable (see Figure 6).

6. Conclusions

The stability of autonomous switched systems is investigated. Some theorems giving sufficient conditions of stability, asymptotic stability, or instability of switched systems are established. Some results are similar to conditions of parametric stabilization or excitation of stable systems.

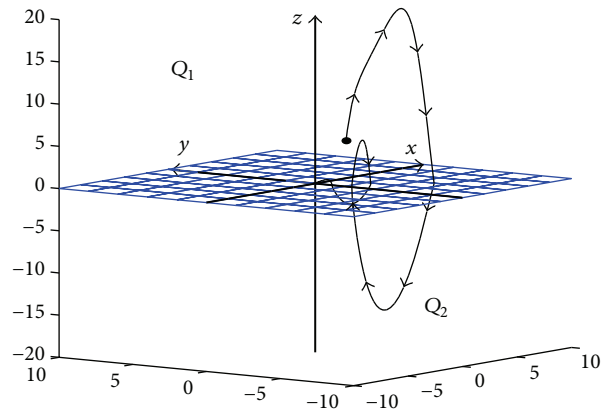


Figure 6: Trajectory of switched system (56).

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was partially supported by CONACYT (Consejo Nacional de Ciencia y Tecnologa) through scholarship SNI (Sistema Nacional de Investigadores) and by IPN (Instituto Politécnico Nacional) through research Projects and scholarships EDI (Estímulo al Desempeño de los Investigadores), COFAA (Comisión de Operación y Fomento de Actividades Académicas), and BEIFI (Beca de Estímulo Institucional de Formación de Investigadores).

References

[1] P. J. Antsalis, W. Kohn, A. Nerode, and S. Sastre, Eds., *Hybrid Systems IV*, vol. 1273 of *Lecture Notes in Computer Science*, Springer, New York, NY, USA, 1997.

[2] D. Cheng, L. Guo, Y. Lin, and Y. Wang, "Stabilization of switched linear systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 5, pp. 661–666, 2005.

[3] R. L. Grossman, A. Nerode, A. P. Ravin, and H. Richel, Eds., *Hybrid Systems*, vol. 736 of *Lecture Notes in Computer Science*, Springer, New York, NY, USA, 1993.

[4] R. Guo and Y. Wang, "Region stability analysis for switched linear systems with multiple equilibria," in *Proceedings of the 29th Chinese Control Conference (CCC '10)*, pp. 986–991, IEEE, Beijing, China, July 2010.

[5] F. Z. Taousser, M. Defoort, and M. Djemai, "Stability analysis of a class of uncertain switched systems on time scale using lyapunov functions," *Nonlinear Analysis: Hybrid Systems*, vol. 16, pp. 13–23, 2015.

[6] D. D. Bainov and P. S. Simeonov, *Systems with Impulse Effect*, Ellis Horwoord, Chichester, UK, 1989.

[7] Y. Or and A. D. Ames, "Stability and completion of zeno equilibria in lagrangian hybrid systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 6, pp. 1322–1336, 2011.

[8] P. Tinglong, Y. Kun, S. Yanxia, G. Zairui, and J. Zhicheng, "Finite-time stability analysis for a class of continuous switched

descriptor systems," *Mathematical Problems in Engineering*, vol. 2014, Article ID 979130, 9 pages, 2014.

[9] V. I. Utkin, J. Guldner, and J. Shi, *Sliding Mode Control in Electromechanical Systems*, Taylor & Francis, Philadelphia, Pa, USA, 1999.

[10] F. Zhu and P. J. Antsaklis, "Optimal control of hybrid switched systems: a brief survey," *Discrete Event Dynamic Systems: Theory and Applications*, vol. 25, no. 3, pp. 345–364, 2015.

[11] P. Peleties and R. DeCarlo, "Asymptotic stability of m-switched systems using lyapunov-like functions," in *Proceedings of the American Control Conference*, pp. 1679–1684, Boston, Mass, USA, June 1991.

[12] M. A. Wicks, P. Peleties, and R. A. DeCarlo, "Construction of piecewise lyapunov functions for stabilizing switched systems," in *Proceedings of the 33rd IEEE Conference on Decision and Control*, pp. 3492–3497, Lake Buena Vista, Fla, USA, December 1994.

[13] M. S. Branicky, "Multiple lyapunov functions and other analysis tools for switched and hybrid systems," *IEEE Transactions on Automatic Control*, vol. 43, no. 4, pp. 475–482, 1998.

[14] V. R. Nosov, H. Dominguez, and J. A. Ortega, "Estabilidad de un péndulo con conmutaciones," in *Memorias del 3er Congreso Internacional en Matemáticas Aplicadas*, pp. 245–253, Mexico City, Mexico, 2007.

[15] V. R. Nosov, H. Dominguez, J. A. Ortega-Herrera, and J. A. Meda-Campaña, "Complex dynamics and chaos in commutable pendulum," *Revista Mexicana de Fisica*, vol. 58, no. 1, pp. 6–12, 2012.

[16] V. N. Afanas'ev, V. B. Kolmanovskii, and V. R. Nosov, *Mathematical Theory of Control Systems Design*, Kluwer Academic Publishers Group, Dorderecht, The Netherlands, 1996.

[17] S. Lefschetz, *Stability of Nonlinear Control Systems*, Academic Press, New York, NY, USA, 1965.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 181737, 8 pages
http://dx.doi.org/10.1155/2015/181737

Research Article

Nonsingular Terminal Sliding Mode Control of Uncertain Second-Order Nonlinear Systems

Minh-Duc Tran[1] and Hee-Jun Kang[2]

[1] *University of Ulsan, Ulsan 680-749, Republic of Korea*
[2] *School of Electrical Engineering, University of Ulsan, Ulsan 680-749, Republic of Korea*

Correspondence should be addressed to Hee-Jun Kang; hjkang@ulsan.ac.kr

Received 27 April 2015; Revised 10 July 2015; Accepted 13 July 2015

Academic Editor: Rongwei Guo

This paper presents a high-performance nonsingular terminal sliding mode control method for uncertain second-order nonlinear systems. First, a nonsingular terminal sliding mode surface is introduced to eliminate the singularity problem that exists in conventional terminal sliding mode control. By using this method, the system not only can guarantee that the tracking errors reach the reference value in a finite time with high-precision tracking performance but also can overcome the complex-value and the restrictions of the exponent (the exponent should be fractional number with an odd numerator and an odd denominator) in traditional terminal sliding mode. Then, in order to eliminate the chattering phenomenon, a super-twisting higher-order nonsingular terminal sliding mode control method is proposed. The stability of the closed-loop system is established using the Lyapunov theory. Finally, simulation results are presented to illustrate the effectiveness of the proposed method.

1. Introduction

As the development of control schemes has progressed, a variety of control systems have been developed for robotic manipulators, including proportional-integral-derivative (PID) control [1], adaptive control [2], computed torque control [3, 4], fuzzy control [5], and neural network control [6]. Sliding mode control (SMC) is an efficient control method that has been widely applied to control for both linear and nonlinear systems. In order to design sliding mode control systems, establishment of suitable sliding surfaces to ensure the desired dynamics is considered first, and then a sliding mode controller is designed to drive the system states to the sliding surface. The main characteristic of SMC is to use discontinuous control effort to keep the system states on the sliding surfaces, whereby SMC has strong robustness with respect to system uncertainties and external disturbances, fast response, and good transient performance. However, the conventional SMC method cannot guarantee the invariance properties during the reaching phase and even against disturbances can degrade the performance of system [7–9]. Moreover, this method adopts a linear sliding surface, which can only provide asymptotic stability of the system in the sliding phase.

Terminal sliding mode control (TSMC) methods, which use nonlinear sliding surfaces instead of a linear surface, were first introduced by Venkataraman and Gulati [10] and further developed by Man et al. [11, 12] and Wu et al. [13]. Compared with linear SMC, TSMC schemes not only ensure that the system states arrive at the equilibrium point in a finite time but also offer some attractive properties, such as their fast response and higher precision. However, the traditional TSMC methods may have slower convergence performance when the system states are not near the equilibrium point, and they also suffer from the singularity problem and have restrictions on the range of the power function. In order to avoid these drawbacks, some new TSMC methods have been proposed [14–16]. Yu and Zhihong [14] have developed fast terminal sliding mode (FTSM), which can improve the convergence speed when the system states are far from the equilibrium point. This method, however, still has the singularity problem. To overcome this, Feng et al. [16] introduced nonsingular terminal sliding mode (NTSM) control.

However, this surface has a limitation on the power function; that is, p and q must be positive odd integers.

Discontinuous terminal sliding mode control (TSMC) has been widely applied to nonlinear systems. Nevertheless, the main drawback of discontinuous TSMC is the chattering phenomenon, which comes from high frequency switching of the control signal. It shows undesirable oscillation on the system, leads to low control accuracy, causes high wear of the moving mechanical parts, and may damage the actuators. To deal with this problem, the most common methods replace the sign function in the switching control with a saturating approximation [17] or boundary layer technique [18]. The boundary layer method was proposed to eliminate the chattering by defining a boundary layer around the sliding surface and then approximate the discontinuous control by continuous function within this boundary layer. As a result, the chattering elimination is achieved; however, there is a trade-off between chattering elimination and tracking performance; a thicker boundary layer can eliminate the chattering phenomenon but the tracking error will be increased. Recently, intelligent control schemes (neural network and fuzzy logic) have been applied to attenuate the chattering phenomenon [19–21]. However, some controller designs based on intelligence techniques were quite complicated and fell into difficulties in stability analysis. Therefore, in this study, high-order sliding mode (HOSM) techniques have been studied and applied. The main characteristic of HOSM is that they are working with the discontinuous control in the higher-order time derivative [22–27], so the chattering can be reduced because the control signal is continuous. Furthermore, HOSM can bring better accuracy than conventional SMC while the robustness of the control system is similar to SMC. It has been presented in [23–25] for the control of rigid robot manipulators.

In this paper, the above-mentioned problems are addressed based on a proposed NTSM surface for second-order nonlinear systems. A control law is designed to drive the system states to reach the sliding surface and converge to zero in a finite time. It does not suffer from the singularity problem or the restriction on the power function. Furthermore, a super-twisting second-order sliding mode is also used to reduce the chattering of the controller. The global finite time stability of the closed-loop system is proven. The convergence times of the reaching phase and sliding phase are also given. The simulation results are presented to illustrate the effectiveness of the proposed method on the two-link robot manipulator.

The remainder of this paper is arranged as follows. Preliminaries and problem formulation are given in Section 2. In Section 3, the structure of super-twisting nonsingular terminal sliding mode controller is presented and a stability analysis is performed. In Section 4, simulation results for a two-link robot manipulator are provided to demonstrate the performance of the proposed controller. Finally, some concluding remarks are presented in Section 5.

2. Preliminaries and Problem Formulation

Consider the following nonlinear second-order mechanical systems:

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= f(x,t) + d(x,t) + b(x,t)\,u(t), \end{aligned} \tag{1}$$

where $x = [x_1, x_2]^T$ denotes the system state vector, $f(x,t)$ and $b(x,t)$ are smooth nonlinear functions of x, $u(t)$ is the control input, and $d(x,t)$ presents the uncertainties and disturbances.

Assumption 1. The matrices $b(x,t)$ are invertible $\forall x$.

Assumption 2. The uncertain term is bounded by

$$|d(x,t)| \le \overline{D}, \tag{2}$$

where $\overline{D}$ is a known positive constant.

Assumption 3. The desired state vector $x_d(t) \in R$ is a twice continuously differentiable function in terms of t.

The control objective of this paper is to design a controller for system (1) to ensure that the error between the real state vector x and the desired state vector $x_d(t)$ converges to zero in finite time.

3. Main Results

In this section, the design of super-twisting nonsingular terminal sliding mode controller is presented. First, a new nonsingular terminal sliding mode surface is proposed to eliminate the singularity problem. Then, the conventional SMC and super-twisting nonsingular terminal sliding mode controller are designed to ensure that the tracking error converges to zero in a finite amount time.

3.1. New Form of NTSM Surface. We define the tracking error as $\varepsilon(t) = x_1(t) - x_{1d}(t)$. Thus, a new NTSM surface is proposed as follows:

$$s = \dot{\varepsilon} + \beta_1 \varepsilon + \beta_2 e^{-\lambda t} \left(\varepsilon^T \varepsilon\right)^{-\alpha} \varepsilon, \tag{3}$$

where $s = [s_1, s_2, \ldots, s_n]^T$, $\varepsilon = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n]^T$, $\dot{\varepsilon} = [\dot{\varepsilon}_1, \dot{\varepsilon}_2, \ldots, \dot{\varepsilon}_n]^T$, $\beta_1 = \mathrm{diag}(\beta_{11}, \beta_{12}, \ldots, \beta_{1n})$, $\beta_2 = \mathrm{diag}(\beta_{21}, \beta_{22}, \ldots, \beta_{2n})$ with $\beta_{1i}, \beta_{2i} > 0$ for every $i = 1, 2, \ldots, n$, $0 < \alpha < 1$, and $\lambda > 0$.

When the system operates in sliding mode, the following is true:

$$s = \dot{\varepsilon} + \beta_1 \varepsilon + \beta_2 e^{-\lambda t} \left(\varepsilon^T \varepsilon\right)^{-\alpha} \varepsilon = 0 \tag{4}$$

$$\dot{\varepsilon} = -\beta_1 \varepsilon - \beta_2 e^{-\lambda t} \left(\varepsilon^T \varepsilon\right)^{-\alpha} \varepsilon. \tag{5}$$

Theorem 4. *Considering the sliding mode dynamic equation (5), the system is finite time stable at the equilibrium point $\varepsilon = 0$, and the tracking error ε will converge to zero in finite time if $2\alpha\lambda_{\min}(\beta_1) - \lambda > 0$.*

The finite convergence time is

$$T_s \le \frac{\ln\left(1+\left(e^{2\alpha\lambda_{\min}(\beta_1)t}\cdot V^{\alpha}(0)\right)/a_2\right)}{2\alpha\lambda_{\min}(\beta_1)-\lambda}, \tag{6}$$

where a_2 is expressed by (15).

Proof. Consider the Lyapunov function:

$$V = \frac{1}{2}\varepsilon^T\varepsilon. \tag{7}$$

Taking the derivative of V in (7) and substituting (5) into it yield

$$\dot{V} = \varepsilon^T\dot{\varepsilon} = \varepsilon^T\left[-\beta_1\varepsilon-\beta_2 e^{-\lambda t}\left(\varepsilon^T\varepsilon\right)^{-\alpha}\varepsilon\right], \tag{8}$$

$$\begin{aligned}\dot{V} &= -\varepsilon^T\beta_1\varepsilon-\beta_2 e^{-\lambda t}\left(\varepsilon^T\varepsilon\right)^{-\alpha}\varepsilon^T\varepsilon\\ &\le -\lambda_{\min}(\beta_1)\varepsilon^T\varepsilon-\lambda_{\min}(\beta_2)e^{-\lambda t}\left(\varepsilon^T\varepsilon\right)^{1-\alpha}\\ &\le -2\lambda_{\min}(\beta_1)V-2^{1-\alpha}\lambda_{\min}(\beta_2)e^{-\lambda t}V^{1-\alpha}\le 0.\end{aligned} \tag{9}$$

Therefore, according to the Lyapunov stability, it is obvious that the origin is at globally stable equilibrium. Next, we will show that the system states converge to zero in finite time.

Multiplying both sides of (9) by $\alpha V^{\alpha-1}$, we have

$$\begin{aligned}\alpha V^{\alpha-1}\frac{dV}{dt} &\le -2\alpha\lambda_{\min}(\beta_1)V^{\alpha}-2^{1-\alpha}\alpha\lambda_{\min}(\beta_2)e^{-\lambda t},\\ \frac{dV^{\alpha}}{dt}+2\alpha\lambda_{\min}(\beta_1)V^{\alpha} &\le -2^{1-\alpha}\alpha\lambda_{\min}(\beta_2)e^{-\lambda t}.\end{aligned} \tag{10}$$

Multiplying both sides of (10) by $e^{2\alpha\lambda_{\min}(\beta_1)t}$ yields

$$\begin{aligned}&e^{2\alpha\lambda_{\min}(\beta_1)t}\left(\frac{dV^{\alpha}}{dt}+2\alpha\lambda_{\min}(\beta_1)V^{\alpha}\right)\\ &\quad\le -2^{1-\alpha}\alpha\lambda_{\min}(\beta_2)e^{[2\alpha\lambda_{\min}(\beta_1)-\lambda]t},\end{aligned} \tag{11}$$

$$\frac{d\left(e^{2\alpha\lambda_{\min}(\beta_1)t}\cdot V^{\alpha}\right)}{dt} \le -2^{1-\alpha}\alpha\lambda_{\min}(\beta_2)e^{[2\alpha\lambda_{\min}(\beta_1)-\lambda]t}. \tag{12}$$

Taking the integral on both sides of (12) from 0 to T_s and knowing $V(T_s)=0$ yield

$$-e^{2\alpha\lambda_{\min}(\beta_1)t}\cdot V^{\alpha}(0) \le -a_2\left[e^{[2\alpha\lambda_{\min}(\beta_1)-\lambda]T_s}-1\right], \tag{13}$$

$$e^{[2\alpha\lambda_{\min}(\beta_1)-\lambda]T_s} \le 1+\frac{e^{2\alpha\lambda_{\min}(\beta_1)t}\cdot V^{\alpha}(0)}{a_2}, \tag{14}$$

where

$$a_2 = \frac{2^{1-\alpha}\alpha\lambda_{\min}(\beta_2)}{2\alpha\lambda_{\min}(\beta_1)-\lambda} > 0. \tag{15}$$

Taking the natural logarithm of both sides of (14) yields

$$\begin{aligned}&\left[2\alpha\lambda_{\min}(\beta_1)-\lambda\right]T_s \le \ln\left(1+\frac{e^{2\alpha\lambda_{\min}(\beta_1)t}\cdot V^{\alpha}(0)}{a_2}\right).\\ &T_s \le \frac{\ln\left(1+\left(e^{2\alpha\lambda_{\min}(\beta_1)t}\cdot V^{\alpha}(0)\right)/a_2\right)}{2\alpha\lambda_{\min}(\beta_1)-\lambda}\end{aligned} \tag{16}$$

This completes the proof. □

Remark 5. The expression in (3) is different from the previously reported TSM and fast TSM in [14], which are expressed, respectively, as

$$\begin{aligned}s &= \dot{x}+\beta x^{q/p},\\ s &= \dot{x}+\alpha x+\beta x^{q/p},\end{aligned} \tag{17}$$

where α and β are positive constants and p and q are positive odd integers that satisfy the following condition: $1 < p/q < 2$. We can easily see that, for $x < 0$, the fractional power q/p may lead to the term $x^{q/p} \notin R$, which means $\dot{x} \notin R$. In addition, the TSM control signals in [14] contain ${x_1}^{q/p-1}x_2$, which may cause a singularity to occur if $x_2 \ne 0$ when $x_1 = 0$.

To solve the complex-value problem in (17), Yu et al. [28] proposed the TSM surface as

$$\begin{aligned}s &= \dot{x}+\beta|x|^{\gamma}\operatorname{sign}(x),\\ s &= \dot{x}+\alpha x+\beta|x|^{\gamma}\operatorname{sign}(x).\end{aligned} \tag{18}$$

The sliding surface in (18) could solve the complex-value number, but the control input can suffer from the singularity problem if $x_2 \ne 0$ when $x_1 = 0$.

Recently, a nonsingular terminal sliding surface was proposed to overcome the singularity problem [16]:

$$s = x+\frac{1}{\beta}\dot{x}^{p/q}. \tag{19}$$

However, this surface still has the limitation for the exponent of the power function; that is, p and q should be positive odd integers. Thus, our proposed TSM surface does not contain any of the mentioned singularities, and the exponent can be any real number in the interval $0 < \alpha < 1$.

Remark 6. Comparing with linear sliding mode, NTSM has higher convergence rate when the system state is far away from the equilibrium point, while NTSM has lower convergence speed when the system state is close to the equilibrium point [29, 30].

It is obvious that the term $e^{-\lambda t}$ in the proposed surface will go backward to zero after a certain time. Thus, the nonsingular terminal sliding mode surface will become linear sliding mode after a period of time. By choosing a suitable λ, the proposed surface will have the advantage of both NTMS and linear sliding surface.

3.2. NTSM Control (NTSMC) Design. One suitable sliding manifold is established. The next step is to design the control to drive the nonlinear system (1) to the expected sliding surface (3) in a finite amount time. The proposed control method is summarized as follows.

Theorem 7. *For the system (1), if the control signal is designed as (20) and the gain η of the controller is larger than the upper*

bounds of the uncertainties, the tracking error $\varepsilon(t)$ will converge to zero in finite time:

$$u(t) = -b(x,t)^{-1} \cdot \left[f(x,t) - \ddot{x}_d + \beta_1\dot{\varepsilon} + \beta_2 A + \eta\,\mathrm{sign}(s)\right], \tag{20}$$

where $\eta = \mathrm{diag}(\eta_1, \eta_2, \ldots, \eta_n)$, $\eta_i > 0$. *Therefore,*

$$A = \left[(-\lambda)\, e^{-\lambda t}\left(\varepsilon^T\varepsilon\right)^{-\alpha}\varepsilon + e^{-\lambda t}(-2\alpha)\, x\left(\varepsilon^T\varepsilon\right)^{-\alpha-1}\left(\varepsilon^T\dot{\varepsilon}\right)\varepsilon + e^{-\lambda t}\left(\varepsilon^T\varepsilon\right)^{-\alpha}\dot{\varepsilon}\right]. \tag{21}$$

Proof. Consider the following Lyapunov candidate function:

$$V = \frac{1}{2}s^T s. \tag{22}$$

The time derivative of the sliding surface (3) with respect to time can be expressed as

$$\begin{aligned}\dot{s} &= \ddot{\varepsilon} + \beta_1\dot{\varepsilon} + \beta_2 A \\ &= f(x,t) + d(x,t) + b(x,t)\,u(t) - \ddot{x}_d(t) + \beta_1\dot{\varepsilon} \\ &\quad + \beta_2 A.\end{aligned} \tag{23}$$

Differentiating V with respect to time and substituting (20) and (23) into it yield

$$\dot{V} = s^T\left(-\eta\,\mathrm{sign}(s) + d(t)\right) \le -\left(\eta - \overline{D}\right)|s| \le 0. \tag{24}$$

Therefore, the condition for Lyapunov stability is satisfied; in the following, we will show that the error converges to zero in finite time.

From (24), we have

$$\begin{aligned}\dot{V} &\le -\sqrt{2}\left(\eta - \overline{D}\right)V^{1/2}, \\ dt &\le -\frac{dV}{\sqrt{2}\left(\eta - \overline{D}\right)V^{1/2}} = -\frac{\sqrt{2}dV^{1/2}}{\left(\eta - \overline{D}\right)}.\end{aligned} \tag{25}$$

Taking the integral of both sides of (25) from T_r to T_s, we have

$$T_s - T_r \le -\int_{V(T_r)}^{V(T_s)} \frac{\sqrt{2}dV^{1/2}}{\left(\eta - \overline{D}\right)} = \frac{\sqrt{2}}{\left(\eta - \overline{D}\right)}V^{1/2}(T_r). \tag{26}$$

Note that $V(T_s) = 0$; therefore, the TSM will reach zero in the finite time:

$$T_s \le \frac{\sqrt{2}}{\left(\eta - \overline{D}\right)}V^{1/2}(T_r) + T_r. \tag{27}$$

This completes the proof. □

Remark 8. In order to eliminate the chattering, a saturation function sat or $s/(\|s\| + \varepsilon)$ (ε is a small positive constant) can be used to replace the sign function.

3.3. Super-Twisting NTSM Control (ST-NTSMC) Design. The main drawback of the conventional sliding mode is the chattering phenomenon which is caused by discontinuous control action when the system state operates near the sliding surface. Even though the chattering reduction can be achieved by using Remark 8, there is a trade-off between chattering elimination and tracking performance; increasing the thickness of the boundary layer can eliminate the chattering phenomenon but will increase the tracking error. Therefore, in this subsection, super-twisting control is applied to attenuate chattering and to increase the tracking performance.

The ST-NSTSMC is designed as

$$u = u_{\mathrm{eq}} + u_{\mathrm{STW}}, \tag{28}$$

where

$$u_{\mathrm{eq}} = -b(x,t)^{-1}\left[f(x,t) - \ddot{x}_d + \beta_1\dot{\varepsilon} + \beta_2 A\right]. \tag{29}$$

Based on [27], the super-twisting controller is designed as

$$\begin{aligned}u_{\mathrm{STW}} &= -b(x,t)^{-1}\left(k_1|s|^{1/2}\,\mathrm{sign}(s) + z\right), \\ \dot{z} &= -k_2\,\mathrm{sign}(s).\end{aligned} \tag{30}$$

The differentiation of the sliding surface is now obtained as

$$\begin{aligned}\dot{s} &= f(x,t) + d(x,t) + b(x,t)\left(u_{\mathrm{eq}} + u_{\mathrm{SMW}}\right) - \ddot{x}_d(t) \\ &\quad + \beta_1\dot{\varepsilon} + \beta_2 A.\end{aligned} \tag{31}$$

Substituting (29) and (30) into (31) yields

$$\begin{aligned}\dot{s} &= -k_1|s|^{1/2}\,\mathrm{sign}(s) + z + d(x,t), \\ \dot{z} &= -k_2\,\mathrm{sign}(s).\end{aligned} \tag{32}$$

The stability and convergence of the closed-loop system in (32) are given in Theorem 9.

Theorem 9. *Suppose that Assumption 1 is guaranteed and the uncertain terms are bounded by*

$$\begin{aligned}d(x,t) &\le \delta|s|^{1/2}, \\ \delta &= \mathrm{diag}(\delta_1, \delta_2, \ldots, \delta_n), \quad \delta_i > 0.\end{aligned} \tag{33}$$

For system (1), with the terminal sliding mode surface chosen as in (3) and the proposed control signal designed as in (28), if the sliding gains of u_{STW} given in (30) satisfy condition (34), then the sliding surface s will converge to zero in a finite time:

$$\begin{aligned}k_1 &> 2\delta, \\ k_2 &> k_1\frac{5k_1 + 4\delta}{2(k_1 - 2\delta)}\delta.\end{aligned} \tag{34}$$

Proof. Now, referring to Moreno's work [27], let us consider the Lyapunov candidate function:

$$V = \xi^T P\xi, \tag{35}$$

where

$$\xi = \left[|s|^{1/2}\,\mathrm{sign}\,(s)\,, z\right]^T,$$
$$P = \frac{1}{2}\begin{bmatrix} {k_1}^2 + 4k_2 & -k_1 \\ -k_1 & 2 \end{bmatrix}. \tag{36}$$

As we know, V is positive definite and radially unbounded:

$$\lambda_{\min}(P)\,\|\zeta\|^2 \le V \le \lambda_{\max}(P)\,\|\zeta\|^2, \tag{37}$$

where $\|\zeta\|^2 = |s| + z^2$. The time derivative of V becomes

$$\dot{V} = -\frac{1}{|s|^{1/2}}\left(\xi^T Q_1 \xi - d(x,t)\,{Q_2}^T \xi\right), \tag{38}$$

where

$$Q_1 = \frac{k_1}{2}\begin{bmatrix} {k_1}^2 + 2k_2 & -k_1 \\ -k_1 & 1 \end{bmatrix},$$
$${Q_2}^T = \begin{bmatrix} \frac{{k_1}^2}{2} + 2k_2 & -\frac{k_1}{2} \end{bmatrix}. \tag{39}$$

Using condition (33), it can be shown that

$$\dot{V} \le -\frac{1}{|s|^{1/2}}\xi^T Q \xi \le -\frac{1}{|s|^{1/2}}\lambda_{\min}(Q)\,\|\xi\|^2, \tag{40}$$

where

$$Q = \frac{k_1}{2}\begin{bmatrix} {k_1}^2 + 2k_2 - \left(\frac{4k_2}{k_1} + k_1\right)\delta & -(k_1 + 2\delta) \\ -(k_1 + 2\delta) & 1 \end{bmatrix}. \tag{41}$$

In the case in which the condition in (34) is satisfied, $Q > 0$, so $\dot{V}$ is negative definite.

We can use (37) and the fact that

$$|s|^{1/2} \le \|\zeta\| \le \frac{V^{1/2}}{{\lambda_{\min}}^{1/2}(P)}$$
$$\|\zeta\| \ge \frac{V^{1/2}}{{\lambda_{\max}}^{1/2}(P)}. \tag{42}$$

Then, substituting (42) into (40) yields

$$\dot{V} \le -\kappa V^{1/2}, \tag{43}$$

where

$$\kappa = \frac{{\lambda_{\min}}^{1/2}(P)\,\lambda_{\min}(Q)}{\lambda_{\max}(P)}. \tag{44}$$

Since the solution of the differential equation

$$\dot{v} \le -\kappa v^{1/2},$$
$$v(0) = v_0 > 0 \tag{45}$$

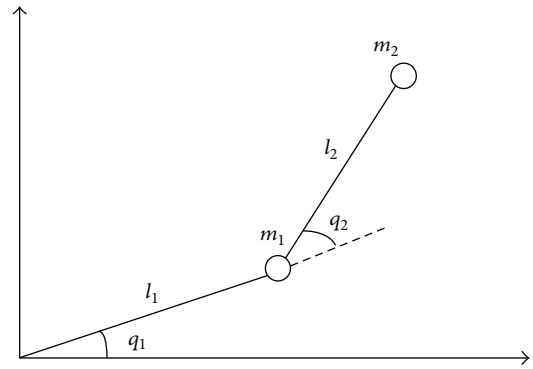


Figure 1: Configuration of the two-link robotic system [3].

is given as

$$v(t) = \left({v_0}^{1/2} - \frac{\kappa}{2}t\right)^2, \tag{46}$$

here, $v(t)$ converges to zero in a finite time and reaches zero after $T = 2V^{1/2}(x_0)/\kappa$. It follows from the comparison principle [18] that $V(t) \le v(t)$ when $V(x_0) \le v_0$. From (46), we can determine that $V(t)$ and therefore s converge to zero in a finite time and reach that value at most after $T = 2V^{1/2}(x_0)/\kappa$. □

4. Simulation Results

In this section, to verify the validity and effectiveness of the proposed method, the two-link planar robot manipulator shown in Figure 1 is considered.

The dynamic equation of the two-link robot is described as follows [3]:

$$M(q)\,\ddot{q} + C(q,\dot{q}) + G(q) = \tau(t) + \tau_d + F(\dot{q}), \tag{47}$$

where

$$M(q) = \begin{bmatrix} l_2^2 m_2 + 2l_1 l_2 m_2 c_2 + l_1^2(m_1 + m_2) & l_2^2 m_2 + l_1 l_2 m_2 c_2 \\ l_2^2 m_2 + l_1 l_2 m_2 c_2 & l_2^2 m_2 \end{bmatrix},$$
$$C(q,\dot{q}) = \begin{bmatrix} -m_2 l_1 l_2 s_2 \dot{q}_2^2 - 2m_2 l_1 l_2 s_2 \dot{q}_1 \dot{q}_2 \\ m_2 l_1 l_2 s_2 \dot{q}_1^2 \end{bmatrix},$$
$$G(q) = \begin{bmatrix} m_2 l_2 g c_{12} + (m_1 + m_2) l_1 g c_1 \\ m_2 l_2 g c_{12} \end{bmatrix}, \tag{48}$$

and $q = (q_1, q_2)^T$ is the joint variable vector, $M(q)$ is the inertial matrix, $C(q,\dot{q})$ represents the centripetal and Coriolis torque matrix, $G(q)$ represents the gravity torque vector, τ_d is the vector of the bounded external disturbance, $F(\dot{q})$ is the friction, and τ is the control torque. m_1 and m_2 are the link masses, l_1 and l_2 are the link lengths, gravity $g = 9.81(m/s^2)$, and the symbols s_1, s_2, s_{12} and c_1, c_2, c_{12} are, respectively, defined as $s_1 = \sin(q_1)$, $s_2 = \sin(q_2)$, $s_{12} = \sin(q_{12})$, $c_1 = \cos(q_1)$, $c_2 = \cos(q_2)$, and $c_{12} = \cos(q_{12})$.

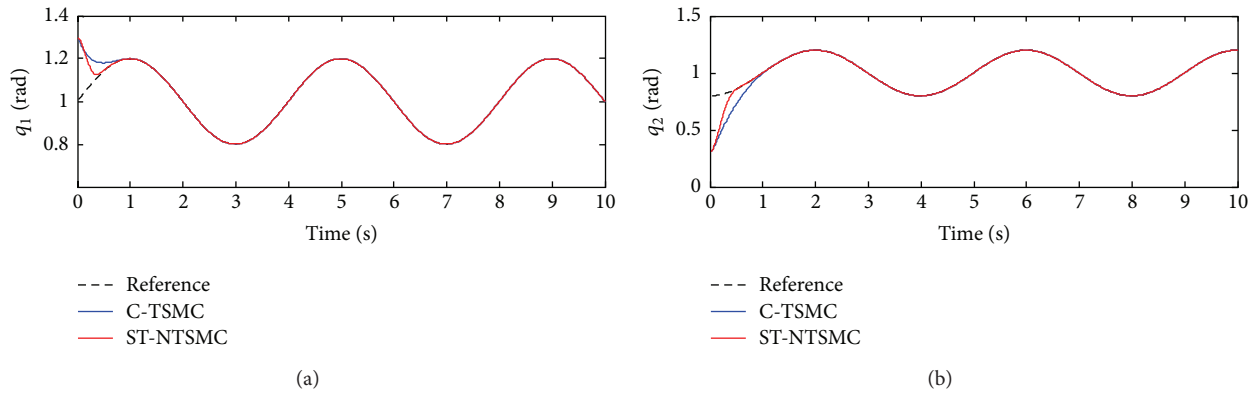


Figure 2: Tracking performance of two-link robot manipulator, (a) at joint 1, (b) at joint 2.

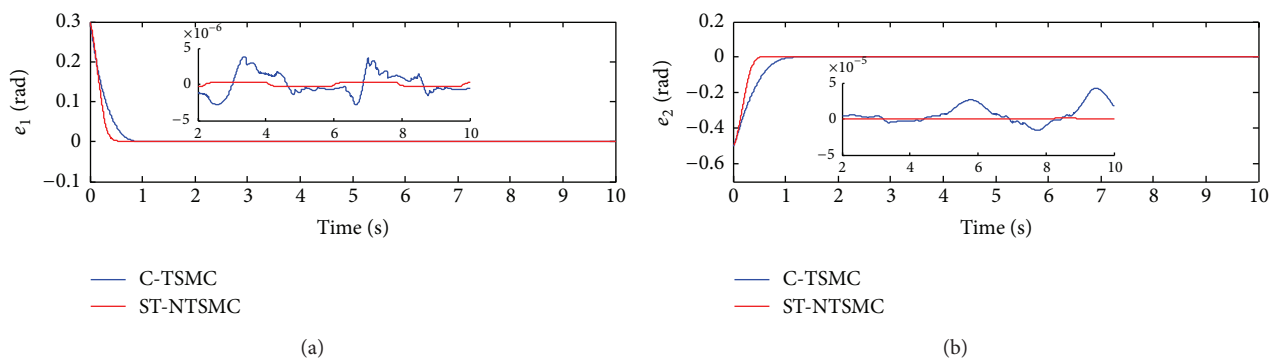


Figure 3: Tracking errors of two-link robot manipulator, (a) at joint 1, (b) at joint 2.

The friction and external disturbance are chosen as

$$F(\dot{q}) = \begin{bmatrix} \dot{q}_1 + 2\sin(q_1) \\ 0.5\sin(q_2) \end{bmatrix}, \qquad \tau_d = \begin{bmatrix} 0.2\sin(t) \\ 0.2\cos(2t) \end{bmatrix}. \tag{49}$$

The parameter values employed to simulate the robot are given as $m_1 = m_2 = 1$ (m) and $l_1 = l_2 = 1$ (kg), and the design reference signals are given by

$$q_{1d} = 1 + 0.2\sin(0.5\pi t), \qquad q_{2d} = 1 - 0.2\cos(0.5\pi t). \tag{50}$$

The initial states of the system are chosen as

$$q_1(0) = 1.3, \quad q_2(0) = 0.3, \quad \dot{q}_1(0) = 0, \quad \dot{q}_2(0) = 0. \tag{51}$$

To this end, Matlab/Simulink is used to perform all of the simulations, and with the sampling time set to 10^{-4} s, the simulation compares the proposed ST-NTSMC control scheme with the previously proposed control method in [28]. Yu et al. [28] suggested the continuous terminal sliding mode control (C-TSMC), which was designed for a two-link robot manipulator as follows:

$$\tau = C_0(q,\dot{q}) + G_0(q) + M_0(q)\ddot{q}_d - M_0\beta^{-1}\gamma^{-1}\operatorname{sig}(\dot{e})^{(2-\gamma)} - M_0(\eta_1 s + \eta_2 \operatorname{sig}(s)^{\varphi}), \tag{52}$$

where $s = e + \beta\operatorname{sig}(\dot{e})^{\gamma}$, $\operatorname{sig}(x)^{\gamma} = [|x_1|^{\gamma}\operatorname{sign}(x_1), |x_2|^{\gamma}\operatorname{sign}(x_2), \ldots, |x_n|^{\gamma}\operatorname{sign}(x_n)]$, $\beta = \operatorname{diag}(\beta_1, \beta_2, \ldots, \beta_n)$, $\eta_1 = \operatorname{diag}(\eta_{11}, \eta_{12}, \ldots, \eta_{1n})$, $\eta_2 = \operatorname{diag}(\eta_{21}, \eta_{22}, \ldots, \eta_{2n})$, $\beta_i, \eta_{1i}, \eta_{2i} > 0$, $0 < \varphi < 1$, and $1 < \gamma < 2$.

The control parameters are selected as shown in Table 1.

Table 1: Control parameters.

Control schemes	Parameters
C-TSMC [28]	$\beta = 10I_2$, $\eta_1 = 25I_2$, $\eta_2 = 45I_2$, $\gamma = 1.5$, $\varphi = 0.3$.
ST-NTSMC	$\beta_1 = 12I_2$, $\beta_2 = 10I_2$, $k_1 = 9$, $k_2 = 5$, $\lambda = 3$, $\alpha = 0.3$.

The simulation results are shown in Figures 2–5. In Figure 2, the tracking results of the robot manipulator using the two control laws above are compared. It shows that the state trajectories can reach the design reference signals in the presence of model parameter uncertainties and external

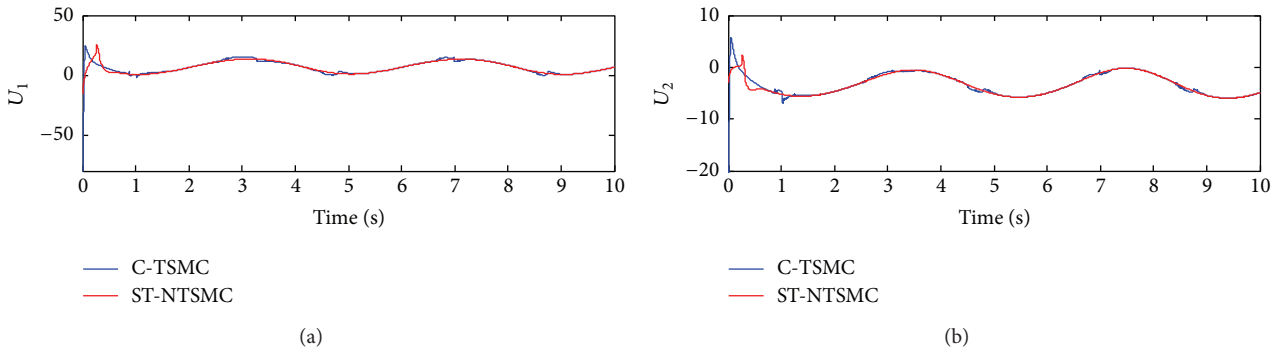


Figure 4: Control inputs, (a) at joint 1, (b) at joint 2.

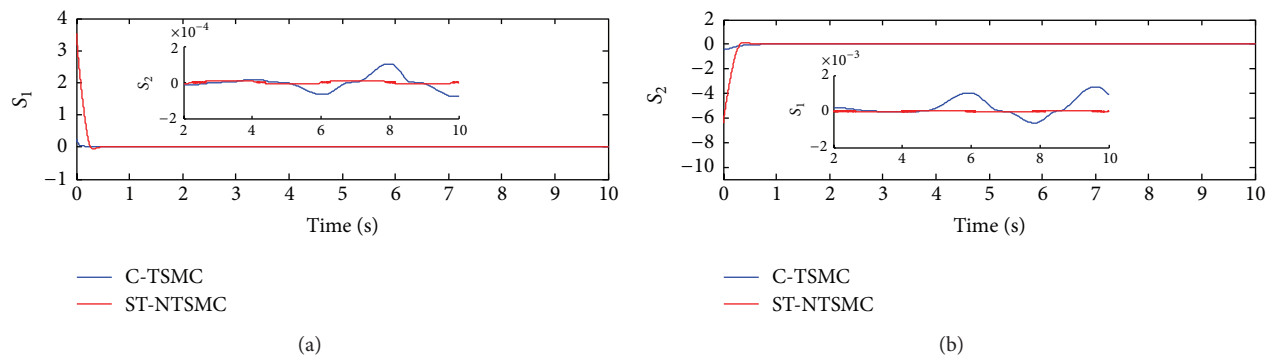


Figure 5: Time responses of the terminal sliding mode surface, (a) at joint 1, (b) at joint 2.

disturbances. The tracking errors via two controllers are compared in Figure 3. One can easily see that the ST-NTSMC produces tracking performance with faster convergence and higher precision. Figure 4 shows the time histories of the applied control inputs and shows that the proposed ST-NTSMC method achieves superior control input performance with smaller control efforts, higher precision tracking, and smoother than the C-TSMC method. The time responses of the sliding manifolds are shown in Figure 5. Clearly, the sliding surface of the proposed method was also much smaller than C-TSMC.

5. Conclusions

In this paper, we presented the ST-NTSMC method for second-order nonlinear systems. This method has been successfully applied in a two-link robot manipulator. The designed nonsingular terminal sliding surface not only avoids the singularity problem, but also can overcome the complex-value and the restriction on the exponent of a power function in conventional TSMC. The performance of the proposed method was evaluated in comparison with recently proposed approaches [28]. The simulation results show that the proposed method achieves highly precise tracking, fast and finite time convergence, and robustness against parameter uncertainties and external disturbances. Furthermore, ST-NTSMC is used to smooth the discontinuous control term in order to attenuate the chattering phenomenon.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This paper is a result of a study on the "Leaders Industry-University Cooperation" Project, supported by the Ministry of Education (MOE).

References

[1] S. Arimoto and F. Miyazaki, "Stability and robustness of PID feedback control for robot manipulators of sensory capability," in *Robotic Research*, M. Brady and R. P. Paul, Eds., MIT Press, Cambridge, Mass, USA, 1984.

[2] J.-J. E. Slotine and W. Li, "On the adaptive control of robot manipulators," *International Journal of Robotics Research*, vol. 6, no. 3, pp. 49–59, 1987.

[3] J. J. Craig, *Introduction to Robotics*, Addion-Wesley, Reading, Mass, USA, 1989.

[4] J. J. Spong and M. Vidyasagar, *Robot Dynamics and Control*, Wiley, New York, NY, USA, 1989.

[5] J. H. Lilly, *Fuzzy Control and Identification*, Wiley, 2010.

[6] L. Jinkun, *Radial Basis Function Neural Network Control for Mechanical Systems*, Tsinghua University Press, Beijing, China, 2013.

[7] V. I. Utkin, "Variable structure systems with sliding modes," *IEEE Transactions on Automatic Control*, vol. 22, no. 2, pp. 212–222, 1977.

[8] J. Y. Hung, W. Gao, and J. C. Hung, "Variable structure control: a survey," *IEEE Transactions on Industrial Electronics*, vol. 40, no. 1, pp. 2–22, 1993.

[9] A. Šabanovic, "Variable structure systems with sliding modes in motion control—a survey," *IEEE Transactions on Industrial Informatics*, vol. 7, no. 2, pp. 212–223, 2011.

[10] S. T. Venkataraman and S. Gulati, "Control of nonlinear systems using terminal sliding modes," *Transactions of the ASME—Journal of Dynamic Systems, Measurement and Control*, vol. 115, no. 3, pp. 554–560, 1993.

[11] Z. Man, A. P. Paplinski, and H. R. Wu, "A robust MIMO terminal sliding mode control scheme for rigid robotic manipulators," *IEEE Transactions on Automatic Control*, vol. 39, no. 12, pp. 2464–2469, 1994.

[12] Z. Man and X. Yu, "Terminal sliding mode control of MIMO linear systems," *IEEE Transactions on Circuits and Systems. I. Fundamental Theory and Applications*, vol. 44, no. 11, pp. 1065–1070, 1997.

[13] Y. Wu, X. Yu, and Z. Man, "Terminal sliding mode control design for uncertain dynamic systems," *Systems & Control Letters*, vol. 34, no. 5, pp. 281–287, 1998.

[14] X. Yu and M. Zhihong, "Fast terminal sliding-mode control design for nonlinear dynamical systems," *IEEE Transactions on Circuits and Systems. I. Fundamental Theory and Applications*, vol. 49, no. 2, pp. 261–264, 2002.

[15] L. Yang and J. Yang, "Nonsingular fast terminal sliding-mode control for nonlinear dynamical systems," *International Journal of Robust and Nonlinear Control*, vol. 21, no. 16, pp. 1865–1879, 2011.

[16] Y. Feng, X. Yu, and Z. Man, "Non-singular terminal sliding mode control of rigid manipulators," *Automatica*, vol. 38, no. 12, pp. 2159–2167, 2002.

[17] T. H. S. Li and Y.-C. Huang, "MIMO adaptive fuzzy terminal sliding-mode controller for robotic manipulators," *Information Sciences*, vol. 180, no. 23, pp. 4641–4660, 2010.

[18] J. J. E. Slotine and W. Li, *Applied Nonlinear Control*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1991.

[19] Y. Jiang, Q. Wang, and C. Dong, "A reaching law neural network terminal sliding mode guidance law design," in *Proceedings of the IEEE Region 10 Conference (TENCON '13)*, Xi'an, China, October 2013.

[20] B. Yoo and W. Ham, "Adaptive fuzzy sliding mode control of nonlinear system," *IEEE Transactions on Fuzzy Systems*, vol. 6, no. 2, pp. 315–321, 1998.

[21] M. Roopaei, M. Zolghadri Jahromi, and S. Jafari, "Adaptive gain fuzzy sliding mode control for the synchronization of nonlinear chaotic gyros," *Chaos*, vol. 19, no. 1, Article ID 013125, 2009.

[22] J. Rivera, C. Mora, J. J. Raygoza, and S. Ortega, "Supper-twisting sliding mode in motion control systems," in *Sliding Mode Control*, A. Bartoszewicz, Ed., pp. 978–953, InTech, 2011.

[23] D. Hernandez, W. Yu, and M. A. Moreno-Amendariz, "Neural PD control with second-order sliding mode compensation for robot manipulators," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN '11)*, pp. 2392–2402, San Jose, Calif, USA, August 2011.

[24] L. M. Capisani, A. Ferrara, and L. Magnani, "Second order sliding mode motion control of rigid robot manipulators," in *Proceedings of the 46th IEEE Conference on Decision and Control (CDC '07)*, pp. 3691–3696, December 2007.

[25] M. Van, H.-J. Kang, and Y.-S. Suh, "Second order sliding mode-based output feedback tracking control for uncertain robot manipulators," *International Journal of Advanced Robotic Systems*, vol. 10, article 16, 2013.

[26] J. Davila, L. Fridman, and A. Levant, "Second-order sliding-mode observer for mechanical systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 11, pp. 1785–1789, 2005.

[27] J. A. Moreno and M. Osorio, "A Lyapunov approach to second-order sliding mode controllers and observers," in *Proceedings of the 47th IEEE Conference on Decision and Control*, pp. 2856–2861, December 2008.

[28] S. Yu, X. Yu, B. Shirinzadeh, and Z. Man, "Continuous finite-time control for robotic manipulators with terminal sliding mode," *Automatica*, vol. 41, no. 11, pp. 1957–1964, 2005.

[29] X. Zhao, Y. X. Jiang, Y. J. Wu, and Y. Q. Zhou, "Fast nonsingular terminal sliding mode control based on multi-slide-mode," *Journal of Beijing University of Aeronautics and Astronautics*, vol. 37, no. 1, pp. 110–113, 2011.

[30] H. J. Shi, L. F. Qian, Y. D. Xu, and L. M. Chen, "Fuzzy moving fast terminal sliding mode control for robotic manipulators," in *Proceedings of the IEEE International Conference on Robotics and Biomimetics (ROBIO '12)*, pp. 1943–1949, IEEE, Guangzhou, China, December 2012.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 947815, 12 pages
http://dx.doi.org/10.1155/2015/947815

Research Article

H_∞ Excitation Control Design for Stochastic Power Systems with Input Delay Based on Nonlinear Hamiltonian System Theory

Weiwei Sun,[1,2] Lianghong Peng,[3] Ying Zhang,[4] and Huaidan Jia[1]

[1] *Institute of Automation, Qufu Normal University, Qufu 273165, China*
[2] *School of Engineering, Qufu Normal University, Rizhao 276826, China*
[3] *School of Automation, Southeast University, Nanjing 210096, China*
[4] *Basic Teaching Department, Shandong Water Polytechnic, Rizhao 276826, China*

Correspondence should be addressed to Weiwei Sun; wwsun@hotmail.com

Received 13 May 2015; Revised 10 August 2015; Accepted 11 August 2015

Academic Editor: Xinguang Zhang

This paper presents H_∞ excitation control design problem for power systems with input time delay and disturbances by using nonlinear Hamiltonian system theory. The impact of time delays introduced by remote signal transmission and processing in wide-area measurement system (WAMS) is well considered. Meanwhile, the systems under investigation are disturbed by random fluctuation. First, under prefeedback technique, the power systems are described as a nonlinear Hamiltonian system. Then the H_∞ excitation controller of generators connected to distant power systems with time delay and stochasticity is designed. Based on Lyapunov functional method, some sufficient conditions are proposed to guarantee the rationality and validity of the proposed control law. The closed-loop systems under the control law are asymptotically stable in mean square independent of the time delay. And we through a simulation of a two-machine power system prove the effectiveness of the results proposed in this paper.

1. Introduction

Time delay always exists in power systems control area. It is often ignored when controller is mainly applied in local systems where the communication time delay is very small compared to the system time constants (see, e.g., [1, 2] and the references therein). Due to the further study of phase measurement unit (PMU) and WAMS, coordinated stability control has got a lot of attention. It uses remote measuring information given by WAMS/PMU. Unlike the small delay in local control, the time delay in wide-area power systems can vary from tens to several hundred milliseconds or more. Since that the large time delay will go against the stability of the system and reduce the performance of the system, so it is very necessary to consider the influence of it on the power system stability analysis and controller design. Besides, the generators are interfered with speed regulation, fluctuation of load, mechanical torsional vibration, the changes of damping coefficients, and so on in the transient process. These random fluctuations can be regarded as a kind of random process [3]. However, the application of the Itô differential formula will lead to the appearance of gravitation and the Hessian term. What is more, the stochastic disturbance (Wiener process) will cause no definition of the system states' derivative [4]. Therefore, stochastic and delay factors increase the difficulties of the analysis and synthesis [5]. Some results, which took signal transmission time delays or stochasticity in power systems into account, have been obtained. Reference [6] presented a free-weighting matrix method based on linear control design approach for the wide-area robust damping controller associated with flexible alternating current transmission system device to improve the dynamical performance of the large-scale power systems. Reference [7] proposed a delay-independent decentralized coordinated robust approach to design excitation controller in terms of H_∞ optimization method incorporating linear matrix inequality (LMI) technique. Considering the nonlinear effects of randomized torsional oscillation on the excitation regulation dynamic process of a generator rotor and exploiting Monte-Carlo principle and numerical methods,

the algorithms and workflow of the proposed excitation control system's transient stability analysis approach were presented in [3]. Reference [8] presented a stochastic cost model and a solution technique for optimal scheduling of the generators in a wind integrated power system considering the demand and wind generation uncertainties.

Based on the linearization at steady state operating point, lots of the techniques are by far achieved and applied to controller design in power systems. These techniques have some disadvantages, such as ignoring some nonlinearities of the system and just expressing the partial structures of the system. What is more, the designed controllers are generally relatively complicated and not very easy to realize online operation. Therefore, some nonlinear methods should be worked out to achieve good control performance for the power systems in consideration of time-delay, stochastics, and disturbances. In recent years, energy-based Lyapunov function method has obtained numerous attention, and remarkable achievements have been reached with this method in the analysis and synthesis of nonlinear systems, as well as in the power systems (see, e.g., [9–13] and the references therein). The method can thoroughly take advantage of the internal structural properties of the systems and make the control design relatively simple. An important step in using energy-based control strategy is to transform the system into a dissipative Hamiltonian system formulation. This kind of system, proposed by [14], has great benefits for that its Hamilton function can be used as the sum of potential energy (excluding gravitational potential energy) and kinetic energy in physical systems and also can be taken as a Lyapunov function (see, e.g., [11, 15–18]). Using the energy-based Hamilton function method, [11] investigated the adaptive H_∞ excitation control of multimachine power systems with disturbances. Simulations show that the control strategy proposed in [11] was more effective than some other control schemes. Considering the impact of time delays in acquisition and transmission of key signals in power systems, [19] deals with the H_∞ excitation control problem of n-machine power system with time-delay and disturbances.

The purpose of this paper is to present a suitable controller structure for the stochastic power systems with input delay and disturbances using the nonlinear Hamiltonian system theory in order to weaken the impact of stochasticity and delay on the control performance of the power systems. Firstly, the prefeedback with delay method is to be used to describe the system as a dissipative Hamiltonian system formulation. Next, based on the obtained new system formulation, we will deal with the H_∞ control problem by using Newton-Leibniz formula, a few properties of norm and matrices. The main results will be proposed for the Hamiltonian system and the power system as well. Finally, we will test and verify the obtained results in this paper by an example of a two-machine power system with delay, stochasticity, and disturbances.

The rest of the paper is organized as follows. Section 2 provides the problem formulation, nonlinear Hamilton realization and some preliminaries. Section 3 gives the main results. Analysis of the achieved results by a two-machine power system example and the conclusion are given in Sections 4 and 5, respectively.

Notations. Throughout the paper the superscript "T" stands for matrix transposition. $\mathscr{R}$ denotes the set of real numbers, $\mathscr{R}_+$ the set of all nonnegative real numbers, $\mathscr{R}^n$ the n-dimensional Euclidean space, and $\mathscr{R}^{n\times m}$ the real matrices with dimension $n \times m$. Diag$\{\cdots\}$ stands for diagonal matrix in which the diagonal elements are the elements in $\{\cdots\}$; $\|\cdot\|$ stands for either the Euclidean vector norm or the induced matrix 2-norm. For any symmetric matrices X and Y, $X \geq Y$ (resp., $X > Y$) means that the matrix X-Y is positive semidefinite (resp., positive definite). tr$[X]$ denotes the trace for square matrix X. $\lambda_{\max}(P)$ ($\lambda_{\min}(P)$) denotes the maximum (minimum) of eigenvalue of a real symmetric matrix P. $\mathscr{C}_{n,\tau} = \mathscr{C}([-\tau, 0], \mathscr{R}^n)$ means the Banach space of continuous functions from $[-\tau, 0]$ to $\mathscr{R}^n$. $\mathrm{C}^b_{\mathscr{F}_0}([-\tau, 0]; \mathscr{R}^n)$ denotes the family of all $\mathscr{F}_0$-measurable bounded $\mathrm{C}([-\tau, 0]; \mathscr{R}^n)$-valued random variables $\phi = \{\phi(t) : t \in [-\tau, 0]\}$. C^i denotes the set of all functions with continuous ith partial derivatives; $\mathrm{C}^{2,1}$ is the family of all functions which are C^2 in the first argument and C^1 in the second argument; $\mathrm{C}^{2,1}(\mathscr{R}^n \times [-\tau, \infty); \mathscr{R}_+)$ stands for the family of all nonnegative functions $V(x, t)$ on $\mathscr{R}^n \times [-\tau, \infty)$ which are C^2 in x and C^1 in t. What is more, for the sake of simplicity, throughout the paper, we denote $\partial H/\partial x$ by ∇H.

2. Problem Formulation and Nonlinear Hamilton Realization

Consider the following n-machine power systems, each generator of which is described by a third-order dynamic model (see [1, 20]):

$$\begin{aligned}
\dot{\delta}_i &= \omega_i - \omega_0, \\
\dot{\omega}_i &= \frac{\omega_0}{M_i} P_{mi} - \frac{D_i}{M_i}(\omega_i - \omega_0) - \frac{\omega_0}{M_i} P_{ei} + \epsilon_{i1}, \\
\dot{E}'_{qi} &= -\frac{1}{T_{d0i}} E_{qi} + \frac{1}{T_{d0i}} u_{fi}(t) + \epsilon_{i2},
\end{aligned} \tag{1}$$

where

$$\begin{aligned}
E_{qi} &= E'_{qi} + I_{di}\left(x_{di} - x'_{di}\right), \\
I_{di} &= B_{ii} E'_{qi} - \sum_{j=1, j\neq i}^{n} B_{ij} E'_{qj} \cos\left(\delta_i - \delta_j\right), \\
P_{ei} &= G_{ii} E'^2_{qi} + E'_{qi} \sum_{j=1, j\neq i}^{n} B_{ij} E'_{qj} \sin\left(\delta_i - \delta_j\right), \\
& \qquad i = 1, 2, \ldots, n,
\end{aligned} \tag{2}$$

δ_i is the power angle of the ith generator (radians), ω_i is the rotor speed of the ith generator (rad/s), $\omega_0 = 2\pi f_0$, E'_{qi} is the q-axis internal transient voltage of the ith generator (per unit), x_{di} is the d-axis transient reactance (per unit), x'_{di} is the d-axis transient reactance of the ith generator (per unit),

u_{fi} is the voltage of the field circuit of the ith generator, the control input (per unit), M_i is the inertia coefficient of the ith generator (s), D_i is the damping constant (per unit), T_{d0i} is the field circuit time constant (s), P_{mi} is the mechanical power, assumed to be constant (per unit), P_{ei} is the active electrical power (per unit), ϵ_{i1} and ϵ_{i2} are bounded disturbances, I_{di} is the d-axis current (per unit), E_{qi} is the internal voltage (per unit), $Y_{ij} = G_{ij} + jB_{ij}$ is the admittance of line i-j (per unit), and $Y_{ii} = G_{ii} + jB_{ii}$ is the self-admittance of bus i (per unit).

There are signal transmission delays and random process in the modern power systems. The delays in the measuring data exist in such case that the exciter inputs are taken from remote buses. And assume that all the feedback wide-area signals have the time delay τ. Meanwhile, the generator torque can be regarded as a kind of random process because of random fluctuation in transient process, such as speed regulation, fluctuation of load, mechanical torsional vibration, and the changes of damping coefficients. Moreover, considering the imaginary control input is u_{fi} which feeds back both the local measurement information and the wide-area measurement signals, so the power system (1) should be modeled into differential-algebraic equations with time delay and stochasticity as follows:

$$\begin{aligned} d\delta_i &= (\omega_i - \omega_0)\, dt, \\ d\omega_i &= \left[\frac{\omega_0}{M_i}P_{mi} - \frac{D_i}{M_i}(\omega_i - \omega_0) - \frac{\omega_0}{M_i}P_{ei} + \epsilon_{i1}\right] dt \\ &\quad + \frac{\xi}{M_i}(\omega_i - \omega_0)\, dw(t), \\ dE'_{qi} &= \left[-\frac{1}{T_{d0i}}E_{qi} + \frac{1}{T_{d0i}}u_{fi}(t-\tau) + \epsilon_{i2}\right] dt, \end{aligned} \tag{3}$$

where ξ is random disturbance intensity and $w(t)$ is a zero-mean Wiener process on a probability space $(\Omega, \mathscr{F}, \mathscr{P})$ relative to an increasing family $(\mathscr{F}_t)_{t>0}$ of σ algebras $(\mathscr{F}_t)_{t>0} \subset \mathscr{F}$; here Ω is the samples space, $\mathscr{F}$ is σ algebra of subsets of the sample space, and P is the probability measure on $\mathscr{F}$. Moreover, we assume $E\{dw(t)\} = 0$, $E\{[dw(t)]^2\} = dt$, where E is the expectation operator.

Assume that $(\delta_i^{(0)}, \omega_0, E_{qi}'^{(0)})$, $i = 1, 2, \ldots, n$, are the preassigned operating points of system (3).

Setting $x_{i1} = \delta_i$, $x_{i2} = \omega_i - \omega_0$, $x_{i3} = E'_{qi}$, $(\omega_0/M_i)P_{mi} = a_i$, $D_i/M_i = b_i$, $(\omega_0/M_i)G_{ii} = c_i$, $\omega_0/M_i = d_i$, $1/T_{d0i} = e_i$, $(x_{di} - x'_{di})/T_{d0i} = h_i$, and $(1/T_{d0i})u_{fi}(t-\tau) = v_i(t-\tau)$, $i = 1, 2, \ldots, n$, then system (3) can be rewritten as follows:

$$\begin{aligned} dx_{i1} &= x_{i2}dt, \\ dx_{i2} &= \Bigg[a_i - b_i x_{i2} - c_i x_{i3}^2 + \epsilon_{i1} \\ &\quad - d_i x_{i3} \sum_{j=1, j\neq i}^{n} B_{ij} x_{j3} \sin\left(x_{i1} - x_{j1}\right)\Bigg] dt + \frac{\xi}{M_i} \\ &\quad \cdot x_{i2} dw(t), \\ dx_{i3} &= \Bigg[-\left(e_i + h_i B_{ii}\right) x_{i3} + v_i(t-\tau) + \epsilon_{i2} \\ &\quad + h_i \sum_{j=1, j\neq i}^{n} B_{ij} x_{j3} \cos\left(x_{i1} - x_{j1}\right)\Bigg] dt, \\ &\qquad i = 1, 2, \ldots, n. \end{aligned} \tag{4}$$

Inspired by [11], we introduce a prefeedback control law:

$$\begin{aligned} v_i(t-\tau) &= -\frac{2c_i h_i}{d_i} x_{i1}(t-\tau)\, x_{i3}(t-\tau) - k_i x_{i3}(t-\tau) \\ &\quad + \overline{u}_i + u_i(t-\tau), \quad i = 1, 2, \ldots, n, \end{aligned} \tag{5}$$

where the first term is to make system (4) have a Hamilton structure, the second and third terms are to guarantee the operating point of the system unchanged, $u_i(t-\tau)$ is the new reference input, and $\overline{u}_i$ and k_i are undetermined constants. To make the operating point of the system invariant, $\overline{u}_i$ and k_i have to satisfy

$$\begin{aligned} &-\left(e_i + h_i B_{ii}\right) E_{qi}'^{(0)} - \frac{2c_i h_i}{d_i}\delta_i^{(0)} E_{qi}'^{(0)} - k_i E_{qi}'^{(0)} + \overline{u}_i \\ &\quad + h_i \sum_{j=1, j\neq i}^{n} B_{ij} E_{qj}'^{(0)} \cos\left(\delta_i^{(0)} - \delta_j^{(0)}\right) = 0, \\ &\qquad i = 1, 2, \ldots, n, \end{aligned} \tag{6}$$

and $k_i = k_{i0}$ which is spelled out in [11]; what is more, this reference provides a kind of choice of $\overline{u}_i$ and k_i.

Furthermore, (5) can be rewritten as

$$\begin{aligned} &v_i(t-\tau) \\ &\quad = -\frac{2c_i h_i}{d_i} x_{i1}(t)\, x_{i3}(t) - k_i x_{i3}(t) + \overline{u}_i \\ &\quad - \frac{2c_i h_i}{d_i}\left[x_{i1}(t-\tau)\, x_{i3}(t-\tau) - x_{i1}(t)\, x_{i3}(t)\right] \\ &\quad - k_i\left[x_{i3}(t-\tau) - x_{i3}(t)\right] + u_i(t-\tau), \\ &\qquad i = 1, 2, \ldots, n. \end{aligned} \tag{7}$$

Let $x_i = [x_{i1}, x_{i2}, x_{i3}]^{\mathrm{T}}$, $\epsilon_i = [\epsilon_{i1}, \epsilon_{i2}]^{\mathrm{T}}$, then system (4) can be expressed as a dissipative Hamiltonian system as follows:

$$\begin{aligned} dx_i &= \Bigg\{\left(J_i - R_i\right)\nabla H_i(x_i) + g_1 u_i(t-\tau) + \frac{2c_i h_i}{d_i} \\ &\quad \cdot g_1\left[g_3^{\mathrm{T}} x_{i1}(t)\, g_1^{\mathrm{T}} x_{i3}(t)\right. \\ &\quad \left. - g_3^{\mathrm{T}} x_{i1}(t-\tau)\, g_1^{\mathrm{T}} x_{i3}(t-\tau)\right] + k_i g_1\left[g_1^{\mathrm{T}} x_{i3}(t)\right. \\ &\quad \left. - g_1^{\mathrm{T}} x_{i3}(t-\tau)\right] + g_2 \epsilon_i\Bigg\} dt + g_4^{(i)}(x_i)\, dw(t), \\ &\qquad i = 1, 2, \ldots, n, \end{aligned} \tag{8}$$

where

$$\begin{aligned}
H_i(x_i) &= -\frac{a_i}{d_i}x_{i1} + \frac{c_i}{d_i}x_{i1}x_{i3}^2 + \frac{e_i + h_iB_{ii} + k_i}{2h_i}x_{i3}^2 \\
&\quad + \frac{1}{2d_i}x_{i2}^2 - \frac{\overline{u}_i}{h_i}x_{i3} \\
&\quad - x_{i3}\sum_{j=1,j\neq i}^{n} B_{ij}x_{j3}\cos\left(x_{i1} - x_{j1}\right), \\
J_i &= \begin{pmatrix} 0 & d_i & 0 \\ -d_i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\
R_i &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & b_id_i & 0 \\ 0 & 0 & h_i \end{pmatrix}, \\
g_1 &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \\
g_2 &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \\
g_3 &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \\
g_4^{(i)}(x_i) &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & \dfrac{\xi}{M_i} & 0 \\ 0 & 0 & 0 \end{pmatrix} x_i.
\end{aligned} \tag{9}$$

$\nabla H_i(x_i)$ is the gradient of the Hamilton function $H_i(x_i)$, which satisfies $H_i(0) = 0$, $i = 1, 2, \ldots, n$.

Owing to each individual subsystem having the cross-variables, this structure does not provide the overall system a Hamilton structure. Thus, we need to find out a common Hamilton function for the n generators, which is regarded as the total energy of the whole system.

Let

$$\begin{aligned}
H(x) &= \sum_{i=1}^{n} H_i + \frac{1}{2}\sum_{i=1}^{n} x_{i3}\sum_{j=1,j\neq i}^{n} B_{ij}x_{j3}\cos\left(x_{i1} - x_{j1}\right) \\
&= \sum_{i=1}^{n}\left[-\frac{a_i}{d_i}x_{i1} + \frac{c_i}{d_i}x_{i1}x_{i3}^2 + \frac{1}{2d_i}x_{i2}^2\right. \\
&\quad + \frac{e_i + h_iB_{ii} + k_i}{2h_i}x_{i3}^2 - \frac{\overline{u}_i}{h_i}x_{i3} \\
&\quad \left. - \frac{1}{2}x_{i3}\sum_{j=1,j\neq i}^{n} B_{ij}x_{j3}\cos\left(x_{i1} - x_{j1}\right)\right],
\end{aligned} \tag{10}$$

where $x = [x_1^{\mathrm{T}}, \ldots, x_n^{\mathrm{T}}]^{\mathrm{T}}$. By using relation $B_{ij} = B_{ji}$, we can verify that

$$\frac{\partial H(x)}{\partial x_{ij}} = \frac{\partial H_i(x_i)}{\partial x_{ij}}, \quad i = 1, 2, \ldots, n, \ j = 1, 2, 3, \tag{11}$$

which imply that $H(x)$ is the common Hamilton function for the n generators. Furthermore, $H(x) \in \mathrm{C}^2$ holds obviously.

Setting

$$\begin{aligned}
u &= \left[u_1, \ldots, u_n\right]^{\mathrm{T}}, \\
\epsilon &= \left[\epsilon_1^{\mathrm{T}}, \ldots, \epsilon_n^{\mathrm{T}}\right]^{\mathrm{T}}, \\
y &= \left[y_1^{\mathrm{T}}, \ldots, y_n^{\mathrm{T}}\right]^{\mathrm{T}},
\end{aligned} \tag{12}$$

then system (8) can be rewritten as follows:

$$\begin{aligned}
dx(t) &= \left\{(J - R)\nabla H(x) + G_1u(t-\tau)\right. \\
&\quad + 2G_1C\left[G_3^{\mathrm{T}}x(t)G_1^{\mathrm{T}}x(t) - G_3^{\mathrm{T}}x(t-\tau)G_1^{\mathrm{T}}(t-\tau)\right] \\
&\quad \left. + G_1KG_1^{\mathrm{T}}\left[x(t) - x(t-\tau)\right] + G_2\epsilon\right\}dt + G_4(x)\,dw(t), \\
x(t) &= \phi(t), \quad t \in [-\tau, 0],
\end{aligned} \tag{13}$$

where $J = \mathrm{Diag}\{J_1, \ldots, J_n\}$, $R = \mathrm{Diag}\{R_1, \ldots, R_n\}$, $C = \mathrm{Diag}\{c_1h_1/d_1, \ldots, c_nh_n/d_n\}$, $K = \mathrm{Diag}\{k_1, \ldots, k_n\}$,

$$\begin{aligned}
G_1 &= \begin{pmatrix} g_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & g_1 \end{pmatrix}_{3n\times n}, \\
G_2 &= \begin{pmatrix} g_2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & g_2 \end{pmatrix}_{3n\times 2n}, \\
G_3 &= \begin{pmatrix} g_3 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & g_3 \end{pmatrix}_{3n\times n}, \\
G_4(x) &= \begin{pmatrix} g_4^{(1)}(x_1) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & g_4^{(n)}(x_n) \end{pmatrix}_{3n\times n}.
\end{aligned} \tag{14}$$

Obviously, J is a skew-symmetric matrix, and R is a positive semidefinite matrix. In addition, we can choose $y = G_2^{\mathrm{T}}\nabla H(x)$ and $z = PG_1^{\mathrm{T}}\nabla H(x)$ as the output and the penalty signal, respectively, where P is a full column rank weighting matrix.

Definition 1. The stochastic time delay Hamiltonian system (13) is said to be robustly asymptotically stable in mean square, if there exists a controller $u(t-\tau)$ such that

$$\lim_{t\to\infty} E\left\{\left\|x(t) - x_0\right\|^2\right\} = 0, \tag{15}$$

where x_0 is the preassigned equilibrium and $x(t)$ is the solution of system (13) at time t under initial condition.

Consider the following cost function:

$$C(T_0) = E\left\{\int_0^{T_0} z^{\mathrm{T}}(t)\, z(t)\, dt\right\} - \gamma^2 E\left\{\int_0^{T_0} \epsilon^{\mathrm{T}}(t)\, \epsilon(t)\, dt\right\}, \quad \forall T_0 > 0. \tag{16}$$

Then H_∞ control objective of system (13) is to find a feedback controller:

$$u(t-\tau) = \alpha(t-\tau) \tag{17}$$

such that

$$C(\infty) < 0 \quad (T_0 \longrightarrow \infty), \tag{18}$$

for given $\gamma > 0$ and at the same time the closed-loop system is asymptotically stable when $\epsilon = 0$.

We conclude this section by recalling some auxiliary results to be used in this paper.

Lemma 2 (see [21]). *For system*

$$dx(t) = f(x(t), x(t-\tau))\, dt + g(x(t), x(t-\tau))\, dw(t), \quad \forall t \geq 0, \tag{19}$$

assume that $f(x, y)$ and $g(x, y)$ are locally Lipschitz in (x, y). If there exists a function $V(x,t) \in C^{2,1}(\mathscr{R}^n \times [-\tau, \infty); \mathscr{R}_+)$ such that for some constant $K > 0$ and any $t \geq 0$,

$$\begin{aligned} &\mathscr{L}V \leq K(1 + V(x(t), t) + V(x(t-\tau), t-\tau)), \\ &\lim_{|x| \to \infty} \inf_{t \geq 0} V(x,t) = \infty, \end{aligned} \tag{20}$$

where the differential operator $\mathscr{L}$ is defined as

$$\mathscr{L}V = \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} f(x(t), x(t-\tau)) + \frac{1}{2} \cdot tr\left\{g^{\mathrm{T}}(x(t), x(t-\tau)) \frac{\partial^2 V}{\partial x^2} g(x(t), x(t-\tau))\right\}, \tag{21}$$

then there exists a unique solution on $[-\tau, \infty)$ for any initial data $\{x(t) = \phi(t) : t \in [-\tau, 0]\} \in C^b_{\mathscr{F}_0}([-\tau, 0]; \mathscr{R}^n)$.

Lemma 3. *For any given matrices $A \in \mathscr{R}^{n\times r}$ and $B \in \mathscr{R}^{n\times r}$, there holds*

$$tr\left(A^{\mathrm{T}}B\right) \leq \frac{1}{2}\left[tr\left(A^{\mathrm{T}}A\right) + tr\left(B^{\mathrm{T}}B\right)\right]. \tag{22}$$

Proof. This proof can be achieved by using the properties of matrix's trace. □

3. Main Results

3.1. Hamiltonian System. The H_∞ controller is given below for the stochastic Hamiltonian system (13) with input delay.

Theorem 4. *Consider system (13) and the following assumptions are satisfied:*

(A1) $\nabla H(x_0) = 0$;

(A2) $Hess(H(x_0)) > 0$;

(A3) $H(x) - H(x_0) \geq (\alpha_1/2)\|x - x_0\|^2$;

(A4) $\nabla^{\mathrm{T}} H(x) \cdot \nabla H(x) \geq \beta_1 \|x - x_0\|^2$.

If

$$2R + \frac{1}{\gamma^2} G_1 G_1^{\mathrm{T}} - \frac{1}{\gamma^2} G_2 G_2^{\mathrm{T}} \geq 0 \tag{23}$$

holds, then the H_∞ control problem of system (13) can be solved by the feedback control law:

$$\begin{aligned} u(t-\tau) = &-\frac{1}{2}\left(\frac{1}{\gamma^2} G_1^{\mathrm{T}} + P^{\mathrm{T}} P G_1^{\mathrm{T}}\right) \nabla H(x(t-\tau)) \\ &- \frac{1}{4}\left(G_1^{\mathrm{T}} G_1\right)^{-1} G_1^{\mathrm{T}} G_5 \nabla H(x(t-\tau)) \\ &+ 2CX(t-\tau) + K G_1^{\mathrm{T}} x(t-\tau) - M - 2N \\ &- \frac{1}{2}\tau\lambda_1\lambda_2 T - \frac{1}{4}\tau\lambda_1\lambda_2\left(G_1^{\mathrm{T}} G_1\right)^{-1} G_1^{\mathrm{T}} G_5, \end{aligned} \tag{24}$$

where x_0 is the preassigned equilibrium of system (13), $G_5 = \mathrm{Diag}\{g_5^{(1)}, \ldots, g_5^{(n)}\}$,

$$\begin{aligned} g_5 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & \dfrac{\left(d_i^2 + 1\right)\epsilon^2}{M_i^2} & 0 \\ 0 & 0 & 0 \end{pmatrix}_{3\times 3}, \\ X(t) &= \begin{pmatrix} x_{11}(t)\, x_{13}(t) \\ x_{21}(t)\, x_{23}(t) \\ \ddots \\ x_{n1}(t)\, x_{n3}(t) \end{pmatrix}_{n\times 1}, \end{aligned} \tag{25}$$

M, N, T are all positive constant matrices which satisfy $\|M\| \geq \|KG_1^{\mathrm{T}} x(t)\|$, $\|N\| \geq \|CX(t)\|$, $\|T\| \geq \|(1/\gamma^2)G_1^{\mathrm{T}} + P^{\mathrm{T}} P G_1^{\mathrm{T}}\|$, and λ_1 and λ_2 are constants which satisfy $\lambda_1 \geq \sup_{t\geq -\tau}\|Hess(H(x(t)))\|$, $\lambda_2 \geq \sup_{t\geq -\tau}\|\dot{x}(t)\|$.

Proof. Take a Lyapunov candidate function as follows:

$$V(x) = 2H(x) - 2H(x_0). \tag{26}$$

According to Itô differential formula, it follows that

$$dV(x) = \mathscr{L}V(x)\, dt + \nabla V(x)\, G_4(x)\, dw(t). \tag{27}$$

According to (21) in Lemma 2, one has

$$\begin{aligned}
&\mathscr{L}V(x)\\
&= \frac{1}{2}\operatorname{tr}\left\{g^{\mathrm{T}}(x(t),x(t-\tau))\frac{\partial^2 V}{\partial x^2}g(x(t),x(t-\tau))\right\}\\
&\quad+\frac{\partial V}{\partial t}+\frac{\partial V}{\partial x}f(x(t),x(t-\tau))\\
&= \operatorname{tr}\left[G_4^{\mathrm{T}}(x)\operatorname{Hess}(H(x))G_4(x)\right]\\
&\quad+2\nabla^{\mathrm{T}}H(x)(J-R)\nabla H(x)\\
&\quad+2\nabla^{\mathrm{T}}H(x)G_1KG_1^{\mathrm{T}}[x(t)-x(t-\tau)]\\
&\quad+4\nabla^{\mathrm{T}}H(x)G_1C[X(t)-X(t-\tau)]\\
&\quad+2\nabla^{\mathrm{T}}H(x)G_2\epsilon\\
&\quad-\nabla^{\mathrm{T}}H(x)G_1\left(\frac{1}{\gamma^2}G_1^{\mathrm{T}}+P^{\mathrm{T}}PG_1^{\mathrm{T}}\right)\nabla H(t-\tau)\\
&\quad-\frac{1}{2}\nabla^{\mathrm{T}}H(x)G_1\left(G_1^{\mathrm{T}}G_1\right)^{-1}G_1^{\mathrm{T}}G_5\nabla H(t-\tau)\\
&\quad+4\nabla^{\mathrm{T}}H(x)G_1CX(t-\tau)\\
&\quad+2\nabla^{\mathrm{T}}H(x)G_1KG_1^{\mathrm{T}}x(t-\tau)\\
&\quad-2\nabla^{\mathrm{T}}H(x)G_1(M+2N)-\tau\lambda_1\lambda_2\nabla^{\mathrm{T}}H(x)G_1T\\
&\quad-\frac{1}{2}\tau\lambda_1\lambda_2\nabla^{\mathrm{T}}H(x)G_1\left(G_1^{\mathrm{T}}G_1\right)^{-1}G_1^{\mathrm{T}}G_5.
\end{aligned}\tag{28}$$

Based on the facts of Lemma 3 and Condition (22), we can achieve

$$\begin{aligned}
&\operatorname{tr}\left[G_4^{\mathrm{T}}(x)\operatorname{Hess}(H(x))G_4(x)\right]\\
&\quad\le\frac{1}{2}\operatorname{tr}\left[G_4^{\mathrm{T}}(x)\operatorname{Hess}(H(x))\operatorname{Hess}^{\mathrm{T}}(H(x))G_4(x)\right]\\
&\quad+\frac{1}{2}\operatorname{tr}\left[G_4^{\mathrm{T}}(x)G_4(x)\right]=\nabla^{\mathrm{T}}H(x)G_5\nabla H(x).
\end{aligned}\tag{29}$$

According to Newton-Leibniz formula, it follows that

$$\nabla H(x_\tau)=\nabla H(x)-\int_{t-\tau}^{t}\operatorname{Hess}(H(x(s)))\dot{x}(s)\,ds.\tag{30}$$

Therefore, the following equalities hold:

$$\begin{aligned}
&-\nabla^{\mathrm{T}}H(x)G_1\left(\frac{1}{\gamma^2}G_1^{\mathrm{T}}+P^{\mathrm{T}}PG_1^{\mathrm{T}}\right)\nabla H(x(t-\tau))\\
&\quad=-\nabla^{\mathrm{T}}H(x)G_1\left(\frac{1}{\gamma^2}G_1^{\mathrm{T}}+P^{\mathrm{T}}PG_1^{\mathrm{T}}\right)\\
&\quad\cdot\left[\nabla H(x)-\int_{t-\tau}^{t}\operatorname{Hess}(H(x(s)))\dot{x}(s)\,ds\right],\\
&-\nabla^{\mathrm{T}}H(x)G_1\left(G_1^{\mathrm{T}}G_1\right)^{-1}G_1^{\mathrm{T}}G_5\nabla H(x(t-\tau))\\
&\quad=-\nabla^{\mathrm{T}}H(x)G_1\left[G_1^{\mathrm{T}}G_1\right]^{-1}\\
&\quad\cdot G_1^{\mathrm{T}}G_5\left[\nabla H(x)-\int_{t-\tau}^{t}\operatorname{Hess}(H(x(s)))\dot{x}(s)\,ds\right].
\end{aligned}\tag{31}$$

According to the Mean Value Theorem of Integrals, there exists $\theta\in[t-\tau,t]$ that satisfies

$$\begin{aligned}
&\int_{t-\tau}^{t}\operatorname{Hess}(H(x(s)))\dot{x}(s)\,ds\\
&\quad=\tau\operatorname{Hess}(H(x(\theta)))\dot{x}(\theta).
\end{aligned}\tag{32}$$

Consequently, we have

$$\begin{aligned}
&\nabla^{\mathrm{T}}H(x)G_1\left(\frac{1}{\gamma^2}G_1^{\mathrm{T}}+P^{\mathrm{T}}PG_1^{\mathrm{T}}\right)\\
&\quad\cdot\int_{t-\tau}^{t}\operatorname{Hess}(H(x(s)))\dot{x}(s)\,ds=\tau\nabla^{\mathrm{T}}H(x)\\
&\quad\cdot G_1\left(\frac{1}{\gamma^2}G_1^{\mathrm{T}}+P^{\mathrm{T}}PG_1^{\mathrm{T}}\right)\operatorname{Hess}(H(x(\theta)))\dot{x}(\theta)\\
&\quad\le\tau\lambda_1\lambda_2\nabla^{\mathrm{T}}H(x)G_1T.
\end{aligned}\tag{33}$$

Similarly, we further obtain

$$\begin{aligned}
&\frac{1}{2}\nabla^{\mathrm{T}}H(x)G_1\left(G_1^{\mathrm{T}}G_1\right)^{-1}\\
&\quad\cdot G_1^{\mathrm{T}}G_5\int_{t-\tau}^{t}\operatorname{Hess}(H(x(s)))\dot{x}(s)\,ds=\frac{1}{2}\tau\nabla^{\mathrm{T}}H(x)\\
&\quad\cdot G_1\left(G_1^{\mathrm{T}}G_1\right)^{-1}G_1^{\mathrm{T}}G_5\operatorname{Hess}(H(x(\theta)))\dot{x}(\theta)\le\frac{1}{2}\\
&\quad\cdot\tau\lambda_1\lambda_2\nabla^{\mathrm{T}}H(x)G_1\left(G_1^{\mathrm{T}}G_1\right)^{-1}G_1^{\mathrm{T}}G_5.
\end{aligned}\tag{34}$$

Combining the above inequalities, we can conclude that

$$\begin{aligned}
&\mathscr{L}V(x)\\
&\quad\le-2\nabla^{\mathrm{T}}H(x)R\nabla H(x)+2\nabla^{\mathrm{T}}H(x)G_2\epsilon\\
&\quad-\nabla^{\mathrm{T}}H(x)G_1\left(\frac{1}{\gamma^2}G_1^{\mathrm{T}}+P^{\mathrm{T}}PG_1^{\mathrm{T}}\right)\nabla H(x)\\
&\quad=-2\nabla^{\mathrm{T}}H(x)R\nabla H(x)-\left\|\gamma\epsilon-\frac{1}{\gamma}G_2^{\mathrm{T}}\nabla H(x)\right\|^2\\
&\quad-\frac{1}{\gamma^2}\nabla^{\mathrm{T}}H(x)G_1G_1^{\mathrm{T}}\nabla H(x)+\left(\gamma^2\epsilon^{\mathrm{T}}\epsilon-z^{\mathrm{T}}z\right)\\
&\quad+\frac{1}{\gamma^2}\nabla^{\mathrm{T}}H(x)G_2G_2^{\mathrm{T}}\nabla H(x)\\
&\quad\le-\nabla^{\mathrm{T}}H(x)\left(2R+\frac{1}{\gamma^2}G_1G_1^{\mathrm{T}}-\frac{1}{\gamma^2}G_2G_2^{\mathrm{T}}\right)\nabla H(x)\\
&\quad+\left(\gamma^2\epsilon^{\mathrm{T}}\epsilon-z^{\mathrm{T}}z\right).
\end{aligned}\tag{35}$$

Taking (23) into account, it yields

$$\mathscr{L}V(x) \le \gamma^2 \epsilon^{\mathrm{T}} \epsilon - z^{\mathrm{T}} z. \tag{36}$$

Integrating (36) from 0 to T_0 leads to (18) which holds as $T_0 \to \infty$.

Next step we prove the closed-loop system where system (13) under the control law (24) is asymptotically stable in mean square when $\epsilon = 0$.

When $\epsilon = 0$, from (35), we can easily get that

$$\begin{aligned}
&\mathscr{L}V(x) \\
&\quad \le -2\nabla^{\mathrm{T}} H(x) R \nabla H(x) \\
&\qquad - \nabla^{\mathrm{T}} H(x) G_1 \left(\frac{1}{\gamma^2} G_1^{\mathrm{T}} + P^{\mathrm{T}} P G_1^{\mathrm{T}} \right) \nabla H(x) \\
&\quad = -\nabla^{\mathrm{T}} H(x) \left[2R + \frac{1}{\gamma^2} G_1 G_1^{\mathrm{T}} - \frac{1}{\gamma^2} G_2 G_2^{\mathrm{T}} \right] \nabla H(x) \\
&\qquad - \frac{1}{\gamma^2} \nabla^{\mathrm{T}} H(x) G_2 G_2^{\mathrm{T}} \nabla H(x) \\
&\qquad - \nabla^{\mathrm{T}} H(x) G_1 P^{\mathrm{T}} P G_1^{\mathrm{T}} \nabla H(x) \\
&\quad \le -\nabla^{\mathrm{T}} H(x) \left(\frac{1}{\gamma^2} G_2 G_2^{\mathrm{T}} + G_1 P^{\mathrm{T}} P G_1^{\mathrm{T}} \right) \nabla H(x).
\end{aligned} \tag{37}$$

Set

$$c_0 = \lambda_{\min} \left\{ \frac{1}{\gamma^2} G_2 G_2^{\mathrm{T}} + G_1 P^{\mathrm{T}} P G_1^{\mathrm{T}} \right\} > 0; \tag{38}$$

then we have

$$\mathscr{L}V(x) \le -c_0 \nabla^{\mathrm{T}} H(x) \nabla H(x). \tag{39}$$

Furthermore, owing to (A4) holding, there is

$$\mathscr{L}V(x) \le -c_0 \beta_1 \|x - x_0\|^2 \tag{40}$$

which implies

$$E\{\mathscr{L}V(x)\} \le -c_0 \beta_1 E\left\{\|x - x_0\|^2\right\}. \tag{41}$$

In addition, because of $E\{dw(t)\} = 0$, we further get

$$E\{dV(x)\} = E\{\mathscr{L}V(x)\}. \tag{42}$$

It is true that, for all $T > t_0$, $t_0 \in [-\tau, 0]$,

$$\begin{aligned}
E\{V(T)\} - E\{V(t_0)\} &= \int_{t_0}^{T} E\{\mathscr{L}V(s)\} ds \\
&\le \int_{t_0}^{T} E\left\{-c_0 \beta_1 \|x(s) - x_0\|^2\right\} ds.
\end{aligned} \tag{43}$$

Hence, one has

$$\frac{d}{dt} E\left\{\|x(T) - x_0\|^2\right\} \le -c_0 \beta_1 E\left\{\|x(T) - x_0\|^2\right\}. \tag{44}$$

From condition (A3), one has

$$\begin{aligned}
&\alpha_1 \|x - x_0\|^2 \le V(x) = 2\left(H(X) - H(X_0)\right), \\
&E\left\{\alpha_1 \|x(T) - x_0\|^2\right\} \le E\{V(T)\}, \\
&\frac{d}{dt} E\left\{\alpha_1 \|x(T) - x_0\|^2\right\} \le \frac{d}{dt} E\{V(T)\} \\
&\qquad \le c_0 \beta_1 E\left\{\|x(T) - x_0\|^2\right\}.
\end{aligned} \tag{45}$$

Set $c_1 = -c_0 \beta_1 / \alpha_1$; it follows that

$$\frac{d}{dt} E\left\{\|x(T) - x_0\|^2\right\} \le c_1 E\left\{\|x(T) - x_0\|^2\right\}. \tag{46}$$

Multiplying $e^{-c_1 T}$ to the two sides of inequality (44) yields

$$\begin{aligned}
&e^{-c_1 T} \frac{d}{dt} E\left\{\|x(T) - x_0\|^2\right\} - e^{-c_1 T} c_1 E\left\{\|x(T) - x_0\|^2\right\} \\
&\quad \le 0
\end{aligned} \tag{47}$$

which implies that

$$\frac{d}{dt}\left(e^{-c_1 T} E\left\{\|x(T) - x_0\|^2\right\}\right) \le 0. \tag{48}$$

Integrating inequality (48) from t_0 to T, we have

$$e^{-c_1 T} E\left\{\|x(T) - x_0\|^2\right\} - e^{-c_1 t_0} E\left\{\|x(t_0) - x_0\|^2\right\} \le 0; \tag{49}$$

that is,

$$E\left\{\|x(T) - x_0\|^2\right\} \le e^{c_1 (T - t_0)} E\left\{\|x(t_0) - x_0\|^2\right\}, \quad \forall T > t_0. \tag{50}$$

Due to $c_1 < 0$, there is

$$\lim_{T \to \infty} E\left\{\|x(T) - x_0\|^2\right\} = 0. \tag{51}$$

According to Definition 1, we can conclude that system (13) under the control law (24) is robustly asymptotically stable in mean square with respect to x_0. This completes the proof. □

Remark 5. $\nabla H(x_0) = 0$ and $\mathrm{Hess}(H(x_0)) > 0$ guarantee that the equilibrium x_0 is the minimal point of $H(x)$. Moreover, in view of conditions (A1)–(A4), there hold $\nabla V(x_0) = 0$ and $\mathrm{Hess}(V(x_0)) > 0$, which together with $V(x_0) = 0$ lead to the fact that $V(x)$ is a positive definite function in some neighborhood of equilibrium x_0.

Remark 6. Owing to the fact of $H(x) \in \mathrm{C}^2$, the solution of the closed-loop system (13) under the control law (24) is existent and unique on $[-\tau, \infty)$ for any initial data $\{x(t) = \phi(t) : t \in [-\tau, 0]\} \in \mathrm{C}^b_{\mathscr{F}_0}([-\tau, 0]; \mathscr{R}^n)$ in some neighborhood of equilibrium x_0.

3.2. N-Machine Power System. In this subsection, we consider the n-machine power system (3).

First, we can verify that

$$H(x)=\sum_{i=1}^{n}\left\{-P_{mi}\delta_i+G_{ii}\delta_iE_{qi}'^2+\frac{M_i}{2\omega_0}(\omega_i-\omega_0)^2+\frac{1+(x_{di}-x_{di}')B_{ii}+k_iT_{d0i}}{2(x_{di}-x_{di}')}E_{qi}'^2-\frac{\overline{u}_iT_{d0i}}{x_{di}-x_{di}'}E_{qi}'-\frac{1}{2}E_{qi}'\sum_{j=1,j\neq i}^{n}B_{ij}E_{qi}'\cos(\delta_i-\delta_j)\right\}\in C^2. \tag{52}$$

Choose the preassigned equilibrium

$$x_0=\left(\delta_i^{(0)},\omega_0,E_{qi}'^{(0)}\right),\quad i=1,2,\ldots,n \tag{53}$$

satisfying

$$\begin{aligned}\operatorname{Hess}(H(x_0))=\operatorname{Hess}\Bigg\{\sum_{i=1}^{n}\Bigg\{&-P_{mi}\delta_i^{(0)}+G_{ii}\delta_i^{(0)}\left(E_{qi}'^{(0)}\right)^2\\&+\frac{1+(x_{di}-x_{di}')B_{ii}+k_iT_{d0i}}{2(x_{di}-x_{di}')}\left(E_{qi}'^{(0)}\right)^2\\&-\frac{1}{2}E_{qi}'^{(0)}\sum_{j=1,j\neq i}^{n}B_{ij}E_{qi}'^{(0)}\cos\left(\delta_i^{(0)}-\delta_j^{(0)}\right)\\&-\frac{\overline{u}_iT_{d0i}}{x_{di}-x_{di}'}E_{qi}'^{(0)}\Bigg\}\Bigg\}>0\end{aligned} \tag{54}$$

and $\nabla^{\mathrm{T}}H(x)=0$; that is

$$\begin{aligned}&P_{mi}+G_{ii}E_{qi}'^{(0)}+E_{qi}'^{(0)}\sum_{j=1,j\neq i}^{n}B_{ij}E_{qi}'^{(0)}\sin\left(\delta_i^{(0)}-\delta_j^{(0)}\right)=0,\\&\frac{1+(x_{di}-x_{di}')B_{ii}+k_iT_{d0i}}{x_{di}-x_{di}'}E_{qi}'(0)-\frac{\overline{u}_iT_{d0i}}{x_{di}-x_{di}'}-\sum_{j=1,j\neq i}^{n}B_{ij}E_{qi}'(0)\cos(\delta_i(0)-\delta_j(0))+2G_{ii}\delta_i(0)E_{qi}'(0)=0.\end{aligned} \tag{55}$$

Meanwhile, we assume that there exist positive constants α_1,β_1 such that $H(x)-H(x_0)\geq(\alpha_1/2)\|x-x_0\|^2$ and $\nabla^{\mathrm{T}}H(x)\cdot\nabla H(x)\geq\beta_1\|x-x_0\|^2$ hold.

An H_∞ controller for system (3) is given in the following theorem.

Theorem 7. *Consider power system (3). If*

$$\begin{pmatrix}0&0&0\\0&\dfrac{2D_i\omega_0}{M_i^2}-\dfrac{1}{\gamma'^2}&0\\0&0&\dfrac{2(x_{di}-x_{di}')}{T_{d0i}'}+\dfrac{1}{\gamma'^2}\end{pmatrix}\geq0,\quad i=1,2,\ldots,n \tag{56}$$

hold, then the H_∞ control problem of system (3) can be solved by the feedback control law

$$\begin{aligned}u_{fi}(t-\tau)=&-2G_{ii}(x_{di}-x_{di}')\delta_i(t-\tau)E_{qi}'(t-\tau)-k_iT_{d0i}E_{qi}'(t-\tau)+T_{d0i}\overline{u}_i-\frac{1}{2}T_{d0i}\Bigg[\frac{1}{\gamma'^2}+p_i^2+\frac{(\omega_0^2+M_i^2)\varepsilon^2}{M_i^4}\Bigg]\Bigg[2G_{ii}\delta_i(t-\tau)E_{qi}'(t-\tau)\\&+\frac{1+(x_{di}-x_{di}')B_{ii}}{x_{di}-x_{di}'}E_{qi}'(t-\tau)-\sum_{j=1,j\neq i}^{n}B_{ij}E_{qi}'(t-\tau)\cos(\delta_i(t-\tau)-\delta_j(t-\tau))\\&+\frac{(k_iE_{qi}'(t-\tau)-\overline{u}_i)T_{d0i}}{x_{di}-x_{di}'}\Bigg]+2G_{ii}(x_{di}-x_{di}')\delta_i(t-\tau)E_{qj}'(t-\tau)+T_{d0i}k_iE_{qj}'(t-\tau)-(m_i+2n_i)\\&-\frac{1}{2}\tau\lambda_1\lambda_2t_i-\frac{1}{4}\tau\lambda_1\lambda_2\frac{(\omega_0^2+M_i^2)\varepsilon^2}{M_i^4},\quad i=1,2,\ldots,n,\end{aligned} \tag{57}$$

where $\lambda_1,\lambda_2,m_i,n_i$, and t_i, are constants, which satisfy

$$\begin{aligned}&\lambda_1\geq\sup_{t\geq-\tau}\Bigg\|Hess\Bigg[\frac{1+(x_{di}-x_{di}')B_{ii}+k_iT_{d0i}}{2(x_{di}-x_{di}')}E_{qi}'^2+G_{ii}\delta_iE_{qi}'^2-\frac{\overline{u}_iT_{d0i}}{x_{di}-x_{di}'}E_{qi}'-P_{mi}\delta_i-E_{qi}'\sum_{j=1,j\neq i}^{n}B_{ij}E_{qj}'\cos(\delta_i-\delta_j)\Bigg]\Bigg\|,\\&\lambda_2\geq\sup_{t\geq-\tau}\left\|\left(\dot\delta_i(t)\quad\dot\omega_i(t)\quad\dot E_{qi}'(t)\right)\right\|,\\&m_i\geq\left|k_iE_{qi}'\right|,\end{aligned}$$

$$n_i \geq \left| \frac{2G_{ii}\left(x_{di} - x'_{di}\right)\delta_i(t-\tau)E'_{qi}(t-\tau)}{T_{d0i}} \right|,$$
$$t_i \geq \left| \frac{1}{\gamma'^2} + p_i^2 \right|, \quad i = 1, 2, \ldots, n. \tag{58}$$

$(\delta_i(t), \omega_i(t), E'_{qi}(t))$, $i = 1, 2, \ldots, n$, is the solution of the closed-loop system at time t under initial condition.

Proof. Taking

$$y_i = \begin{pmatrix} \dfrac{M_i}{\omega_0}\left(\omega_i(t) - \omega_0\right) \\ 2G_{ii}\delta_i(t)E'_{qi}(t) + \dfrac{1 + \left(x_{di} - x'_{di}\right)B_{ii}}{x_{di} - x'_{di}}E'_{qi}(t) \end{pmatrix} - \begin{pmatrix} 0 \\ \displaystyle\sum_{j=1, j\neq i}^{n} B_{ij}E'_{qj}(t)\cos\left(\delta_i(t) - \delta_j(t)\right) \end{pmatrix} + \begin{pmatrix} 0 \\ \dfrac{\left(k_i E'_{qi}(t) - \overline{u}_i\right)T_{d0i}}{x_{di} - x'_{di}} \end{pmatrix},$$
$$z_i = p_i \left[2G_{ii}\delta_i(t)E'_{qi}(t) - \sum_{j=1, j\neq i}^{n} B_{ij}E'_{qj}(t)\cos\left(\delta_i(t) - \delta_j(t)\right) + \frac{1 + \left(x_{di} - x'_{di}\right)B_{ii}}{x_{di} - x'_{di}}E'_{qi}(t) + \frac{\left(k_i E'_{qi}(t) - \overline{u}_i\right)T_{d0i}}{x_{di} - x'_{di}} \right] \tag{59}$$

into consideration, then we can prove the result using the similar method in the proof of Theorem 4, where $p_i \geq 0$, $i = 1, 2, \ldots, n$ are the weighting constants. □

4. Illustrative Examples

To show the effectiveness of the proposed control strategy, we give a two-machine power system as shown in Figure 1. The generators G_1, G_2 are assumed to be connected to distant power systems and disturbed by random fluctuation. In simulating, a temporary short-circuit fault occurs at point K (see Figure 1) during the time 0.5 sec~1 sec. The system parameters used in this simulation are given in Table 1. Choose $\omega_0 = 1$, $\xi = 1$.

Taking the above parameters, system (3) can be expressed as

$$\begin{aligned} d\delta_1 &= \left(\omega_1 - 1\right)dt, \\ d\omega_1 &= \left(\frac{6}{8} - \frac{5}{8}\omega_1 - \frac{1}{8}P_{e1} + \epsilon_{11}\right)dt \\ &\quad + 0.125\left(\omega_1 - 1\right)dw(t), \\ dE'_{q1} &= \left[-\frac{1}{5}E_{q1} + \frac{1}{5}u_{f1}(t-\tau) + \epsilon_{12}\right]dt, \\ d\delta_2 &= \left(\omega_2 - 1\right)dt, \\ d\omega_2 &= \left(\frac{7}{9} - \frac{6}{9}\omega_2 - \frac{1}{9}P_{e2} + \epsilon_{21}\right)dt \\ &\quad + \frac{1}{9}\left(\omega_2 - 1\right)dw(t), \\ dE'_{q2} &= \left[-\frac{1}{6}E_{q2} + \frac{1}{6}u_{f2}(t-\tau) + \epsilon_{22}\right]dt, \end{aligned} \tag{60}$$

$$\begin{aligned} E_{q1} &= E'_{q1} + 0.5I_{d1}, \\ P_{e1} &= E'^2_{q1} + E'_{q1}\sum_{j=1, j\neq i}^{2} E'_{qj}\sin\left(\delta_i - \delta_j\right), \\ I_{d1} &= 4E'_{q1} - \sum_{j=1, j\neq i}^{2} E'_{qj}\cos\left(\delta_i - \delta_j\right), \\ E_{q2} &= E'_{q2} + 0.6I_{d2}, \\ P_{e2} &= E'^2_{q2} + E'_{q2}\sum_{j=1, j\neq i}^{2} E'_{qj}\sin\left(\delta_i - \delta_j\right), \\ I_{d2} &= 10E'_{q2} - \sum_{j=1, j\neq i}^{2} E'_{qj}\cos\left(\delta_i - \delta_j\right). \end{aligned} \tag{61}$$

Table 1: Generators' data (all per unit except M_i, T_{d0i}, $i = 1, 2, \ldots, n$ in seconds).

M_1	P_{m1}	D_1	x_{d1}	x'_{d1}	T_{d01}
8	1	5	1	0.5	5
M_2	P_{m2}	D_2	x_{d2}	x'_{d2}	T_{d02}
9	1	6	1	0.4	6
B_{11}	B_{12}	B_{21}	B_{22}	G_{11}	G_{22}
4	1	1	10	1	1

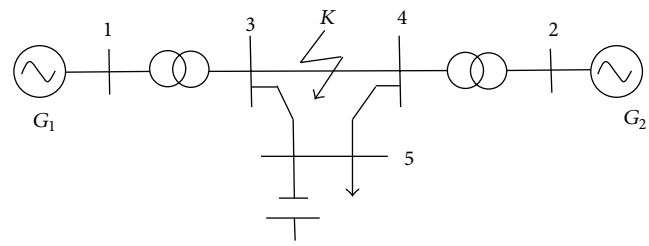


Figure 1: Two-machine power system.

Choosing the following preassigned operating point

$$\left(\delta_1^{(0)}, \omega_0, E'^{(0)}_{q1}, \delta_2^{(0)}, \omega_0, E'^{(0)}_{q2}\right) = \begin{bmatrix} 0.5 & 1 & 1 & 0.5 & 1 & 1 \end{bmatrix}, \tag{62}$$

then $\overline{u}_1 = 0.5$, $\overline{u}_2 = 16/15$, and $k_1 = k_2 = -0.1$.

It is easy to verify that system (60) with the above values satisfies conditions (A1)–(A4) of Theorem 4.

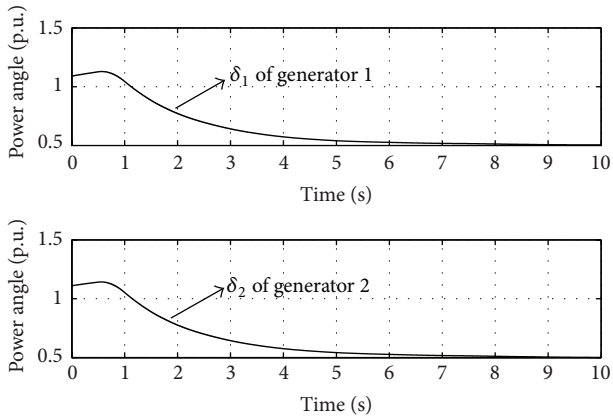


Figure 2: Power angle dynamic behavior while $\tau = 0.05$ s.

The fault indicates a unit step function; that is, $\epsilon_{11} = \epsilon_{12} = \epsilon_{21} = \epsilon_{22} = -1(t-0.5) + 1(t-1)$. For given $\gamma' = 4$, we can find $p_1 = p_2 = 1$ such that inequality (56) is satisfied.

We will test the effectiveness of the proposed control configuration at two different time delays $\tau = 0.5$ s and $\tau = 0.05$ s. The initial condition is $(\delta_1(0), \omega_1(0), E'_{q1}(0), \delta_2(0), \omega_2(0), E'_{q2}(0)) = [1.2 \;\; 1 \;\; 2 \;\; 1.2 \;\; 1 \;\; 2]$.

Take $\mu = 1/8^4$, $\lambda' = 40$, $m'_{11} = m'_{12} = 100$, $m'_{21} = m'_{22} = 100$, and $m'_3 = 40$. According to Theorem 7 proposed in this paper, system (60) is asymptotically stable in mean square for all $\tau \geq 0$ and $\epsilon = 0$ under the feedback control law

$$\begin{aligned}
u_{f1}(t-\tau) &= (4096)^{-1}\Big[-29858.5\delta_1(t-\tau)E'_{q1}(t-\tau) \\
&\quad - 62358.25E'_{q1}(t-\tau) + 12881.25E'_{q2}(t-\tau) \\
&\quad \cdot \cos\big(\delta_1(t-\tau) - \delta_2(t-\tau)\big) - 3202000\tau \\
&\quad \cdot \operatorname{sgn}(5\omega_1 - 5) + 74646.25\Big] - 4480\tau \operatorname{sgn}\Big[5E'_{q1}(t) \\
&\quad + 2\delta_1(t)E'_{q1}(t) - E'_{q2}(t)\cos\big(\delta_1(t) - \delta_2(t)\big) - 5\Big], \\
u_{f2}(t-\tau) &= (4096)^{-1}\Big[-35011\delta_2(t-\tau)E'_{q2}(t-\tau) \\
&\quad - 162422.4E'_{q2}(t-\tau) + 17915.1E'_{q1}(t-\tau) \\
&\quad \cdot \cos\big(\delta_1(t-\tau) - \delta_2(t-\tau)\big) - 3842400\tau \\
&\quad \cdot \operatorname{sgn}(6\omega_2 - 6) + 191094.4\Big] - 5376\tau \\
&\quad \cdot \operatorname{sgn}\Big[\Big(\frac{2}{3} + 10\Big)E'_{q2}(t) + 2\delta_2(t)E'_{q2}(t) \\
&\quad - E'_{q1}(t)\cos\big(\delta_1(t) - \delta_2(t)\big) - \frac{64}{6}\Big].
\end{aligned} \tag{63}$$

Simulations with the above initial condition and the delay $\tau = 0.5$ s and $\tau = 0.05$ s are given in Figures 2–4 and Figures 5–7, separately. Through Figures 2–7, we can see that the states of the system converge to the equilibrium $(\delta_1(0), \omega_1(0), E'_{q1}(0), \delta_2(0), \omega_2(0), E'_{q2}(0)) = [1.2 \;\; 1 \;\; 2 \;\; 1.2 \;\; 1 \;\; 2]$ eventually. Obviously, under the delayed feedback controller by using the proposed method, the robustness of the closed-loop system is guaranteed. It is also seen that the controller possesses insensitivity in regard to the types of time delay and stochastic disturbances.

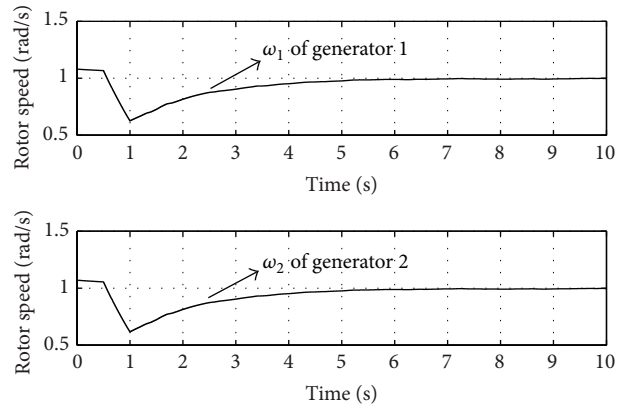


Figure 3: Rotor speed dynamic behavior while $\tau = 0.05$ s.

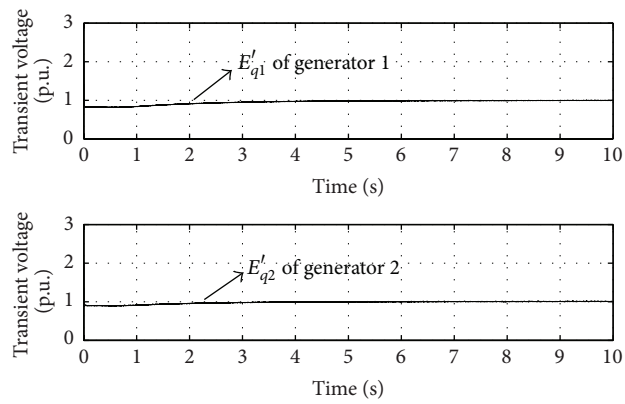


Figure 4: Transient voltage dynamic behavior while $\tau = 0.05$ s.

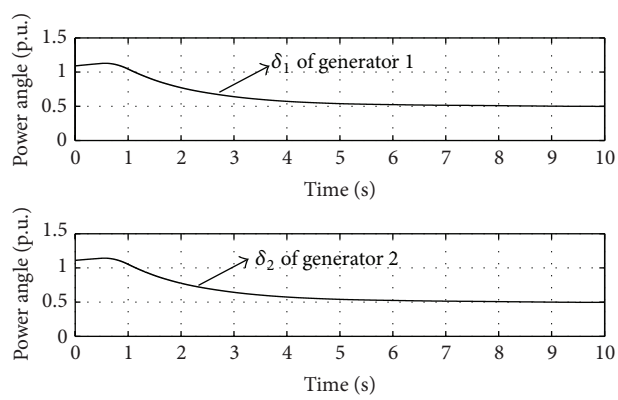


Figure 5: Power angle dynamic behavior while $\tau = 0.5$ s.

5. Conclusion

This paper studied the H_∞ excitation controller design problem of a class of stochastic power systems with time-delay and

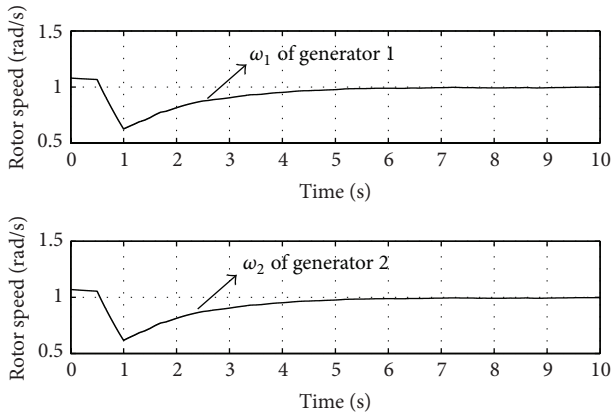


Figure 6: Rotor speed dynamic behavior while $\tau = 0.5$ s.

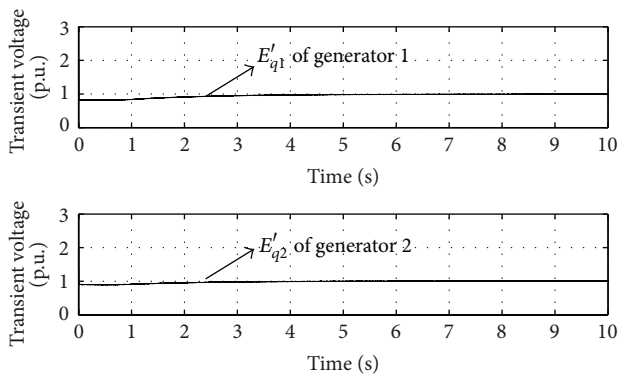


Figure 7: Transient voltage dynamic behavior while $\tau = 0.5$ s.

disturbances. In the design process, we used the prefeedback technique, Newton-Leibniz formula, and a few properties of norm. Besides, we obtain these results by nonlinear Hamilton function approach due to the special structural properties of the Hamiltonian systems. We also give a two-machine power system simulation and it shows that the results achieved in this paper are practicable in analyzing the H_∞ excitation control problem of stochastic power system in consideration of time-delay and disturbances.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported by the National Natural Science Foundation of China under Grant 61203013.

References

[1] Q. Lu and Y. Z. Sun, "Nonlinear stabilizing control of multimachine systems," *IEEE Transactions on Power Systems*, vol. 4, no. 1, pp. 236–241, 1989.

[2] D. G. Padhan and S. Majhi, "A new control scheme for PID load frequency controller of single-area and multi-area power systems," *ISA Transactions*, vol. 52, no. 2, pp. 242–251, 2013.

[3] Y. J. Peng, J. Zeng, and F. Q. Deng, "A numeric method for stochastic transient stability analysis of excitation control system," *Proceedings of the Chinese Society of Electrical Engineering*, vol. 31, no. 19, pp. 60–66, 2011.

[4] T. Li and Y. Zhang, "Fault detection and diagnosis for stochastic systems via output PDFs," *Journal of the Franklin Institute*, vol. 348, no. 6, pp. 1140–1152, 2011.

[5] S. P. Wen, Z. G. Zeng, T. W. Huang, and G. Bao, "Observer-based H_∞ control of a class of mixed delay systems with random data losses and stochastic nonlinearities," *ISA Transactions*, vol. 52, no. 2, pp. 207–214, 2013.

[6] F. Liu, R. Yokoyama, Y. Zhou, and M. Wu, "Free-weighting matrices-based robust wide-area FACTS control design with considering signal time delay for stability enhancement of power systems," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 7, no. 1, pp. 31–39, 2012.

[7] C.-X. Dou, X.-Z. Zhang, S.-L. Guo, and C.-C. Mao, "Delay-independent excitation control for uncertain large power systems using wide-area measurement signals," *International Journal of Electrical Power & Energy Systems*, vol. 32, no. 3, pp. 210–217, 2010.

[8] P. V. Swaroop, I. Erlich, K. Rohrig, and J. Dobschinski, "A stochastic model for the optimal operation of a wind-thermal power system," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 940–950, 2009.

[9] M. Galaz, R. Ortega, A. S. Bazanella, and A. M. Stankovic, "An energy-shaping approach to the design of excitation control of synchronous generators," *Automatica*, vol. 39, no. 1, pp. 111–119, 2003.

[10] S. Fiaz, D. Zonetti, R. Ortega, J. M. A. Scherpen, and A. J. van der Schaft, "A port-Hamiltonian approach to power network modeling and analysis," *European Journal of Control*, vol. 19, no. 6, pp. 477–485, 2013.

[11] Y. Wang, D. Cheng, Y. Liu, and C. Li, "Adaptive H_∞ excitation control of multimachine power systems via the Hamiltonian function method," *International Journal of Control*, vol. 77, no. 4, pp. 336–350, 2004.

[12] Y. Z. Wang, G. Feng, D. Z. Cheng, and Y. Liu, "Adaptive L_2 disturbance attenuation control of multi-machine power systems with SMES units," *Automatica*, vol. 42, no. 7, pp. 1121–1132, 2006.

[13] Z. Xi and J. Lam, "Stabilization of generalized Hamiltonian systems with internally generated energy and applications to power systems," *Nonlinear Analysis: Real World Applications*, vol. 9, no. 3, pp. 1202–1223, 2008.

[14] B. Maschke and A. J. van der Schaft, "Port-controlled Hamiltonian systems: modeling origins and system theoretic properties," in *Proceedings of the IFAC Symposium on Nonlinear Control Systems (NOLCOS '92)*, pp. 282–288, Bordeaux, France, 1992.

[15] W.-W. Sun, Z. Lin, and Y.-Z. Wang, "Global asymptotic and finite-gain L_2 stabilisation of port-controlled Hamiltonian systems subject to actuator saturation," *International Journal of Modelling, Identification and Control*, vol. 12, no. 3, pp. 304–310, 2011.

[16] W. W. Sun, "Stabilization analysis of time-delay Hamiltonian systems in the presence of saturation," *Applied Mathematics and Computation*, vol. 217, no. 23, pp. 9625–9634, 2011.

[17] R. M. Yang and Y. Z. Wang, "Stability for a class of nonlinear time-delay systems via Hamiltonian functional method," *Science China: Information Sciences*, vol. 55, no. 5, pp. 1218–1228, 2012.

[18] R. M. Yang and Y. Z. Wang, "Finite-time stability analysis and H_∞ control for a class of nonlinear time-delay Hamiltonian systems," *Automatica*, vol. 49, no. 2, pp. 390–491, 2012.

[19] L. H. Peng and W. W. Sun, "H_∞ excitation control of multimachine power systems considering delay and disturbances," in *Proceedings of the 32nd Chinese Control Conference (CCC '13)*, pp. 679–684, Xian, China, July 2013.

[20] Q. Lu, Y. Sun, Z. Xu, and T. Mochizuki, "Decentralized nonlinear optimal excitation control," *IEEE Transactions on Power Systems*, vol. 11, no. 4, pp. 1957–1962, 1996.

[21] S.-J. Liu, S. S. Ge, and J.-F. Zhang, "Adaptive output-feedback control for a class of uncertain stochastic non-linear systems with time delays," *International Journal of Control*, vol. 81, no. 8, pp. 1210–1220, 2008.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 628583, 13 pages
http://dx.doi.org/10.1155/2015/628583

Research Article

Cascade Probability Control to Mitigate Bufferbloat under Multiple Real-World TCP Stacks

Hoang-Linh To,[1] Thuyen Minh Thi,[1] and Won-Joo Hwang[2]

[1] *Department of Information and Communication System, Inje University, Gimhae, Gyeongnam 621-749, Republic of Korea*
[2] *Department of Information and Communications Engineering, Inje University, Gimhae, Gyeongnam 621-749, Republic of Korea*

Correspondence should be addressed to Won-Joo Hwang; ichwang@inje.ac.kr

Received 22 June 2015; Accepted 13 August 2015

Academic Editor: Rongwei Guo

Persistently full buffer problem, commonly known as bufferbloat, causes unnecessary additional latency and throughput degradation whenever congestion happens in Internet. Several proposed queue management schemes, with the debloat mission, are almost based on the modification of one-loop feedback control where the instability and bad transient behavior are still big challenges. In this paper, we present a cascade probability control scheme using margin optimal method to address such challenges under different kinds of real-world TCP stacks. Simulation results guarantee the measured round trip time tracking to a low value of delay (e.g., ≈180 ms under TCP Reno, and ≈130 ms under TCP Cubic) and ≈50% delay reduction in comparison to current deployed queue management schemes in network devices.

1. Introduction

Nowadays, interactive and delay-sensitive applications such as VoIP, teleconference, and gaming often perform more poorly than before. From the economic perspective, as the cost of memory has decreased over the past few years, memory with large capacity has been put into network devices such as routers. From the engineering perspective, as traditional analysis, a larger buffer results in less loss rate under congestion. However, recent studies about *bufferbloat* reveal some bad effects of a large buffer. It might destroy interactivity of transport control protocols (TCP) under load and often results in higher latency and lower throughput. Nagle [1] firstly drew attention to effects of infinite buffers in packet-switching networks. Then *bufferbloat* termed by Gettys [2] firstly opens a research field to seriously reconsider the problem of large buffers and hidden "dark" buffers which can appear everywhere in Internet devices.

Conceptually, bufferbloat is a phenomenon firstly realized in packet-switched networks, where excess buffering of packets might cause high latency and jitters, as well as reducing the overall network throughput. When a router device is equipped with a large buffer, it can become practically unusable for many delay-sensitive applications like voice calls, chat, or even web surfing. Traditional rule-of-thumb for setting buffer size in an Internet route is based on bandwidth-delay product (BDP), the product of a data link's capacity and its round-trip time (RTT). The basic premise is that modern networking kit contains too much buffer memory. In [2], it is necessary to restrict bandwidth to improve latency and ping times. Despite this fundamental of sacrificing bandwidth for latency, most users keep asking how to fill their bandwidth quota and get good VoIP/games satisfaction. Long time ago and until now, researchers nonexhaustively attempt to reduce dropping packets by adding more and more buffering in routers. However, each dropped packet is essential for well-behaved functioning of TCP in congestion.

Let us examine an interaction situation between TCP and bufferbloat [3]. TCP relies on timely congestion notification to adjust its transmission rate to the available bandwidth. And bufferbloat means that new arriving packets are continued to be buffered, instead of dropped due to large buffer size. It causes the queue to build up longer at the bottlenecks. Entering the dropping state, more packets are dropped than necessary which would shrink the transmission rates of the TCP sender. In particular, if several TCP applications are

transmitting over the same congestion point, all flows will see drops at the same time. Therefore, all transmission rates would be reduced simultaneously, called TCP global synchronization. Certainly, suitable amount of buffering is helpful to improve the efficiency of data transmissions to smoothen bursts transmissions. Not dropping packets early enough, however, leads to increasing delays for interactive real-time communication since widely deployed congestion control algorithms only rely on packet loss as a signal for congestion. Not dropping packets in a timely fashion also prevents TCP applications from reacting properly to overload.

Solutions to bufferbloat problem or debloat include queue management algorithms, which attempt to manage queue occupancy in passive or active ways. Passive queue management (PQM) such as Drop Tail is the most currently deployed queue management algorithm in Internet router devices. It drops packets when buffer queue is overflow, so because of this simplicity, it is widely used now. The weaknesses of PQM such as lock-out [4] and global synchronization [5] motivate to urgent needs of active queue management schemes (AQMs) deployment. Several AQMs have been deeply studied in the recent decade. The most popular one is random early detection (RED) [5] and its variants [6, 7] with the main idea of using a probabilistic approach to randomly dropping packets for congestion resolution. However, most of them require careful tuning for varying network conditions; otherwise they only work well under a few specific scenarios with default parameters [8]. Game theory approach, which is mainly dedicated to users (user-aware), has also been investigated to tackle this issue [9, 10]. Recently, a new AQM named controlled delay management (CoDel) [11] has been proposed to overcome weakness. Parameterless and easy deployment are two among strong points of CoDel. Even though, for larger RTTs and smaller bandwidths, CoDel has poor link utilization than RED and its variants, the next drop time of CoDel algorithm is derived by using the multiplicative decrease of square root of total number dropped packet counting, which needs more in-depth investigation for improvement.

In this paper, we revisit AQM design problem from control theoretical perspective but consider cascade control to further optimize performance in user application level while considering bufferbloat phenomenon. Control theory is one of the most efficient tools for AQM to bring a better system stability due to well-developed control theory background. Several efforts have been recently put into this direction for RED [12], PID-AQM [13, 14], and controller design in state space [15, 16] or queuing modelling with the impatient customer feature [17] and so forth. Based on the fluid-flow approximation model for additive increase-multiplicative decrease (AIMD) phase of TCP, they converted dynamic equations of window size and queue length into system transfer function or state space by linearization methods. Then they designed the specific controller according to a closed-loop transfer function of the whole TCP system. We realize that these models almost design one-loop control for queue length only which creates some problem of difficult adjusting controller parameters and stability guarantee. One more loop with cascade design in control theory would improve their performance, especially the bad transient behaviour of current debloat schemes. So we consider a two-loop control scheme for AQM, which is mainly dedicated to bufferbloat issue. Dividing into an inner and an outer loop, the inner one adjusts window size based on changing of traffic and feedback window size at time t. The outer one's mission is to adjust queue length based on feedback queue length value at time t. Each inner loop and outer loop are designed using two transfer functions which decomposed from the fluid-flow model and, therefore, have different controllers. One difficulty when considering this cascade design is the interaction in time-scale between two loops. We see that the inner loop operates in the transport layer while the outer one operates in link layer (faster than transport layer). The main motivation is that solving bufferbloat (large buffer) problem cannot be done without considering different network layers simultaneously. Therefore, one-loop control methods like Drop Tail, RED, or classical PID should not be adequate for bufferbloat mitigation. Our proposed CPC, with an inner and an outer loop, acts in both the transport layer (adjusting window size) and the link layer (adjusting queue length). To the best of our knowledge, our work is the first attempt to adapt cascade control method to bufferbloat research field. The weakness of this method is that more complexity is added because of two additional controllers. However, better performance results (shorter queuing delay and larger goodput) are achieved and we are going to see more details in Section 5. The performance metrics we evaluate consist of measured round-trip time, queue length at each instantaneous time, and goodput at TCP application layer which represents the users' (clients) satisfaction level. The primary contributions are summarized as follows:

(i) We propose a cascade probability control (CPC) which has two control loops. The inner loop gets information from current window size and capacity of link while the outer loop is based on difference between average measured queue length and queue length reference value.

(ii) Section 4 presents our proposed CPC controllers which are designed using the optimal margin method in frequency domain. The fast transient and stability in a wide frequency range can be achieved using this method.

(iii) We develop our own model to simulate CPC performance using the popular open-source software, OMNeT++. We also compare our proposal with Drop Tail and RED under three popular TCP stacks (e.g., Reno, Cubic, and FreeBSD.) using the real-world stack module network simulation cradle (NSC). Simulation results demonstrate that bufferbloat delay can be reduced significantly (≈50% compared to Drop Tail and RED) and well controlled by CPC (Section 5).

2. Related Works

Several approaches to AQM, using control theory with the core linearized TCP model by Hollot et al. [12], were proposed

and contributed a large portion to debloat field [18]. The original goal was to propose more concrete design guidelines for the RED parameters to improve stability and responsiveness; however, they also discover clearer understanding of RED's behaviour with changing of network conditions such as round-trip times, offer load as number of TCP flows, and link capacities. Focusing on control theory viewpoint, we categorize them into three types: classical control, robust control, and fuzzy logic control.

Classical PID controller based algorithms were designed as alternative AQM solutions to meet various specifications of the Internet using feedback control. Hollot et al. [12] analysed RED as I-controller and proposed two types, the proportional (P) controller and the proportional integral (PI) controller, for improving RED. The stable region of the control gain is determined for dynamic-RED (DRED) using the Routh stability test. In [19], dynamic-RED (DRED) was proposed, using a load-dependent probability to randomly discard packets when a buffer becomes congested. It maintains the average queue size close to a predetermined threshold but allows transient traffic bursts to be queued without unnecessary packet drops. Its main advantage is that we do not have to collect state information of individual flows. In [15], a feedback control model with PI controller has been recently proposed to improve link performance in wired communication networks.

Robust control approach was also studied to improve classical control. The issue of large delay with large buffer in bufferbloat was addressed by DC-AQM algorithm based on internal mode compensation (IMC) principle, which is an example of robust control approach [20]. Using IMC controller derived, they tried to tune parameters K_p, K_i, and K_d of classical PID controller to reduce delay. To also contend with large delay, gain adaptive Smith predictor with PI controller (GAS-PI) in [21] was built to improve robustness. Then, in [13], a predictive PID controller is proposed for TCP/AQM. They used the generalized predictive control method to determine suitable values for K_p, K_i, and K_d so that they made the system more robust to changes in model parameters such as offered load and round-trip time.

Fuzzy control RED (FCRED) was proposed in [6]. It consists of a fuzzy controller adjusting the $P_{\max}$ parameter of the RED algorithm. The fuzzy controller includes three parts: the fuzzification unit followed by the fuzzy-interference engine with fuzzy-rule base and finally a defuzzification unit. The fuzzification module maps the input values to be controlled to a fuzzy set (i.e., membership functions). The fuzzy rule base provides the connection between the input signals and the appropriate output variable. Fuzzy logic rules are constructed based on trial-and-error, which needs the knowledge and experience of domain-expert in TCP congestion control. Moreover, fuzzy logic-based AQM schemes are sometimes not distributed and hard to implement.

Briefly, it can be seen that, as the years progress, the main direction has been to more and more sophisticated robust control techniques, combined with some classical techniques in control theory as well. In many cases, the linear TCP model in [12] continues to be at the core of control theoretic AQM algorithms. Cascade control, however, has not been considered yet in development of AQM. In this paper, we contribute to the AQM debloat research trends by using cascade control to address large delay bufferbloat issue. Our approach uses two controllers which are connected in cascade style. Our CPC results obtained from theoretical Matlab or simulation using OMNeT++ and the integrated NSC tool are so promising in fast transient behavior and stability.

3. System Model

The TCP/AQM fluid-flow model described by nonlinear differential equations has been extensively studied in particular AQM routers interacting with TCP sources (e.g., [12, 13, 16, 20]). Until now, they can capture the additive increase-multiplicative decrease (AIMD) feature from TCP [22], without slow start and time-out mechanisms. However, this lacking only affects initial start-up of the system. Once the system reaches the stable point, the differential equations solver is able to track changes in the network well [23]:

$$\begin{aligned} \dot{w}(t) &= \frac{1}{R(t)} - \frac{w(t)}{2}\frac{w(t-\tau)}{R(t-\tau)}p(t-\tau); \\ \dot{q}(t) &= N\frac{w(t)}{R(t)} - C_l, \end{aligned} \tag{1}$$

where

(i) C_l is the transmission capacity of link l (packets/sec);

(ii) R is the round-trip time (sec); $R(t) = T_p + q(t)/C_l$ with T_p being the fixed propagation delay;

(iii) N is the number of TCP flows;

(iv) τ is feedback delay (sec).

The operating point (w_0, q_0, p_0) of TCP model (1) can be derived at $\dot{w} = 0$ and $\dot{q} = 0$ as follows:

$$\begin{aligned} w_0^2 p_0 &= 2, \\ w_0 &= \frac{R_0 C_l}{N}, \\ R_0 &= T_p + \frac{q_0}{C_l}. \end{aligned} \tag{2}$$

Doing linearization of the above TCP queue model around operating point (w_0, q_0, p_0) (details in [12]), with $\delta\dot{w} = w - w_0$, we have a linearized small signal model of TCP/AQM:

$$\begin{aligned} \delta\dot{w}(t) &= -\delta w(t)\frac{2N}{R_0^2 C_l} - \delta p(t - R_0)\frac{R_0 C_l^2}{2N^2}, \\ \delta\dot{q}(t) &= \delta w(t)\frac{N}{R_0} - \delta q(t)\frac{1}{R_0}. \end{aligned} \tag{3}$$

Convert (3) using Laplace transformation as follows:

$$\begin{aligned} s \times w(s) &= -w(s)\frac{2N}{R_0^2 C_l} - p(s)e^{-sR_0} \cdot \frac{R_0 C_l^2}{2N^2}, \\ s \times q(s) &= w(s)\frac{N}{R_0} - q(s)\frac{1}{R_0}. \end{aligned} \tag{4}$$

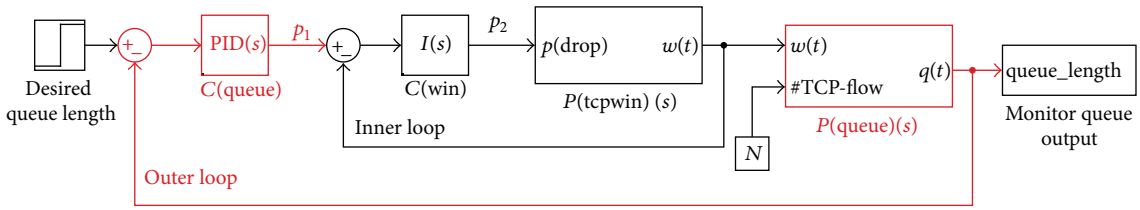


Figure 1: Block diagram: proposed cascade probability control scheme.

So the TCP/AQM system transfer functions in Laplace domain are

$$P(s) = P_{\text{tcpwin}}(s) \cdot P_{\text{queue}}(s) \cdot e^{-sR_0} = \left[\frac{w(s)}{p(s)}\right] \cdot \left[\frac{q(s)}{w(s)}\right] = \left[\frac{A}{s+B}\right] \cdot \left[\frac{C}{s+D}\right] e^{-sR_0}, \tag{5}$$

where $A = R_0 C_l^2/(2N^2)$; $B = 2N/(R_0^2 C_l)$; $C = N/R_0$; and $D = 1/R_0$.

4. Cascade Probability Control (CPC)

In literature, single loop control for AQM was often studied due to less complexity and low oscillation but did not behave well in case of disturbances (e.g., bandwidth fluctuation and bursty traffic) or bad transient. Cascade control comes to rescue and achieves fast rejection of disturbance before it affects the main system model. In this section, we propose a cascade probability control (CPC) to improve dynamic of the open-loop system transfer function (5) which consists of two subsystems, window size and queue length control. The inner loop adjusts the window size by an inner controller C_{win}, while the outer loop receives queue length information at time t and uses an outer controller C_{queue} to maintain $q(t)$ (Figure 1). The final goal is to reach a stable value of queue length, so that bufferbloat phenomenon can be mitigated. Moreover, the margin optimal method for controller design also brings a fast transient behavior to our system.

4.1. CPC Controllers. We design two PI controllers for simplicity and for reducing number of parameters: the inner controls dropping probability p_2 based on traffic information and the outer controls dropping probability p_1 based on difference between measured average queue length and reference queue length. We outline our proposed design framework in Figure 1 and would compete with the adaptive weight PID approach in [24].

4.1.1. Inner Loop. An important design requirement is that the inner loop controller should behave quickly. From (5), the inner control objective is a linear first-order type: $P_{\text{tcpwin}}(s) = k_I/(1+Ts)$, where $k_I = A/B$ and $T = 1/B$. Hence, we design an integral I-controller for inner loop:

$$C_{\text{win}(s)} = \frac{1}{T_{C_{\text{win}}} s}. \tag{6}$$

(i) The close-loop transfer function of the inner loop is

$$I(s) = \frac{w(s)}{p_1(s)} = \frac{P_{\text{tcpwin}}(s) C_{\text{win}(s)}}{1 + P_{\text{tcpwin}}(s) C_{\text{win}(s)}} = \frac{k_I}{T_{C_{\text{win}}} s(1+Ts) + k_I}. \tag{7}$$

(ii) The $I(s)$ transfer function is converted into frequency domain, with ω being frequency:

$$|I(j\omega)| = \frac{k_I}{\sqrt{\left(k_I - T_{C_{\text{win}}} T \cdot \omega^2\right)^2 + \left(\omega \cdot T_{C_{\text{win}}}\right)^2}}, \quad \Longleftrightarrow I(j\omega)^2 = \frac{k_I^2}{k_I^2 + \left(T_{C_{\text{win}}}^2 - 2k_I \cdot T_{C_{\text{win}}} T\right)\omega^2 + T_{C_{\text{win}}}^2 \cdot T^2\omega^4}. \tag{8}$$

One of quality requirements to closed-loop control system, which is represented by $I(s)$, is that the output is the same as the input signal or the controller C_{win} should bring $|I(j\omega)| = 1$, $\forall\omega$, which can be called margin optimal method. However, due to several reasons of real system, that requirement is rarely satisfied for all frequencies ω. An acceptable design is that $|I(j\omega)| \approx 1$, in a wide band of low frequencies ω. Hence, we propose choosing $T_{C_{\text{win}}}$ such that $T_{C_{\text{win}}}^2 - 2k_I \cdot T_{C_{\text{win}}} T = 0$ or $T_{C_{\text{win}}} = 2k_I T$. This close-form expression of $T_{C_{\text{win}}}$ is used to make decision for controller C_{win}.

4.1.2. Outer Loop

(i) The close-loop transfer function of the outer loop is

$$O(s) = \frac{q(s)}{q_{\text{ref}}} = \frac{I(s) P_{\text{queue}(s)} C_{\text{queue}}(s)}{1 + I(s) P_{\text{queue}(s)} C_{\text{queue}}(s)}. \tag{9}$$

(ii) The outer control objective into zero-pole form is

$$I(s) P_{\text{queue}(s)} = \frac{k_I}{T_{C_{\text{win}}} s(1+Ts) + k_I} \cdot \frac{C}{s+D} = \frac{k_O}{\left(1+T_{1_O} s\right)\left(1+T_{2_O} s\right)\left(1+T_{3_O} s\right)}, \tag{10}$$

where $k_O = C/D$; $T_{1_O}T_{2_O} = TT_{C_{\text{win}}}/k_I$; $T_{1_O} + T_{2_O} = T_{C_{\text{win}}}/k_I$; $T_{3_O} = 1/D$.

For the outer loop, the objective function is linear third-order type, due to inclusion of $I(s)$. Hence, we choose proportional-integral-derivative PID controller by using the same method at the inner loop design, or $|O(j\omega)| \approx 1$:

$$C_{\text{queue}(s)} = k_{p_O}\left(1 + \frac{1}{T_{i_O}s} + T_{d_O}s\right), \tag{11}$$

with $k_{p_O} = (T_{1_O} + T_{2_O})/2k_O T_{3_O}$, $T_{i_O} = T_{1_O} + T_{2_O}$, $T_{d_O} = (T_{1_O} \cdot T_{2_O})/(T_{1_O} + T_{2_O})$.

4.2. CPC Numerical Analysis. The proposed CPC controllers can be easily verified by using example parameters from [12] which consist of $N = 60$, $C_l = 3750$ (packets/sec), and $R_0 = 0.246$ (sec). According to the above analysis, the inner controller $C_{\text{win}} = 0.000291$/s, while the outer controller $C_{\text{queue}} = 0.1281(1 + 1/(3.78\,\text{s}) + 1.89\,\text{s})$.

Figure 2(a) presents the queue length output in case of 20 packets queue reference. The CPC manual tune mode uses our above designed controllers, while the CPC autotune mode uses *pidtune* function of Matlab. Firstly, in comparison to RED which can be modeled as a single-loop I-type controller, one advantage of CPC is that fast transient behavior can be achieved. Transient behavior is a major issue of current solutions to bufferbloat. Fast transient means that we can reach the queue length reference quickly in response to the input change of dynamic systems.

Secondly, crossover frequency is a criterion to assess a control system's operation ability in a wide range of frequency. The higher the crossover frequency is, the better stability at which the system can operate is. Let us denote crossover frequency as ω_c. If we choose a frequency higher than ω_c, system would be not stable anymore. The left-hand side of Figure 2(b) is Bode diagram phase-margin of the system. It informs about ω_c information of our proposed CPC scheme. Specifically, CPC manual tune has $\omega_c = 2.37$ (rad/s) which is the highest value, while $\omega_c = 0.627$ (rad/s) for CPC autotune and $\omega_c = 0.261$ (rad/s) for RED. Therefore, CPC can operate better in a wider range of frequency.

Finally, using Nyquist stability criterion, the closed-loop transfer function of the outer loop $O(s)$ is determined by the values of its poles. It states that, for stability, the real part of all poles must be negative or the poles are in the left half-plane of pole-zero map. The right-hand side of Figure 2(b) shows us the distribution of zeros and poles for three schemes. Clearly, all the poles have the real part negative which strongly demonstrates CPC's stability.

In summary, CPC scheme can achieve fast transient, high crossover frequency and still stable. Those motivate us to conduct simulations about CPC behavior to bufferbloat under multiple real-world TCP stacks.

5. Simulation Results

In this section, we implement the proposed scheme in simulator to show bufferbloat phenomenon and the advantages of CPC scheme in comparison to traditional AQMs. We choose the dumbbell topology according to AQM guidelines [25].

Table 1: Simulation parameters.

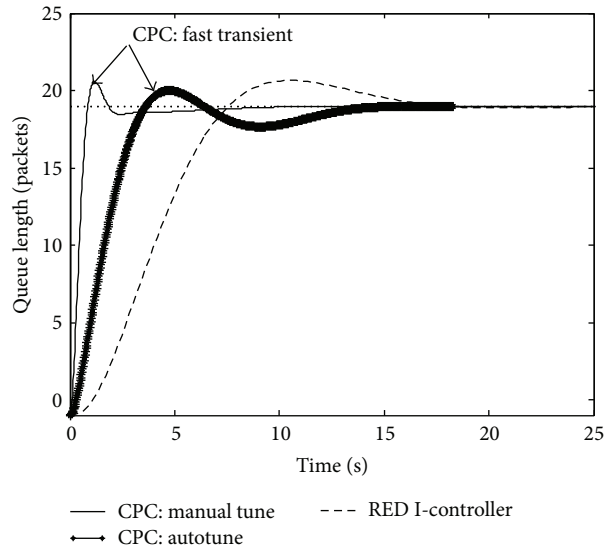
Name	Value	Unit
Maximum buffer size	500	Packets
Desired queue length	20	Packets
Target RTT delay	200	ms

5.1. Simulation Setup. We develop our own simulation model to verify the proposed CPC scheme using the popular OMNeT++ framework [26]. The chosen topology in Figure 3 represents a dumbbell network which is suitable for evaluating a queue management scheme. Three clients simultaneously send 200 × 3 (MByte) data to servers through intermediate routers. The advertised window's server is set up to infinity so that it does not limit the sender's speed. To create an artificial bottleneck, we set up the "high-to-slow" link bandwidth where the bandwidth of links from clients to router 1 are high at 1000 (Mbps) and the bandwidth of the link from router 1 to others is slow at 2 (Mbps). RTT can be measured at each server, while queue size is monitored at the congestion point. Bursty traffic is generated using a generator traffic model inside each client from OMNeT++. At the first glance, we vary the buffer size parameter to show the simulation model working toward bufferbloat phenomenon existence. Then we compare the CPC scheme with two types of traditional queue discipline which are operating inside the current Internet, Drop Tail, and RED, in terms of queue size and round-trip time or latency. Some main parameters are summarized in Table 1.

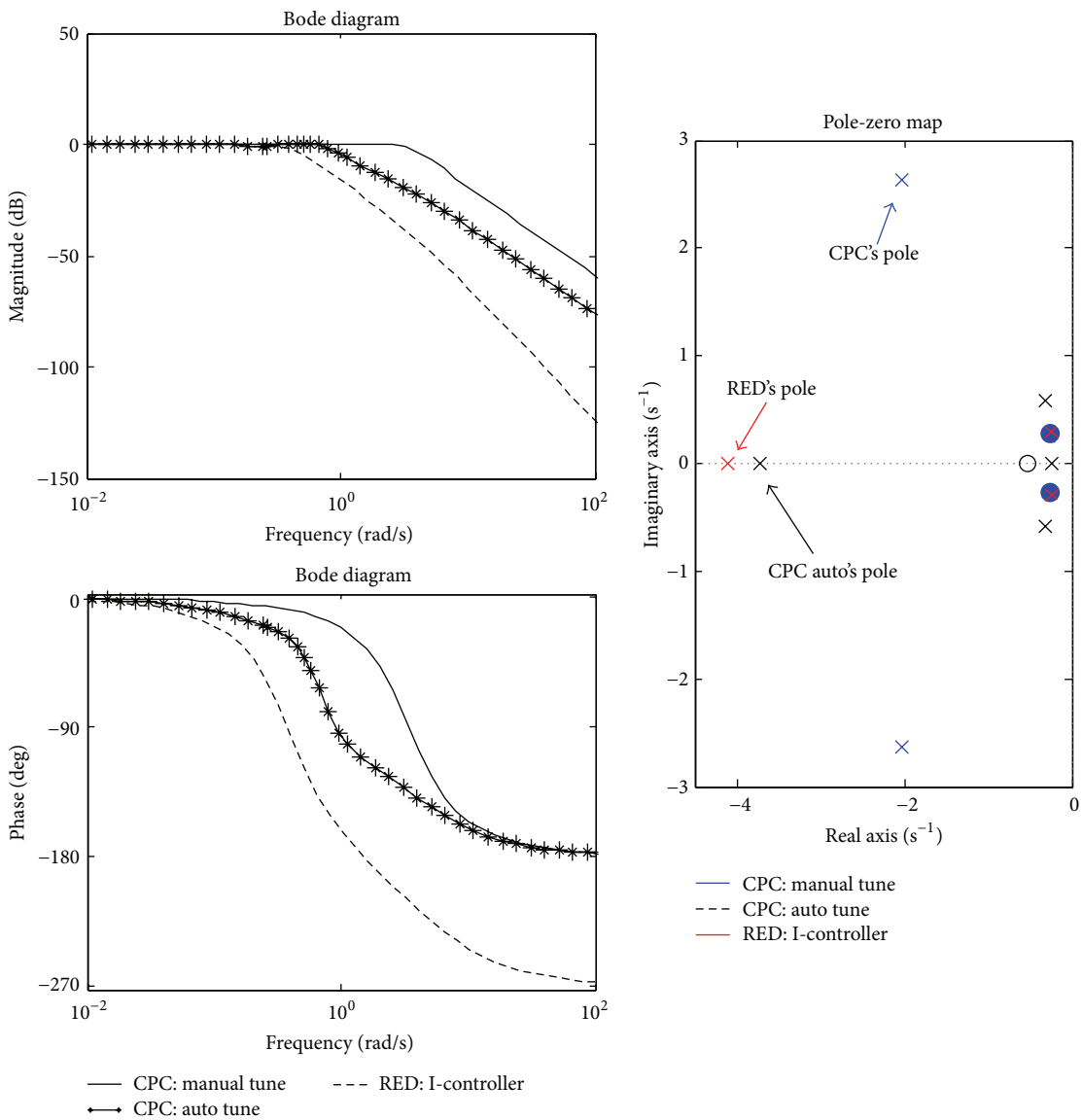
5.2. Bufferbloat Existence

5.2.1. Experiment. To demonstrate clearly bufferbloat existence, we exploit the real-time response under load (RRUL) test specification in [27]. This test puts a network under worst case conditions and then measures latency responsiveness and other relative performances of TCP and UDP streams of varying rates, while under that load. Then, Toke Hoiland-Jorgensen produces a wrapper for netperf tool to implement test cases such as HTTP, VoIP, and FTP under RRUL. In this paper, we conduct the experiment test from our computer-client located in South Korea to two servers which are mainly dedicated for bufferbloat testing, demo.tohojo.dk, and snapon.lab.bufferbloat.net, respectively. Figures 4(a) and 4(b) present our ping results when we run experiment in 300 seconds and sampling period of 0.2 second. Ping delay for both cases are approximately 300 (ms), while 100 → 150 (ms) is an acceptable range for toll-quality voice and delay-sensitive applications. Once again, we see that debloat solutions were not deployed in our current experiment routers.

5.2.2. Simulation. As mentioned before, the main reason for bufferbloat is the unmanaged large buffer at bottleneck



(a) Step response performance



(b) Bode diagram and pole-zero map

Figure 2: Stability analysis of CPC.

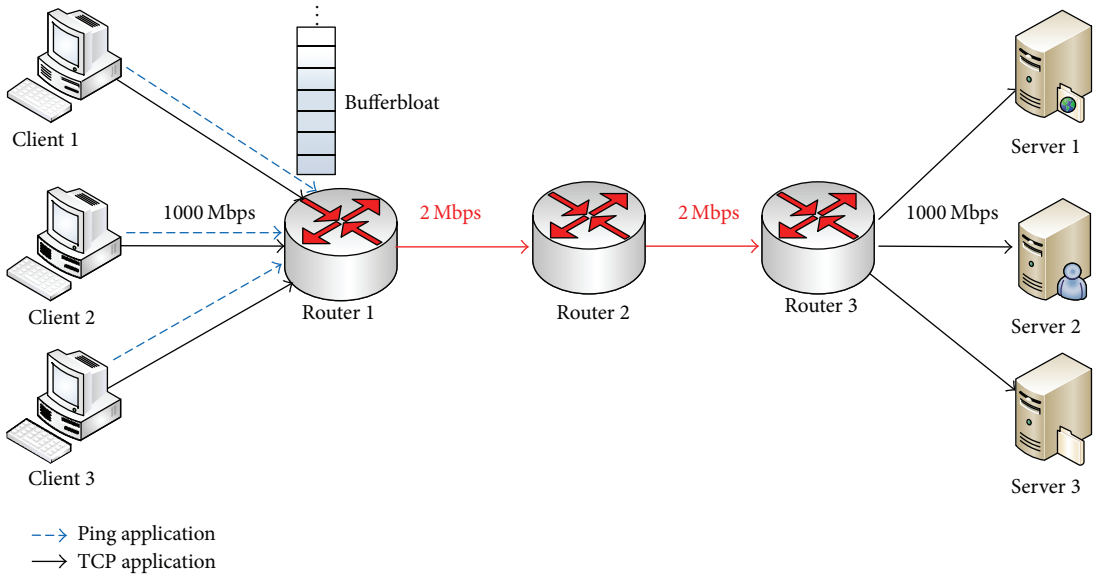
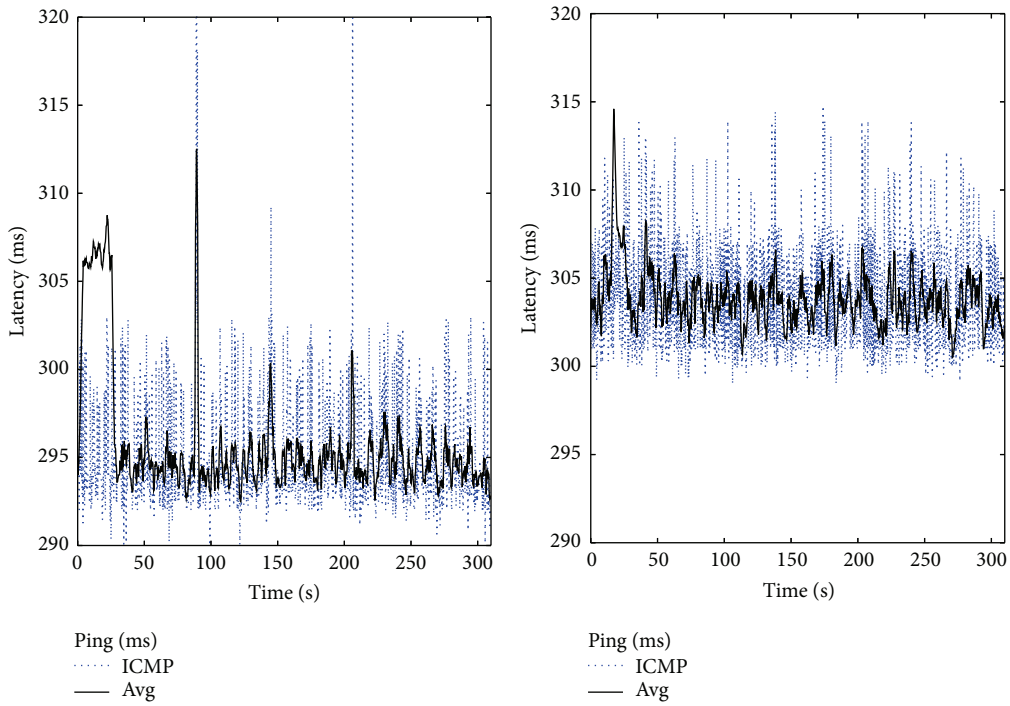


Figure 3: Simulation topology.



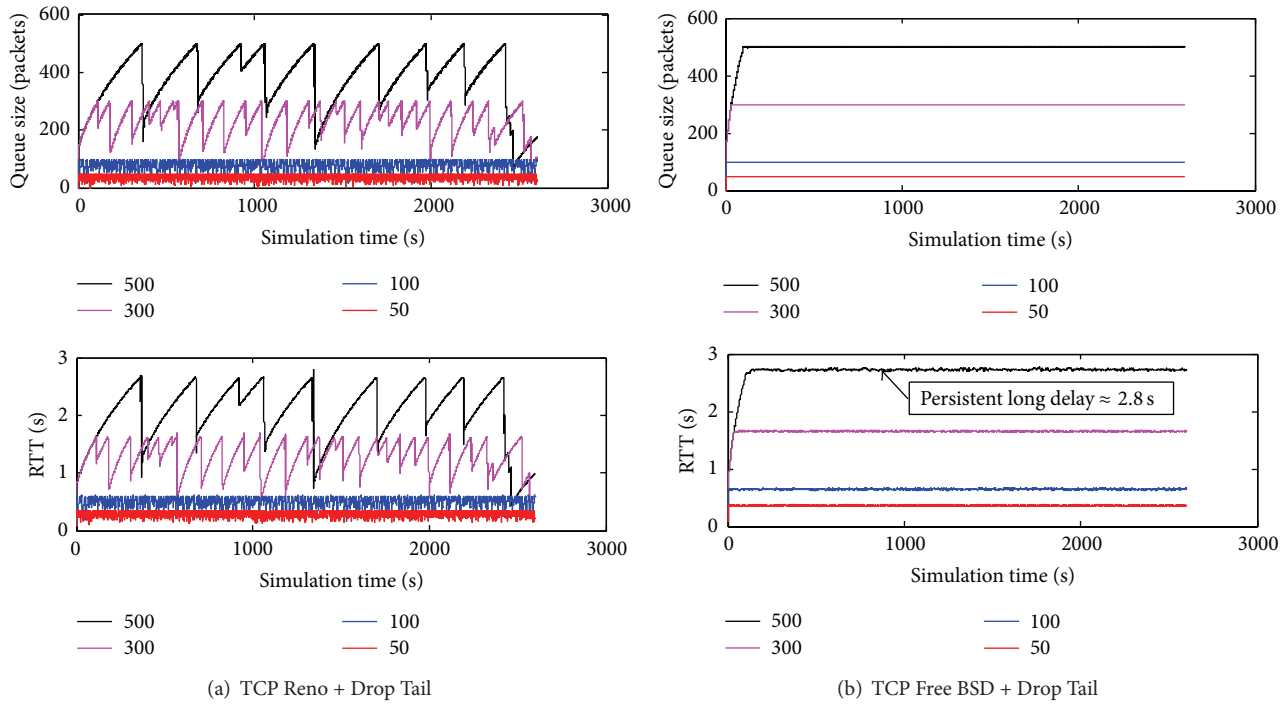
(a) Server: demo.tohojo.dk

(b) Server: snapon.lab.bufferbloat.dk

Figure 4: Bufferbloat existence: an experiment from Korea to two servers.

links that could damage TCP's congestion control/avoidance mechanisms. Hence, a big buffer is necessary to reproduce this phenomenon. In the current Internet, buffers are measured using bytes. OMNeT++ simulation framework, however, uses packets to sizing buffers. The choice of the largest buffer size is approximately 8x BDP product (packets) [11]. Let us denote buffer size as bu. With bottleneck bandwidth 2 Mbps and propagation delay 300 ms, we have the results when following the traditional rule-of-thumb for buffer-size:

$$\begin{aligned} \text{bu (byte)} &= \text{BDP} = \text{BW} \times \text{Delay} = 2.10^{6} \times 300.10^{-3} \\ &= 600 \text{ (kbit)} = 75 \text{ (kbyte)}. \end{aligned} \tag{12}$$



(a) TCP Reno + Drop Tail

(b) TCP Free BSD + Drop Tail

Figure 5: Bufferbloat existence and varied buffer sizes.

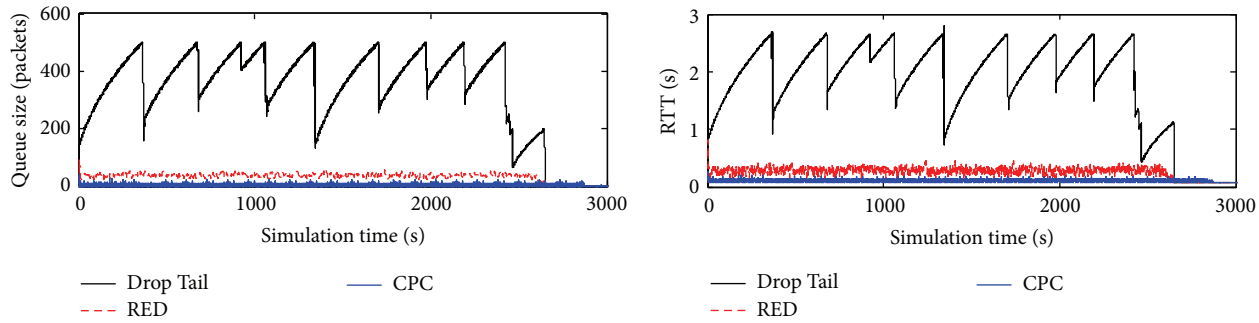
According to [28], one TCP packet size or maximum transmission unit (MTU) is 1500 (bytes):

$$\text{bu (packets)} = \frac{\text{bu (byte)}}{\text{MTU}} = \frac{75.10^3}{1500} = 50 \text{ (packets)}. \quad (13)$$

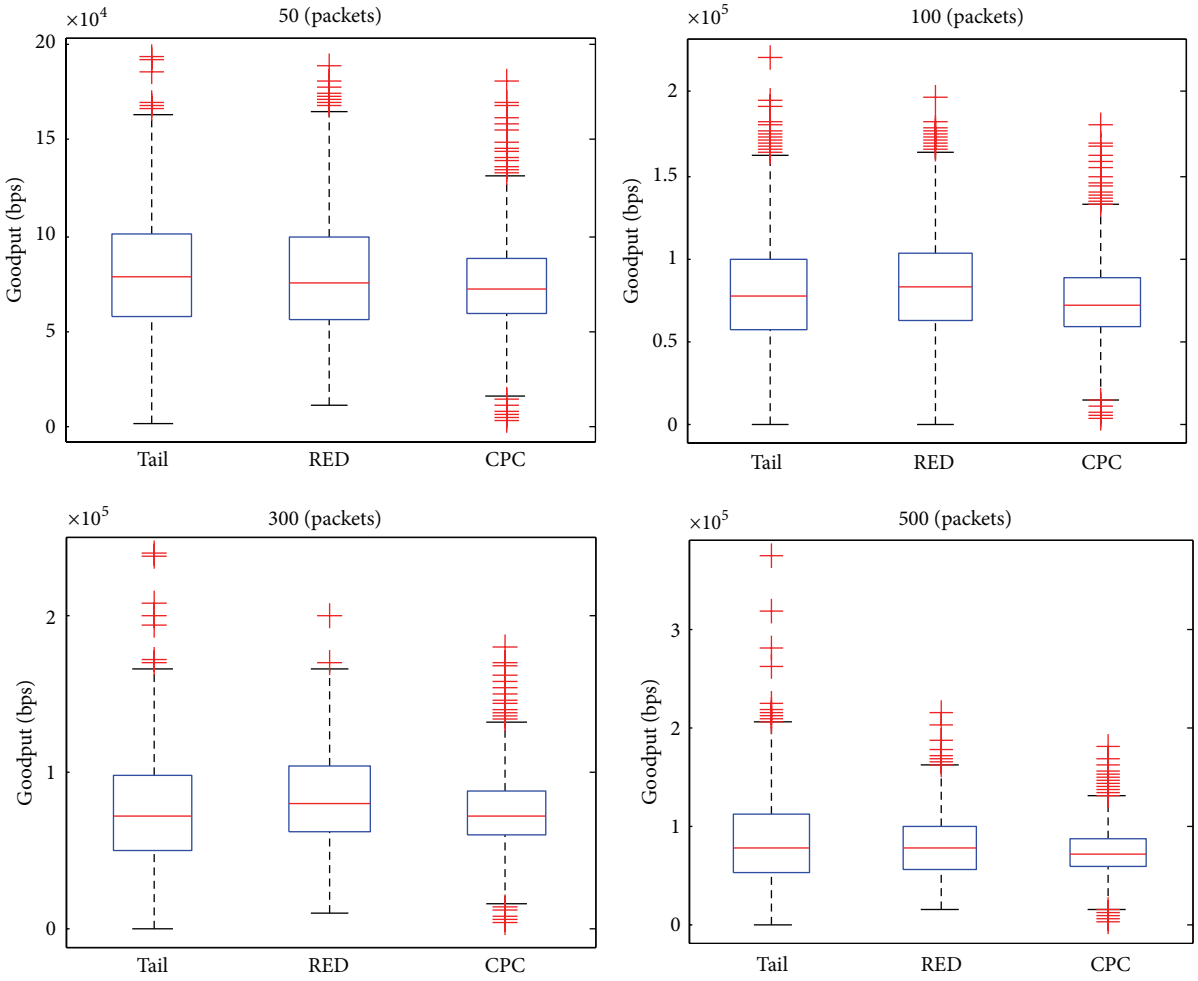
Therefore, in this simulation, we vary the buffer sizes such as 50, 100, 300, and 500 being the largest one (packets) under different TCP congestion-avoidance algorithms with Drop Tail queuing discipline. Firstly, Figure 5(a) shows the effects on queue evolution and round-trip time, under TCP Reno, the most used in Windows and some Linux operating systems (OS). We obtain the saw-tooth result at both queue size and RTT graph because of TCP Reno's characteristics while others do not have saw-tooth type graphs. From Figure 5(a), bufferbloat issue can also be clearly recognized. If buffer size is very large (500 packets), RTT will become pretty large (≈2.5 sec) while queue size grows until it reaches the threshold of buffer, which is the main drawback of Drop Tail. Several AQMs come to rescue at this point. The phenomenon once again appears in TCP FreeBSD, the popular algorithm in Unix OS that does not own a saw-tooth graph. Figure 5(b), which is simulated under TCP FreeBSD, also presents a long delay (≈2.8 sec) when buffer is too large. Those results confirm bufferbloat happening at different OS whenever a long buffer queue is built up.

5.3. CPC Performance under TCP Real-World Stacks. Next, we conduct more necessary simulations to compare our CPC scheme performance with traditional popular queuing disciplines Drop Tail and RED. We monitor the queue size at the bottleneck point (router 1), the round-trip time of packets, and goodput at each destination server. Goodput is the application level throughput, that is, the number of user information bits delivered by the network to a certain destination per unit of time. For example, if a file is transferred, the goodput that the user experiences corresponds to the file size in bits divided by the file transfer time. The goodput is always lower than the throughput (the gross bit rate that is transferred physically), which generally is lower than the channel capacity or bandwidth. Goodput monitoring is necessary to evaluate the actual good throughput performance inside networks. We also do these comparisons under several different TCP versions which include Reno, Cubic, and FreeBSD. The first one is mainly implemented in Windows and some Linux Kernels. The second one is improved and deployed in newer versions of Linux Kernel until now. The last one is popularly implemented under Unix Kernel. All the following results are derived when we set the largest buffer size at congestion point, 500 (packets). In the following, we summarize the background knowledge of those three TCP versions before going into detailed results of bufferbloat delay reduction and goodput stability.

5.3.1. Under Reno Real-World Stack. For each connection, TCP maintains a congestion window *cwnd*, limiting the total number of packets that may be transmitted at a time. TCP Reno algorithm uniqueness is the fast recovery phase. That means when a packet loss is detected by receiving three duplicate ACKs, Reno will halve *cwnd* and set slow start threshold *ssthresh* equal to the new *cwnd* value, perform a fast retransmit, and enter a phase called fast recovery. It also reduces congestion window to one maximum segment size (MSS) on a timeout event. Congestion-avoidance phase, as



(a) Round-trip time



(b) Goodputs

Figure 6: TCP Reno + Drop Tail/RED/CPC.

usual, exploits AIMD to control *cwnd* based on packet loss notification. In fact, compound TCP stack for Windows and Linux also increases AIMD window as TCP Reno does [29].

In Figure 6(a), queue size and RTT according to three queue management schemes are presented. With 500 packets of buffer, Reno and Drop Tail show a saw-tooth graph where the highest RTT is high up to 2.8 (sec). With the target RTT being 200 ms, RED under Reno presents better performance than Drop Tail, RTT (RED) $\approx$ 400 ms, by early dropping some packets to keep the queue size small. CPC under Reno, however, achieves a much better performance while keeping the lowest queue size and lowest RTT (CPC) $\approx$ 180 ms, which is acceptable for delay-sensitive applications. The reason is that CPC continuously controls the next queue size value

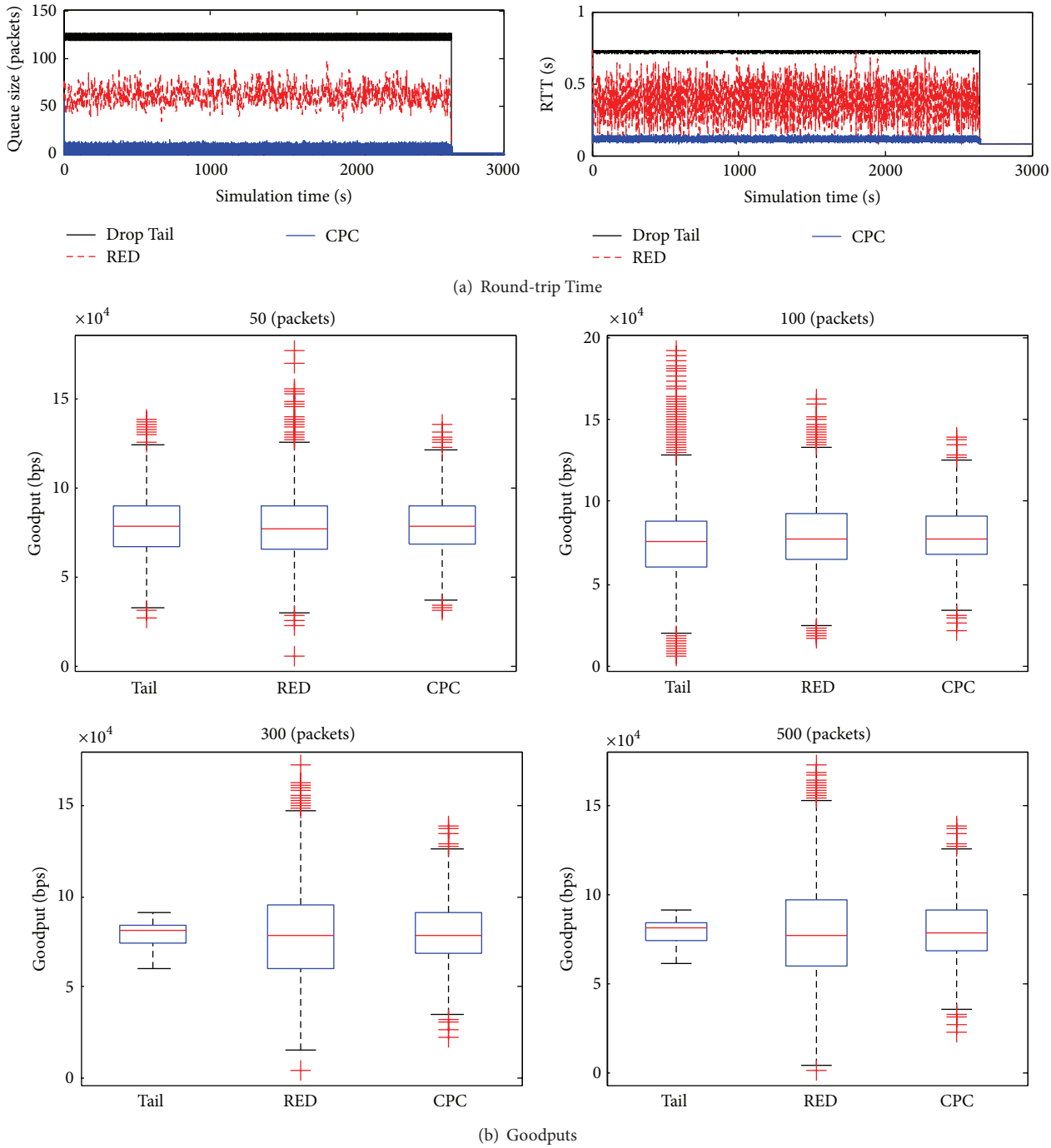
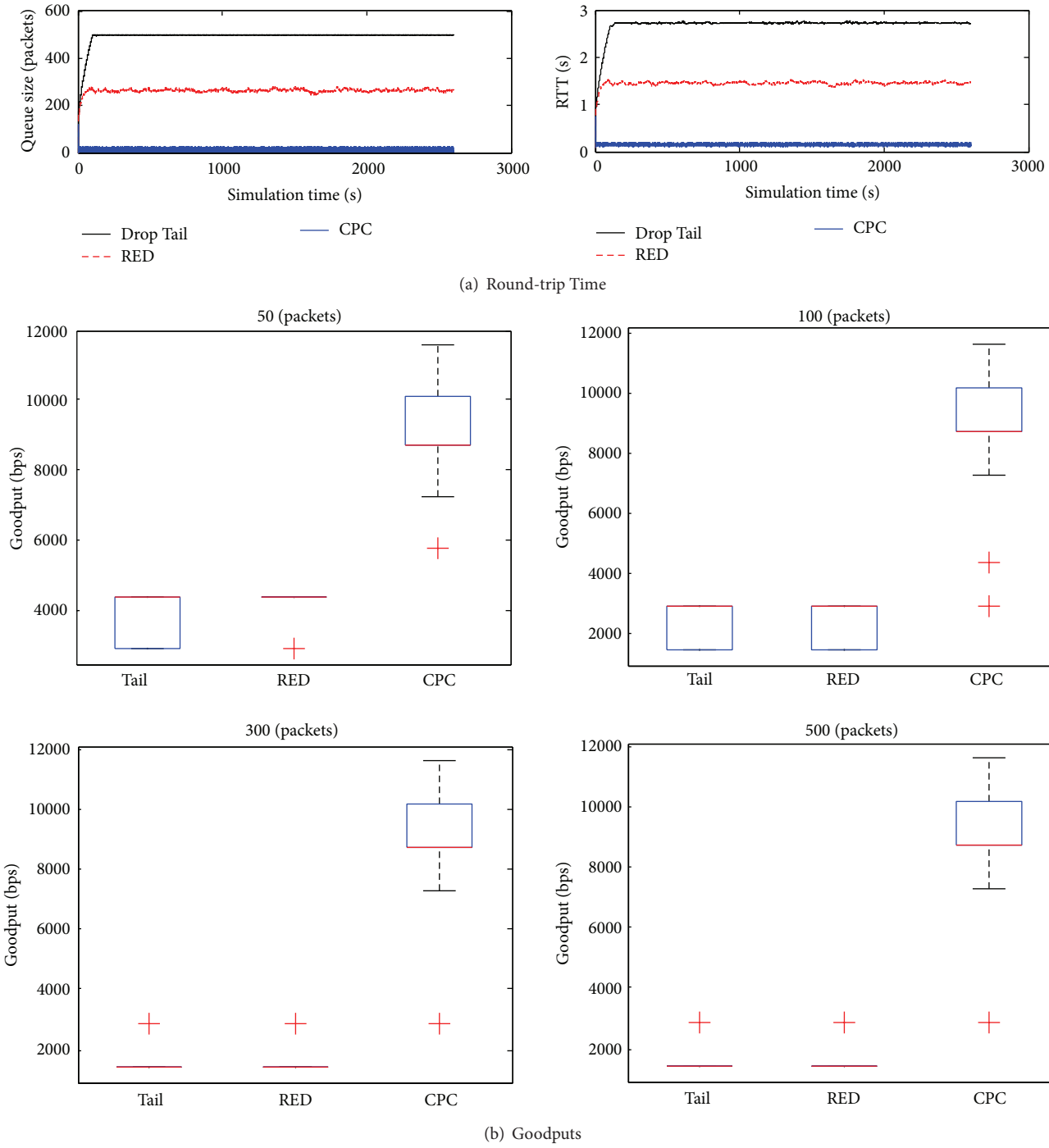


Figure 7: TCP Cubic + Drop Tail/RED/CPC.

according to traffic changes and adjusts the current queue size at the outer loop so that it indirectly affects RTT value of packets.

Next, Figure 6(b) shows us goodput (bit per second) statistic results under TCP Reno according to four different buffer sizes. These graphs use box-and-whisker plot that can effectively depict groups of numerical data through their quartiles. For instance, the center red line of box represents the median goodput value; the width of box shows the spread degree of goodput statistic data; for example, bigger box and longer whisker mean that statistic data oscillate so much or are unstable. Regardless of different buffer size, under TCP Reno, the median of goodput is nearly the same, around 8×10^4 (bps). The CPC scheme, however, achieves the best stability performance compared to the others because the width of CPC boxes is smallest.

5.3.2. Under Cubic Real-World Stack. Cubic is a congestion control protocol for TCP and the current default TCP algorithm in the Linux kernel [30]. It replaces the default



(a) Round-trip Time

(b) Goodputs

Figure 8: TCP FreeBSD + Drop Tail/RED/CPC.

TCP congestion control algorithm described above with a different algorithm based on a cubic function of the time since the last congestion event. Instead of adjusting the congestion window as a function of previous values as each packet is acknowledged, the Cubic algorithm recomputes the congestion window size at each step using the cubic function calibrated. This results in the congestion window responding quickly to changes in available bandwidth. The TCP Cubic is a less aggressive and more systematic derivative of binary increase congestion (BIC). We exploit NSC module to incorporate real-world TCP/IP network stacks into our simulation model. NSC owns TCP Cubic (version 2.6.26) stack which allows us to test our CPC queuing scheme performance.

In Figure 7(a), we can see again that CPC scheme seems to achieve the best performance, keeping lower queue size and RTT only 130 ms, which is acceptable for delay-sensitive traffic. However, the RTT gaps between three schemes are not

so much because of the aggression property of TCP Cubic. While RED shows large variability of RTT, the others keep RTT much more stable.

Goodput performance in Cubic also reveals TCP Cubic's aggression. When buffers are at small values (50 and 100 packets), three boxes' widths are nearly similar. That means that CPC can achieve the same goodput as the others if we use small buffers. Moreover, increasing the buffer size to 300 and 500 packets makes CPC's goodput more stable, with the smaller box width. These figures demonstrate that CPC can mitigate the bufferbloat—big buffers problem pretty well.

5.3.3. Under FreeBSD Real-World Stack. Finally, we investigate one more popular TCP stack. FreeBSD is a free Unix-like operating system, historically standing for "Berkeley Unix." It was chosen as a BSD-derived TCP/IP stack that is widely used and has had much development. There are several applications that are directly based on FreeBSD, an example being the famous instant messenger WhatsApp. Much of FreeBSD became an integral part of other operating systems such as Apple's OS X. The integrated NSC module for OMNeT++ also incorporates TCP FreeBSD real-world stack so that we can turn it on for demonstrating CPC's performance under FreeBSD. Figure 8(a) presents queue size and RTT results for three disciplines. We can see that our proposal CPC achieves a better performance, keeping lower queue size and RTT (FreeBSD) around 150 ms.

So far, the proposed CPC scheme achieved good goodput performance under TCP Reno and Cubic stacks. We, however, find another interesting fact of CPC when performing under TCP FreeBSD stack. In Figure 8(b), we can conclude two things. Firstly, the width of CPC goodput's boxes in four buffer cases is almost bigger than the other schemes. It means the CPC stability under FreeBSD stack's environment is not good. Secondly, our CPC scheme, however, still achieves more goodput than others. Specifically, in four cases, the system's goodput can reach 10 (kbps) using CPC while it can reach only 4 (kbps) and 2 (kbps) under either small or large buffer using Drop Tail or RED schemes, respectively.

6. Conclusions

Bufferbloat problem, with overfilling a queue at the low-speed side, leads to long end-to-end latency because of the persistent long queuing delay. In this paper, we proposed a control scheme called cascade probability control (CPC) as an alternative way to tackle this issue and reduce queuing delay. Our scheme introduces two-loop control model which consists of queue length and window size control to keep round-trip time value around an expected value. In comparison to current deployed schemes, our scheme's advantages are to maintain stability and noise reduction while the queuing delay can still be ensured under several different TCP versions. Further works would extend CPC on bufferbloat and queuing delay control in the cellular communications.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This research was supported by the MSIP (Ministry of Science, ICT, and Future Planning), Korea, under the Global IT Talent support program (R0618-15-1001) supervised by the IITP (Institute for Information and Communications Technology Promotion).

References

[1] J. B. Nagle, "On packet switches with infinite storage," *IEEE Transactions on Communications*, vol. 35, no. 4, pp. 435–438, 1987.

[2] J. Gettys, "Bufferbloat: dark buffers in the Internet," *IEEE Internet Computing*, vol. 15, no. 3, pp. 95–96, 2011.

[3] T. Hoiland-Jorgensen, *Battling bufferbloat: an experimental comparison of four approaches to queue management in linux [Master module project]*, Roskilde University, 2012.

[4] M. Hassan and R. Jain, *High Performance TCP/IP Networking: Concepts, Issues, and Solutions*, Prentice-Hall, 2003.

[5] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," *IEEE/ACM Transactions on Networking*, vol. 1, no. 4, pp. 397–413, 1993.

[6] J. Sun, M. Zukerman, and M. Palaniswami, "Stabilizing RED using a fuzzy controller," in *Proceedings of the IEEE International Conference on Communications (ICC '07)*, pp. 266–271, June 2007.

[7] S. Floyd, R. Gummadi, and S. Shenker, "Adaptive red: an algorithm for increasing the robustness of red's active queue management," Tech. Rep., 2001.

[8] M. May, J. Bolot, C. Diot, and B. Lyles, "Reasons not to deploy red," in *Proceedings of the 7th International Workshop on Quality of Service (IWQoS '99)*, pp. 260–262, 1999.

[9] J. Hwang and S.-S. Byun, "A resilient buffer allocation scheme in active queue management: a stochastic cooperative game theoretic approach," *International Journal of Communication Systems*, vol. 28, no. 6, pp. 1080–1099, 2015.

[10] M. Khosroshahy, "UARA in edge routers: an effective approach to user fairness and traffic shaping," *International Journal of Communication Systems*, vol. 25, no. 2, pp. 169–184, 2012.

[11] K. Nichols and V. Jacobson, "Controlling queue delay," *Queue*, vol. 10, no. 5, p. 20, 2012.

[12] C. V. Hollot, V. Misra, D. Towsley, and W.-B. Gong, "A control theoretic analysis of RED," in *Proceedings of the 20th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '01)*, vol. 3, pp. 1510–1519, IEEE, Anchorage, Alaska, USA, April 2001.

[13] R. Zhu, H. Teng, and J. Fu, "A predictive PID controller for AQM router supporting TCP with ECN," in *Proceedings of the 2004 Joint Conference of the 10th Asia-Pacific Conference on Communications and the 5th International Symposium on Multi-Dimensional Mobile Communications*, vol. 1, pp. 356–360, August 2004.

[14] G. Kahe, A. H. Jahangir, and B. Ebrahimi, "A compensated PID active queue management controller using an improved

queue dynamic model," *International Journal of Communication Systems*, vol. 27, no. 12, pp. 4543–4563, 2014.

[15] H.-L. To, G.-O. Yoon, J.-H. Nam, G. Solongo, and W.-J. Hwang, "Feedback burst loss ratio control for link performance improvement in optical burst switching networks," *Journal of Korea Multimedia Society*, vol. 16, no. 9, pp. 1067–1076, 2013.

[16] K. B. Kim, "Design of feedback controls supporting TCP based on the state-space approach," *IEEE Transactions on Automatic Control*, vol. 51, no. 7, pp. 1086–1099, 2006.

[17] H.-L. To, S.-H. Lee, and W.-J. Hwang, "A burst loss probability model with impatient customer feature for optical burst switching networks," *International Journal of Communication Systems*, vol. 28, no. 11, pp. 1729–1740, 2015.

[18] R. Adams, "Active queue management: a survey," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 3, pp. 1425–1476, 2013.

[19] J. Aweya, M. Ouellette, and D. Y. Montuno, "DRED: a random early detection algorithm for TCP/IP networks," *International Journal of Communication Systems*, vol. 15, no. 4, pp. 287–307, 2002.

[20] F. Ren, C. Lin, and B. Wei, "Design a robust controller for active queue management in large delay networks," in *Proceedings of the 9th International Symposium on Computers and Communications (ISCC '04)*, vol. 2, pp. 748–754, June 2004.

[21] S. Xiang, B. Xu, S. Wu, and D. Peng, "Gain adaptive smith predictor for congestion control in robust active queue management," in *Proceedings of the 6th World Congress on Intelligent Control and Automation (WCICA '06)*, vol. 1, pp. 4489–4493, June 2006.

[22] D.-M. Chiu and R. Jain, "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks," *Computer Networks and ISDN Systems*, vol. 17, no. 1, pp. 1–14, 1989.

[23] G. Patil and G. Raina, "Some guidelines for queue management in high-speed networks," in *Proceedings of the 3rd Asian Himalayas International Conference on Internet (AH-ICI '12)*, pp. 1–6, IEEE, Kathmandu, Nepal, November 2012.

[24] F. Du and J. Sun, "An aqm scheme based on adaptive weight cascaded pid controller," in *Proceedings of the 10th World Congress on Intelligent Control and Automation (WCICA '12)*, pp. 2849–2854, July 2012.

[25] N. Kuhn, P. Natarajan, D. Ros, and N. Khademi, *AQM Characterization Guidelines*, IETF89, 2014.

[26] A. Varga, "Using the OMNeT++ discrete event simulation system in education," *IEEE Transactions on Education*, vol. 42, no. 4, p. 11, 1999.

[27] D. Taht, Realtime Response Under Load (RRUL) Test, draft 07, 2012, https://github.com/dtaht/deBloat/blob/master/spec/rrule.doc.

[28] C. Hornig, "A standard for the transmission of IP datagrams over ethernet networks," Network Working Group RFC894, 1984.

[29] D. J. Leith, L. L. H. Andrew, T. Quetchenbach, R. N. Shorten, and K. Lavi, "Experimental evaluation of delay/loss-based TCP congestion control algorithms," in *Proceedings of the 6th International Workshop on Protocols for Fast Long-Distance Networks (PFLDnet '08)*, University of Manchester, Manchester, UK, March 2008.

[30] S. Ha, I. Rhee, and L. Xu, "Cubic: a new tcp-friendly high-speed tcp variant," *ACM SIGOPS Operating Systems Review*, vol. 42, no. 5, pp. 64–74, 2008.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 206162, 12 pages
http://dx.doi.org/10.1155/2015/206162

Research Article

Global Asymptotic Stabilization Control for a Class of Nonlinear Systems with Dynamic Uncertainties

Jiangbo Yu,[1] Jizhong Wang,[1] and Zhongcai Zhang[2]

[1] *School of Science, Shandong Jianzhu University, Jinan 250101, China*
[2] *School of Automation, Southeast University, Nanjing 210096, China*

Correspondence should be addressed to Jiangbo Yu; jbyu2002@163.com

Received 11 May 2015; Accepted 16 August 2015

Academic Editor: Uchechukwu E. Vincent

This paper is concerned with the global asymptotic stabilization control problem for a class of nonlinear systems with input-to-state stable (ISS) dynamic uncertainties and uncertain time-varying control coefficients. Unlike the existing works, the ISS dynamic uncertainty is characterized by the uncertain supply rates. By using the backstepping control approach, a systematic controller design procedure is developed. The designed control law can guarantee that the system states are asymptotically regulated to the origin from any initial conditions and the other signals are bounded in closed-loop systems. Moreover, it is shown that, under some additional conditions, a linear control law can be designed by the proposed methodology. The simulation example demonstrates its effectiveness.

1. Introduction

The nonlinear control theory is an active research direction in the control field because of its widespread applications in the real world. During the past two decades, various novel methodologies have been generated for the nonlinear feedback control; see the recent survey [1] and references therein for an interesting introduction to this area. One of the influential notions is the input-to-state stability (ISS) and its several variants. Since they are introduced by Sontag in [2, 3], the notion of ISS as well as its integral variant—integral ISS (iISS)—has become a foundational concept upon which much of modern nonlinear feedback analysis and design rest. As noted in [4], ISS provides a nonlinear generalization of finite gains with respect to supremum norms and also of finite L^2 gains, and it plays a central role in recursive design, coprime factorizations, controllers for nonminimum phase systems, and many other areas. Based on the series of works on ISS, the nonlinear small-gain theorem was proposed in the state-space setting and is widely used in the stability analysis and control design for complex interconnected systems in [5]. The stochastic results can be found in [6, 7] and the references therein.

It is noted that a unifying framework is presented in [8] for the global output feedback regulation control problem from ISS to iISS. The framework established in [8] extends many known classes of output feedback form systems. However, the system uncertainties investigated there depend only on the system output and the inverse system state. With unmeasured states dependent growth, in [9, 10], the problem of global stabilization by output/state feedback is investigated for a class of nonlinear systems with uncertain control coefficients. However, there is no dynamic uncertainty for the system under consideration. In [11], this work is further studied for a larger class of nonlinear uncertain systems, in which the observer gain is governed by a Riccati differential equation. Moreover, the output regulation problem is also considered in [12] for this class of nonlinear systems with iISS inverse dynamics. Later, in [13, 14], this work is further investigated for the nonlinear systems with uncertain nonlinearities dependent on all unmeasured states. However, the control coefficients in above results are required to be known a priori or unknown nonzero constants. In [15], the global set-point tracking control is investigated for a class of cascaded nonlinear systems with unknown control coefficients. However, a restrictive condition is that the control coefficients are required to have the same signs.

In this paper, we will further investigate this problem for a class of nonlinear systems with more general nonlinear

uncertainties. Unlike the existing works such as in [9, 12, 15–17], the studied system is with the uncertain control coefficients, which could be unknown time-varying functions. Another feature of this work is that the dynamic uncertainties are characterized by the uncertain ISS supply rates. This is different from the existing results reported in literatures where the ISS dynamic uncertainty is investigated under the hypothesis that supply rates are known a priori such as in [8, 11, 13, 15]. With the help of the backstepping approach [18], we design a robust adaptive controller which could achieve the system states convergent to the origin while the other signals are bounded. Moreover, it is of interest to note that a linear control law can be designed using the developed scheme if some stronger conditions are imposed on the nonlinear system.

The rest of the paper is organized as follows. In Section 2, we provide some mathematical preliminaries and state the problem. The controller design procedure is developed in Section 3, and the main result is presented in Section 4. Section 5 illustrates the obtained results by a numerical example. Section 6 concludes this paper.

Notation. Let $\mathbf{R}(\mathbf{R}_+)$ denote the set of all (positive) real numbers and let $R^{\mathbf{n}}$ denote the real n-dimensional space. For a given vector or matrix X, X^T denotes its transpose. For any column vector $x = (x_1, \ldots, x_n)^T \in \mathbf{R}^{\mathbf{n}}$, $\overline{x}_i$ denotes the column vector consisting of the first i elements of x in the original order; that is, $\overline{x}_i = (x_1, \ldots, x_i)^T$. Specifically, for $x = (x_1, \ldots, x_n)^T$, $x_1 = \overline{x}_1$, $x = \overline{x}_n$. A continuous function $\alpha : [0, a) \to [0, \infty)$ is said to belong to class-K if it is strictly increasing and $\alpha(0) = 0$. It is said to belong to class-K_∞ if $a = \infty$ and $\alpha(r) \to \infty$ as $r \to \infty$. The notation $\widetilde{\beta}(s) = O(\beta(s))$ means that there exist two positive constants k and c such that $\widetilde{\beta}(s) \le k\beta(s)$, $\forall |s| < c$.

2. Problem Formulation

In this paper, we consider the following class of cascaded nonlinear systems with dynamic uncertainties:

$$\begin{aligned} \dot{\eta} &= q(t, \eta, y), \\ \dot{x}_1 &= d_1(t) x_2 + g_1(t, \eta, x), \\ &\vdots \\ \dot{x}_n &= d_n(t) u + g_n(t, \eta, x), \\ y &= x_1, \end{aligned} \tag{1}$$

where $u \in \mathbf{R}$ is the control input, $y \in \mathbf{R}$ is the system output, $x = (x_1, \ldots, x_n) \in \mathbf{R}^{\mathbf{n}}$ are the system states, and $\eta \in \mathbf{R}^{\mathbf{r}}$ is referred to as dynamic uncertainties, which is unmeasured and hence is not available for feedback design. The continuous functions $d_i(t)$ $(i = 1, \ldots, n)$ called the control coefficients are assumed to be unknown; particularly, $d_i(t) \neq 0$; the unmodeled (or uncertain) dynamics $q(\cdot)$ and $g_i(\cdot)$ $(i = 1, \ldots, n)$ are locally Lipschitz.

The control objective in this paper is to find a smooth, dynamic, partial-state feedback law of the form

$$\begin{aligned} \dot{\xi} &= \chi_\xi(\xi, x), \\ u &= \chi_u(\xi, x), \end{aligned} \tag{2}$$

where χ_ξ and χ_u are smooth functions such that all solutions $(\eta(t), x(t), \xi(t))$ in closed-loop system are bounded on $[0, \infty)$ and specially the system states $(\eta(t), x(t))$ asymptotically converge to the origin. Toward this end, throughout the paper, we make the following assumptions on system (1).

Assumption 1. The η-subsystem is input-to-state stable (ISS) with state η and input y; that is, there exists a positive-definite and proper ISS-Lyapunov function V_0, such that

$$\begin{aligned} \underline{\alpha}(|\eta|) \le V_0(\eta) &\le \overline{\alpha}(|\eta|), \\ \frac{\partial V_0}{\partial \eta}(\eta) q(t, \eta, y) &\le -\alpha_0(|\eta|) + \delta_0 \gamma_0(|y|), \end{aligned} \tag{3}$$

where $\underline{\alpha}(\cdot), \overline{\alpha}(\cdot), \alpha_0(\cdot), \gamma_0(\cdot) \in K_\infty$ and $\delta_0 > 0$ is an unknown constant.

Remark 2. According to [19], one knows that η-subsystem satisfying (3) is ISS, and the function pair $(\alpha_0, \delta_0\gamma_0)$ is viewed as the supply rates. Since δ_0 in (3) is unknown, the dynamic uncertainty has uncertain ISS supply rates. This is different from the existing results reported in literatures, where the ISS dynamic uncertainty is investigated with the supply rates assumed to be known a priori, such as [8, 11, 13, 15].

Assumption 3. For the uncertain nonlinearities $g_i(t, \eta, x)$ $(i = 1, \ldots, n)$, there exist unknown positive constants p_{ij} $(i = 1, \ldots, n;\ j = 1, 2)$ such that

$$|g_i(t, \eta, x)| \le p_{i1}\phi_{i1}(|\eta|) + p_{i2}\phi_{i2}(|x_1, \ldots, x_i|), \quad i = 1, \ldots, n, \tag{4}$$

where $\phi_{ij}(\cdot)$ are known smooth functions and $\phi_{ij}(0) = 0$ $(i = 1, \ldots, n;\ j = 1, 2)$.

Assumption 4. There exist known positive constants $\underline{d}_i$ and $\overline{d}_i$ $(i = 1, \ldots, n)$, such that

$$0 < \underline{d}_i \le d_i(t) \le \overline{d}_i, \quad i = 1, \ldots, n. \tag{5}$$

To deal with the unmeasured state η, we have the following lemma, which plays an important role in the coming feedback design and stability analysis.

Lemma 5. *Consider the η-subsystem satisfying Assumption 1. Suppose*

$$\gamma_0(s) = O(s^2), \tag{6}$$

and then we can choose a positive continuous function $\rho(\cdot)$, such that the function

$$\overline{V}_0(\eta) = \int_0^{V_0(\eta)} \rho(s)\, ds \tag{7}$$

is another candidate ISS-Lyapunov function satisfying

$$\frac{\partial \overline{V}_0}{\partial \eta}(\eta)\, q(t,\eta,y) \le -(1-\epsilon)\,\rho\circ\underline{\alpha}(|\eta|)\,\alpha_0(|\eta|) + \delta\gamma(|y|), \tag{8}$$

where ϵ $(0<\epsilon<1)$ is a small design constant, $\delta>0$ is an unknown constant, and γ is a K_∞-function with $\gamma(s)=O(s^2)$.

Remark 6. If $\gamma(s)=O(s^2)$, according to Lemma 2 in [19], there exists a smooth nondecreasing function $\overline{\gamma}(\cdot)$ satisfying

$$\gamma(s) \le \overline{\gamma}(s)\, s^2. \tag{9}$$

Lemma 7. *For any C^1 function $f(x_1,\ldots,x_n)$, there exist continuous functions $f_i(x_1,\ldots,x_n)$ $(1\le i\le n)$, such that*

$$f(x_1,\ldots,x_n) = f(0,\ldots,0) + \sum_{i=1}^{n} x_i f_i(x_1,\ldots,x_n). \tag{10}$$

Remark 8. According to Lemma 7, from $\phi_{i2}(0)=0$ $(i=1,\ldots,n)$ in Assumption 3, it is known that, for each $\phi_{i2}(\cdot)$, there exist smooth functions $\phi_{i2j}(\cdot)$ $(1\le j\le i)$ satisfying

$$\phi_{i2}(x_1,\ldots,x_i) = \sum_{j=1}^{i} x_j \phi_{i2j}(x_1,\ldots,x_i), \quad i=1,\ldots,n. \tag{11}$$

3. Controller Design

In this section, we give the controller design procedure using the backstepping design method.

Step 1. Starting with the x_1-subsystem $\dot{x}_1 = d_1(t)x_2 + g_1(t,\eta,x)$. We consider the variable x_2 as the virtual control input. Let $z_1 = x_1$ and $z_2 = x_2 - \vartheta_1$ where ϑ_1 is the intermediate control input. Considering Lemma 5 and Remark 6, along solutions of (1), the time derivative of the function

$$V_1 = \frac{1}{2}z_1^2 + \overline{V}_0(\eta) \tag{12}$$

satisfies

$$\begin{aligned}\dot{V}_1 &\le z_1\left(d_1(t)\,\vartheta_1 + g_1(t,\eta,x)\right) + d_1(t)\,z_1 z_2 \\ &\quad -(1-\epsilon)\,\rho\circ\underline{\alpha}(|\eta|)\,\alpha_0(|\eta|) + \delta\overline{\gamma}(|x_1|)\,x_1^2.\end{aligned} \tag{13}$$

According to Assumption 3 and the completion of squares, we have

$$\begin{aligned}z_1 g_1(t,\eta,x) &\le |z_1|\left(p_{11}\phi_{11}(|\eta|) + p_{12}|z_1|\,\phi_{121}(x_1)\right) \\ &\le \phi_{11}^2(|\eta|) + \frac{1}{4}z_1^2 {p_{11}}^2 \\ &\quad + z_1^2 p_{12}|\phi_{121}(x_1)|.\end{aligned} \tag{14}$$

Define $p^* = \max\{\delta,\ p_{i1},\ p_{i2},\ p_{i1}^2,\ p_{i2}^2 \mid i=1,\ldots,n\}$, and we get

$$\begin{aligned}&z_1 g_1(t,\eta,x) + \delta\overline{\gamma}(|x_1|)\,x_1^2 \\ &\quad\le \phi_{11}^2(|\eta|) + \frac{1}{4}z_1^2 {p_{11}}^2 + z_1^2 p_{12}|\phi_{121}(x_1)| \\ &\qquad + \delta\overline{\gamma}(|x_1|)\,x_1^2 \\ &\quad\le \phi_{11}^2(|\eta|) + p^* z_1^2\left(\frac{1}{4} + |\phi_{121}(x_1)| + \overline{\gamma}(|x_1|)\right) \\ &\quad\le \phi_{11}^2(|\eta|) + z_1^2\widehat{\phi}_1(x_1)\,p^*\end{aligned} \tag{15}$$

with a new smooth function $\widehat{\phi}_1(x_1) \ge 1/4 + |\phi_{121}(x_1)| + \overline{\gamma}(|x_1|) > 0$. As a result, there holds

$$\begin{aligned}\dot{V}_1 &\le -(1-\epsilon)\,\rho\circ\underline{\alpha}(|\eta|)\,\alpha_0(|\eta|) + d_1(t)\,z_1\vartheta_1 \\ &\quad + d_1(t)\,z_1 z_2 + \phi_{11}^2(|\eta|) + z_1^2\widehat{\phi}_1(x_1)\,p^*.\end{aligned} \tag{16}$$

Considering the unknown constant p^* in (16), we use an adaptive signal $\widehat{p}$ to estimate p^*. Consequently, we augment V_1 with the parameter estimation error $\widetilde{p} = p^* - \widehat{p}$, such as

$$\overline{V}_1 = V_1 + \frac{1}{2\Upsilon}\widetilde{p}^2, \tag{17}$$

where $\Upsilon>0$ is the design parameter. In view of (16) and $p^* = \widehat{p} + \widetilde{p}$, a direct substitution leads to

$$\begin{aligned}\dot{\overline{V}}_1 &\le -(1-\epsilon)\,\rho\circ\underline{\alpha}(|\eta|)\,\alpha_0(|\eta|) + d_1(t)\,z_1\vartheta_1 \\ &\quad + z_1^2\widehat{\phi}_1(x_1)\,\widehat{p} + d_1(t)\,z_1 z_2 + \phi_{11}^2(|\eta|) \\ &\quad + \frac{1}{\Upsilon}\widetilde{p}\left(\Upsilon\widehat{\phi}_1(x_1)\,z_1^2 - \dot{\widehat{p}}\right).\end{aligned} \tag{18}$$

Considering Assumption 4, we take the virtual control

$$\vartheta_1 = -\frac{1}{\underline{d}_1}\left(\nu_1 + \widehat{\phi}_1(x_1)\,\widehat{p}\right)z_1, \tag{19}$$

where $\nu_1>0$ is a design constant to be determined later. Let

$$\tau_1 = \Upsilon\widehat{\phi}_1(x_1)\,z_1^2, \tag{20}$$

and then we get

$$\begin{aligned}\dot{\overline{V}}_1 &\le -(1-\epsilon)\,\rho\circ\underline{\alpha}(|\eta|)\,\alpha_0(|\eta|) - \nu_1 z_1^2 + \phi_{11}^2(|\eta|) \\ &\quad + d_1(t)\,z_1 z_2 + \frac{1}{\Upsilon}\widetilde{p}\left(\tau_1 - \dot{\widehat{p}}\right).\end{aligned} \tag{21}$$

Remark 9. It is noted that, in (19), we assume that $\widehat{p}(t)\ge 0$. In fact, from the updating law of $\dot{\widehat{p}}$ given later, this property can be guaranteed by choosing the initial condition $\widehat{p}(0)\ge 0$. Alternatively, using the idea in [20], we also can apply the $\widehat{p}\le\sqrt{1+\widehat{p}^2}$ or $\widehat{p}\le(1+\widehat{p}^2)/2$ instead of $\widehat{p}$.

Step 2. Let $z_3 = x_3 - \vartheta_2$, where ϑ_2 is the virtual control law. We consider the Lyapunov function

$$V_2 = \overline{V}_1 + \frac{1}{2}z_2^2. \tag{22}$$

In view of (21), we have

$$\begin{aligned}\dot{V}_2 \le & -(1-\epsilon)\rho\circ\underline{\alpha}(|\eta|)\alpha_0(|\eta|) - \nu_1 z_1^2 + \phi_{11}^2(|\eta|) \\ & + d_2(t)z_2z_3 + \frac{1}{\Upsilon}\tilde{p}\left(\tau_1 - \dot{\hat{p}}\right) + z_2\Bigg(d_2(t)\vartheta_2 \\ & + g_2(t,\eta,x) - \frac{\partial\vartheta_1}{\partial x_1}g_1(t,\eta,x) + d_1(t)z_1 \\ & - \frac{\partial\vartheta_1}{\partial x_1}d_1(t)x_2 - \frac{\partial\vartheta_1}{\partial\hat{p}}\dot{\hat{p}}\Bigg).\end{aligned} \tag{23}$$

From (4) and (11), the following calculations hold:

$$\begin{aligned}z_2g_2(t,\eta,x) \le & |z_2|p_{21}\phi_{21}(|\eta|) \\ & + |z_2|p_{22}|x_1|\phi_{121}(\overline{x}_2) \\ & + |z_2|p_{22}(|z_2+\vartheta_1|)\phi_{122}(\overline{x}_2).\end{aligned} \tag{24}$$

Like the calculations in (14), by completing the squares, we have

$$\begin{aligned}& -z_2\frac{\partial\vartheta_1}{\partial x_1}g_1(t,\eta,x) \\ & \le \phi_{11}^2(|\eta|) + z_1^2 \\ & \quad + p^*z_2^2\left(\frac{1}{4}\left(\frac{\partial\vartheta_1}{\partial x_1}\right)^2 + \frac{1}{4}\left(\frac{\partial\vartheta_1}{\partial x_1}\right)^2\phi_{121}^2(x_1)\right).\end{aligned} \tag{25}$$

Define $\hat{\phi}_{21}(x_1,\hat{p}) = (1/4)(\partial\vartheta_1/\partial x_1)^2(1+\phi_{121}^2(x_1))$, and then we have

$$\begin{aligned}-z_2\frac{\partial\vartheta_1}{\partial x_1}g_1(t,\eta,x) \le & \phi_{11}^2(|\eta|) + z_1^2 \\ & + z_2^2\hat{\phi}_{21}(x_1,\hat{p})p^*.\end{aligned} \tag{26}$$

In the same manner, using the completion of squares again, it can be verified that

$$\begin{aligned}& |z_2|p_{21}\phi_{21}(|\eta|) \le \phi_{21}^2(|\eta|) + \frac{1}{4}p_{21}^2z_2^2, \\ & |z_2|p_{22}|z_1|\phi_{121}(\overline{x}_2) \le z_1^2 + \frac{1}{4}p_{22}^2z_2^2\phi_{121}^2(\overline{x}_2), \\ & |z_2|p_{22}(|z_2+\vartheta_1|)\phi_{122}(\overline{x}_2) \\ & \quad \le z_2^2p_{22}\phi_{122}(\overline{x}_2) + |z_2|p_{22}|\vartheta_1|\phi_{122}(\overline{x}_2) \\ & \quad \le z_2^2p_{22}\phi_{122}(\overline{x}_2) \\ & \qquad + |z_2|p_{22}\left|-\frac{1}{\underline{d}_1}\left(\nu_1 + \hat{\phi}_1(x_1)\hat{p}\right)z_1\right|\phi_{122}(\overline{x}_2) \\ & \quad \le z_2^2p_{22}\phi_{122}(\overline{x}_2) + z_1^2 \\ & \qquad + \frac{1}{4}p_{22}^2z_2^2\frac{1}{\underline{d}_1^2}\left(\nu_1 + \hat{\phi}_1(x_1)\hat{p}\right)^2\phi_{122}^2(\overline{x}_2).\end{aligned} \tag{27}$$

As a result,

$$\begin{aligned}z_2g_2(t,\eta,x) \le & \phi_{21}^2(|\eta|) + 2z_1^2 + z_2^2p^*\Bigg(\frac{1}{4} \\ & + \frac{1}{4}\phi_{121}^2(\overline{x}_2) + \phi_{122}(\overline{x}_2) \\ & + \frac{1}{4}\frac{1}{\underline{d}_1^2}\left(\nu_1 + \hat{\phi}_1(x_1)\hat{p}\right)^2\phi_{122}^2(\overline{x}_2)\Bigg).\end{aligned} \tag{28}$$

Define $\hat{\phi}_{22}(\overline{x}_2,\hat{p}) = 1/4 + (1/4)\phi_{121}^2(\overline{x}_2) + \phi_{122}(\overline{x}_2) + (1/4)(1/\underline{d}_1^2)(\nu_1 + \hat{\phi}_1(x_1)\hat{p})^2\phi_{122}^2(\overline{x}_2)$, and then

$$z_2g_2(t,\eta,x) \le \phi_{21}^2(|\eta|) + 2z_1^2 + z_2^2p^*\hat{\phi}_{22}(\overline{x}_2,\hat{p}). \tag{29}$$

Let $\hat{\phi}_2(\overline{x}_2,\hat{p}) = \hat{\phi}_{21}(x_1,\hat{p}) + \hat{\phi}_{22}(\overline{x}_2,\hat{p}) > 0$, and then we get

$$\begin{aligned}& z_2\left(g_2(t,\eta,x) - \frac{\partial\vartheta_1}{\partial x_1}g_1(t,\eta,x)\right) \\ & \quad \le \sum_{j=1}^{2}\phi_{j1}^2(|\eta|) + 3z_1^2 + z_2^2\hat{\phi}_2(\overline{x}_2,\hat{p})p^*.\end{aligned} \tag{30}$$

From Assumption 4, it is deduced that

$$\begin{aligned}& z_2\left(d_1(t)z_1 - \frac{\partial\vartheta_1}{\partial x_1}d_1(t)x_2\right) \\ & \quad \le |z_2||d_1(t)||z_1| + |z_2|\left|\frac{\partial\vartheta_1}{\partial x_1}\right||d_1(t)|(|z_2| + |\vartheta_1|) \\ & \quad \le z_1^2 + \frac{\overline{d}_1^2}{4}z_2^2 + z_2^2\frac{\overline{d}_1}{2}\left(1 + \left(\frac{\partial\vartheta_1}{\partial x_1}\right)^2\right) + z_1^2 \\ & \qquad + \frac{1}{4}z_2^2\left(\frac{\partial\vartheta_1}{\partial x_1}\right)^2\frac{\overline{d}_1^2}{\underline{d}_1^2}\left(\nu_1 + \hat{\phi}_1(x_1)\hat{p}\right)^2.\end{aligned} \tag{31}$$

Define $\overline{\phi}_{21}(\overline{x}_2,\hat{p}) = \overline{d}_1^2/4 + (1/4)(\overline{d}_1^2/\underline{d}_1^2)(\partial\vartheta_1/\partial x_1)^2(\nu_1 + \hat{\phi}_1(x_1)\hat{p})^2 + (\overline{d}_1/2)(1 + (\partial\vartheta_1/\partial x_1)^2)$, and there holds

$$z_2\left(d_1(t)z_1 - \frac{\partial\vartheta_1}{\partial x_1}d_1(t)x_2\right) \le 2z_1^2 + z_2^2\overline{\phi}_{21}(\overline{x}_2,\hat{p}). \tag{32}$$

Take the following notation:

$$\tau_2 = \tau_1 + \Upsilon z_2^2\hat{\phi}_2(\overline{x}_2,\hat{p}), \tag{33}$$

and furthermore, in view of (30) and (32), we obtain

$$\begin{aligned}\dot{V}_2 \le & -(1-\epsilon)\rho\circ\underline{\alpha}(|\eta|)\alpha_0(|\eta|)-(\nu_1-5)z_1^2\\ & +2\phi_{11}^2(|\eta|)+\phi_{21}^2(|\eta|)+d_2(t)z_2z_3+z_2\Big(d_2(t)\vartheta_2\\ & +z_2\widehat{\phi}_2(\overline{x}_2,\widehat{p})\widehat{p}+z_2\overline{\phi}_{21}(\overline{x}_2,\widehat{p})-\frac{\partial\vartheta_1}{\partial\widehat{p}}\dot{\widehat{p}}\Big)+\frac{1}{\Upsilon}\\ & \cdot\widetilde{p}\left(\tau_2-\dot{\widehat{p}}\right).\end{aligned} \tag{34}$$

Considering $d_2(t)$ is unknown, the term of $-(\partial\vartheta_1/\partial\widehat{p})\dot{\widehat{p}}$ can not be canceled. In fact, we express (34) as

$$\begin{aligned}\dot{V}_2 \le & -(1-\epsilon)\rho\circ\underline{\alpha}(|\eta|)\alpha_0(|\eta|)-(\nu_1-5)z_1^2\\ & +2\phi_{11}^2(|\eta|)+\phi_{21}^2(|\eta|)+d_2(t)z_2z_3+z_2\Big(d_2(t)\vartheta_2\\ & +z_2\widehat{\phi}_2(\overline{x}_2,\widehat{p})\widehat{p}+z_2\overline{\phi}_{21}(\overline{x}_2,\widehat{p})-\frac{\partial\vartheta_1}{\partial\widehat{p}}\tau_2\Big)+\Big(\frac{1}{\Upsilon}\widetilde{p}\\ & +z_2\frac{\partial\vartheta_1}{\partial\widehat{p}}\Big)\left(\tau_2-\dot{\widehat{p}}\right).\end{aligned} \tag{35}$$

For the term of $-z_2(\partial\vartheta_1/\partial\widehat{p})\tau_2$, according to (33), it can be dealt with as follows:

$$\begin{aligned}-z_2\frac{\partial\vartheta_1}{\partial\widehat{p}}\tau_2 &= -z_2\frac{\partial\vartheta_1}{\partial\widehat{p}}\Upsilon\left(z_1^2\widehat{\phi}_1(x_1)+z_2^2\widehat{\phi}_2(\overline{x}_2,\widehat{p})\right)\\ &\le z_1^2+\frac{1}{4}\Upsilon^2z_2^2\left(\frac{\partial\vartheta_1}{\partial\widehat{p}}\right)^2z_1^2\widehat{\phi}_1^2(x_1)+z_2^2\frac{1}{4}\Upsilon\left(1+z_2^2\right)\\ &\quad\cdot\widehat{\phi}_2(\overline{x}_2,\widehat{p})\left(1+\left(\frac{\partial\vartheta_1}{\partial\widehat{p}}\right)^2\right)=z_1^2\\ &\quad+z_2^2\left(\frac{1}{4}\Upsilon^2\left(\frac{\partial\vartheta_1}{\partial\widehat{p}}\right)^2z_1^2\widehat{\phi}_1^2(x_1)\right.\\ &\quad\left.+\frac{1}{4}\Upsilon\left(1+z_2^2\right)\widehat{\phi}_2(\overline{x}_2,\widehat{p})\left(1+\left(\frac{\partial\vartheta_1}{\partial\widehat{p}}\right)^2\right)\right).\end{aligned} \tag{36}$$

Define the following smooth function: $\overline{\phi}_{22}(\overline{x}_2,\widehat{p}) = (1/4)\Upsilon^2z_1^2(\partial\vartheta_1/\partial\widehat{p})^2\widehat{\phi}_1^2(x_1)+(1/4)\Upsilon(1+z_2^2)\widehat{\phi}_2(\overline{x}_2,\widehat{p})(1+(\partial\vartheta_1/\partial\widehat{p})^2)$, and we get

$$-z_2\frac{\partial\vartheta_1}{\partial\widehat{p}}\tau_2\le z_1^2+z_2^2\overline{\phi}_{22}(\overline{x}_2,\widehat{p}). \tag{37}$$

Denote $\overline{\phi}_2(\overline{x}_2,\widehat{p})=\overline{\phi}_{21}(\overline{x}_2,\widehat{p})+\overline{\phi}_{22}(\overline{x}_2,\widehat{p})>0$, and a direct substitution leads to

$$\begin{aligned}\dot{V}_2 \le & -(1-\epsilon)\rho\circ\underline{\alpha}(|\eta|)\alpha_0(|\eta|)-(\nu_1-6)z_1^2\\ & +2\phi_{11}^2(|\eta|)+\phi_{21}^2(|\eta|)+d_2(t)z_2z_3\\ & +z_2\left(d_2(t)\vartheta_2+z_2\widehat{\phi}_2(\overline{x}_2,\widehat{p})\widehat{p}+z_2\overline{\phi}_2(\overline{x}_2,\widehat{p})\right)\\ & +\left(\frac{1}{\Upsilon}\widetilde{p}+z_2\frac{\partial\vartheta_1}{\partial\widehat{p}}\right)\left(\tau_2-\dot{\widehat{p}}\right).\end{aligned} \tag{38}$$

Take the virtual control

$$\vartheta_2=-\frac{1}{\underline{d}_2}\left(\nu_2+\widehat{\phi}_2(\overline{x}_2,\widehat{p})\widehat{p}+\overline{\phi}_2(\overline{x}_2,\widehat{p})\right)z_2, \tag{39}$$

which is such that

$$\begin{aligned}\dot{V}_2 \le & -(1-\epsilon)\rho\circ\underline{\alpha}(|\eta|)\alpha_0(|\eta|)-(\nu_1-6)z_1^2-\nu_2z_2^2\\ & +2\phi_{11}^2(|\eta|)+\phi_{21}^2(|\eta|)+d_2(t)z_2z_3\\ & +\left(\frac{1}{\Upsilon}\widetilde{p}+z_2\frac{\partial\vartheta_1}{\partial\widehat{p}}\right)\left(\tau_2-\dot{\widehat{p}}\right).\end{aligned} \tag{40}$$

Step i $(3\le i\le n)$. Assume that, in Step $i-1$, we have designed the virtual control ϑ_j $(j=1,\ldots,i-1)$ and the tuning function τ_j $(j=1,\ldots,i-1)$, such that the Lyapunov function

$$V_{i-1}=\sum_{j=1}^{i-1}\frac{1}{2}z_j^2+\frac{1}{2\Upsilon}\widetilde{p}^2 \tag{41}$$

satisfies

$$\begin{aligned}\dot{V}_{i-1} \le & -(1-\epsilon)\rho\circ\underline{\alpha}(|\eta|)\alpha_0(|\eta|)\\ & -\left(\nu_1-6-\sum_{j=3}^{i-1}(2j+1)\right)z_1^2\\ & -\left(\nu_2-7-\sum_{j=4}^{i-1}(2j)\right)z_2^2-\cdots\\ & -(\nu_{i-2}-7)z_{i-2}^2-\nu_{i-1}z_{i-1}^2\\ & +\left(\frac{1}{\Upsilon}\widetilde{p}+\sum_{j=2}^{i-1}z_j\frac{\partial\vartheta_{j-1}}{\partial\widehat{p}}\right)\left(\tau_{i-1}-\dot{\widehat{p}}\right)\\ & +d_{i-1}(t)z_{i-1}z_i+\sum_{j=1}^{i-1}(i-j)\phi_{j1}^2(|\eta|).\end{aligned} \tag{42}$$

In what follows, it will be shown that the property (42) also holds in Step i.

Let ϑ_i be the virtual control and $z_{i+1}=x_{i+1}-\vartheta_i$. Consider the Lyapunov function

$$V_i=V_{i-1}+\frac{1}{2}z_i^2. \tag{43}$$

To begin with, the dynamics of z_i can be expressed as

$$\begin{aligned}\dot{z}_i=d_i(t)\vartheta_i+g_i(t,\eta,x)-\sum_{j=1}^{i-1}\frac{\partial\vartheta_{i-1}}{\partial x_j}g_j(t,\eta,x)\\ -\sum_{j=1}^{i-1}\frac{\partial\vartheta_{i-1}}{\partial x_j}d_j(t)x_{j+1}-\frac{\partial\vartheta_{i-1}}{\partial\widehat{p}}\dot{\widehat{p}}+d_i(t)z_{i+1}.\end{aligned} \tag{44}$$

For notational convenience, denote $\vartheta_0=0$. From $x_k=z_k+\vartheta_{k-1}=z_k-(1/\underline{d}_{k-1})(\nu_{k-1}+\widehat{\phi}_{k-1}(x_{[k-1]},\widehat{p})\widehat{p}+$

$\overline{\phi}_i(x_{[k-1]}, \widehat{p}))z_{k-1}$ $(k = 1, \ldots, i-1)$, like the calculations in (26), it can be verified that there exist smooth functions $\widehat{\phi}_{ij}(\cdot)$ $(j = 1, \ldots, i) > 0$, such that, for $j = 1, \ldots, i-1$,

$$\begin{aligned} -z_i \frac{\partial \vartheta_{i-1}}{\partial x_j} g_j(t, \eta, x) &\le \phi_{j1}^2(|\eta|) + \sum_{k=1}^{j-1}\left(2z_k^2\right) + z_j^2 \\ &\quad + z_i^2 \widehat{\phi}_{ij}\left(\overline{x}_j, \widehat{p}\right) p^*, \\ z_i g_i(t, \eta, x) &\le \phi_{i1}^2(|\eta|) \\ &\quad + \sum_{k=1}^{i-1}\left(2z_k^2\right) + z_i^2 \widehat{\phi}_{ii}(\overline{x}_i, \widehat{p}) p^*. \end{aligned} \tag{45}$$

Define $\widehat{\phi}_i(\overline{x}_i, \widehat{p}) = \sum_{j=1}^{i} \widehat{\phi}_{ij}(\overline{x}_j, \widehat{p})$; from (45), it follows that

$$\begin{aligned} & z_i\left(g_i(t, \eta, x) - \sum_{j=1}^{i-1} \frac{\partial \vartheta_{i-1}}{\partial x_j} g_j(t, \eta, x)\right) \\ &\quad \le \sum_{j=1}^{i} \phi_{j1}^2(|\eta|) \\ &\quad + \sum_{j=1}^{i-1}(1 + 2(i-j)) z_j^2 + z_i^2 \widehat{\phi}_i(\overline{x}_i, \widehat{p}) p^*. \end{aligned} \tag{46}$$

Similar to (31) and (32), there exists a smooth function $\overline{\phi}_{i1}(\overline{x}_i, \widehat{p})$ such that

$$\begin{aligned} & z_i\left(d_{i-1}(t) z_{i-1} - \sum_{j=1}^{i-1} \frac{\partial \vartheta_{i-1}}{\partial x_j} d_j(t) x_{j+1}\right) \\ &\quad \le z_1^2 + \sum_{k=2}^{i-2}\left(2z_k^2\right) + 3z_{i-1}^2 + z_i^2 \overline{\phi}_{i1}(\overline{x}_i, \widehat{p}). \end{aligned} \tag{47}$$

Let

$$\tau_i = \tau_{i-1} + \Upsilon z_i^2 \widehat{\phi}_i(\overline{x}_i, \widehat{p}), \tag{48}$$

and then, from (47) and (48), there holds

$$\begin{aligned} \dot{V}_i &\le -(1-\epsilon)\rho \circ \underline{\alpha}(|\eta|)\alpha_0(|\eta|) - \Bigg(\nu_1 - 6 \\ &\quad - \sum_{j=3}^{i-1}(2j+1) - 2i\Bigg) z_1^2 - \Bigg(\nu_2 - 7 \\ &\quad - \sum_{j=4}^{i-1}(2j) - 1 - 2(i-1)\Bigg) z_2^2 - \cdots - (\nu_{i-2} - 7 - 7) \\ &\quad \cdot z_{i-2}^2 - (\nu_{i-1} - 6) z_{i-1}^2 + \sum_{j=1}^{i}(i-j+1)\phi_{j1}^2(|\eta|) \\ &\quad + d_i(t) z_i z_{i+1} + z_i\Bigg(d_i(t)\vartheta_i + z_i \widehat{\phi}_i(\overline{x}_i, \widehat{p})\widehat{p} \\ &\quad + z_i \overline{\phi}_{i1}(\overline{x}_i, \widehat{p}) \\ &\quad - \sum_{j=2}^{i-1} z_j \frac{\partial \vartheta_{j-1}}{\partial \widehat{p}} \cdot \Upsilon z_i \widehat{\phi}_i(\overline{x}_i, \widehat{p}) - \frac{\partial \vartheta_{i-1}}{\partial \widehat{p}}\tau_i\Bigg) + \Bigg(\frac{1}{\Upsilon}\widetilde{p} \\ &\quad + \sum_{j=2}^{i} z_j \frac{\partial \vartheta_{j-1}}{\partial \widehat{p}}\Bigg)\left(\tau_i - \dot{\widehat{p}}\right). \end{aligned} \tag{49}$$

Remark 10. In (49), we subtract two terms $-\sum_{j=2}^{i-1} z_j(\partial\vartheta_{j-1}/\partial\widehat{p})\cdot\Upsilon z_i\widehat{\phi}_i(\overline{x}_i, \widehat{p})$ and $-(\partial\vartheta_{i-1}/\partial\widehat{p})\tau_i$ in the brackets to generate the term $((1/\Upsilon)\widetilde{p} + \sum_{j=2}^{i} z_j(\partial\vartheta_{j-1}/\partial\widehat{p}))(\tau_i - \dot{\widehat{p}})$.

However, from (49), it can be seen that, due to the unknown control coefficient $d_i(t)$, the terms $-\sum_{j=2}^{i-1} z_j(\partial\vartheta_{j-1}/\partial\widehat{p}) \cdot z_i\widehat{\phi}_i(\overline{x}_i, \widehat{p})$ and $-(\partial\vartheta_{i-1}/\partial\widehat{p})\tau_i$ could not be directly canceled by the coming virtual control ϑ_i. We get around this burden by the following estimates:

$$\begin{aligned} & -z_i \sum_{j=2}^{i-1} z_j \frac{\partial \vartheta_{j-1}}{\partial \widehat{p}} \cdot \Upsilon z_i \widehat{\phi}_i(\overline{x}_i, \widehat{p}) \\ &\quad \le z_i^2 \frac{1}{2}\left(1 + \Upsilon^2\left(\sum_{j=2}^{i-1} z_j \frac{\partial \vartheta_{j-1}}{\partial \widehat{p}}\right)^2 \widehat{\phi}_i^2(\overline{x}_i, \widehat{p})\right), \\ & -z_i \frac{\partial \vartheta_{i-1}}{\partial \widehat{p}}\tau_i = -z_i \frac{\partial \vartheta_{i-1}}{\partial \widehat{p}} \Upsilon\left(z_1^2 \widehat{\phi}_1 + \cdots + z_i^2 \widehat{\phi}_i\right) \\ &\quad \le \sum_{j=1}^{i-1} z_j^2 + z_i^2 \frac{1}{4}\Upsilon^2\left(\frac{\partial \vartheta_{i-1}}{\partial \widehat{p}}\right)^2 \sum_{j=1}^{i} z_j^2 \widehat{\phi}_j^2(\overline{x}_j, \widehat{p}). \end{aligned} \tag{50}$$

Let $\overline{\phi}_i(\overline{x}_i, \widehat{p}) = (1/4)\Upsilon^2(\partial\vartheta_{i-1}/\partial\widehat{p})^2 \sum_{j=1}^{i} z_j^2 \widehat{\phi}_j^2(\overline{x}_j, \widehat{p}) + (1/2)(1 + \Upsilon^2(\sum_{j=2}^{i-1} z_j(\partial\vartheta_{j-1}/\partial\widehat{p}))^2 \widehat{\phi}_i^2(\overline{x}_i, \widehat{p})) + \overline{\phi}_{i1}(\overline{x}_i, \widehat{p}) > 0$, and then we get

$$\begin{aligned} & z_i\Bigg(z_i \overline{\phi}_{i1}(\overline{x}_i, \widehat{p}) - \sum_{j=2}^{i-1} z_j \frac{\partial \vartheta_{j-1}}{\partial \widehat{p}} \cdot z_i \widehat{\phi}_i(\overline{x}_i, \widehat{p}) \\ &\quad - \frac{\partial \vartheta_{i-1}}{\partial \widehat{p}}\tau_i\Bigg) \le \sum_{j=1}^{i-1} z_j^2 + z_i^2 \overline{\phi}_i(\overline{x}_i, \widehat{p}). \end{aligned} \tag{51}$$

By substituting (51) into (49), it follows that

$$\begin{aligned} \dot{V}_i &\le -(1-\epsilon)\rho \circ \underline{\alpha}(|\eta|)\alpha_0(|\eta|) \\ &\quad - \left(\nu_1 - 6 - \sum_{j=3}^{i}(2j+1)\right) z_1^2 \\ &\quad - \left(\nu_2 - 7 - \sum_{j=4}^{i}(2j)\right) z_2^2 - \cdots \\ &\quad - (\nu_{i-2} - 7 - 8) z_{i-2}^2 - (\nu_{i-1} - 7) z_{i-1}^2 \\ &\quad + z_i\left(d_i(t)\vartheta_i + z_i \widehat{\phi}_i(\overline{x}_i, \widehat{p})\widehat{p} + z_i \overline{\phi}_i(\overline{x}_i, \widehat{p})\right) \end{aligned}$$

$$+\left(\frac{1}{\Upsilon}\widetilde{p}+\sum_{j=2}^{i}z_j\frac{\partial\vartheta_{j-1}}{\partial\widehat{p}}\right)\left(\tau_i-\dot{\widehat{p}}\right)$$
$$+\sum_{j=1}^{i}(i-j+1)\,\phi_{j1}^2(|\eta|)+d_i(t)\,z_iz_{i+1}. \tag{52}$$

Take the virtual control

$$\vartheta_i=-\frac{1}{\underline{d}_i}\left(\nu_i+\widehat{\phi}_i(\overline{x}_i,\widehat{p})\,\widehat{p}+\overline{\phi}_i(\overline{x}_i,\widehat{p})\right)z_i, \tag{53}$$

and then the following holds:

$$\begin{aligned}\dot{V}_i\le&-(1-\epsilon)\,\rho\circ\underline{\alpha}(|\eta|)\,\alpha_0(|\eta|)\\&-\left(\nu_1-6-\sum_{j=3}^{i}(2j+1)\right)z_1^2\\&-\left(\nu_2-7-\sum_{j=4}^{i}(2j)\right)z_2^2-\cdots-(\nu_{i-1}-7)\,z_{i-1}^2\\&-\nu_iz_i^2+\sum_{j=1}^{i}(i-j+1)\,\phi_{j1}^2(|\eta|)\\&+\left(\frac{1}{\Upsilon}\widetilde{p}+\sum_{j=2}^{i}z_j\frac{\partial\vartheta_{j-1}}{\partial\widehat{p}}\right)\left(\tau_i-\dot{\widehat{p}}\right)+d_i(t)\,z_iz_{i+1}.\end{aligned} \tag{54}$$

In particular, when $i=n$, the actual control u appears, and we choose the controller u and updating law τ_n for $\widehat{p}(t)$ as follows:

$$u=-\frac{1}{\underline{d}_n}\left(\nu_n+\widehat{\phi}_n(x,\widehat{p})\,\widehat{p}+\overline{\phi}_n(x,\widehat{p})\right)z_n, \tag{55}$$

$$\dot{\widehat{p}}=\tau_n=\sum_{j=1}^{n}\Upsilon z_j^2\widehat{\phi}_j\left(\overline{x}_j,\widehat{p}\right), \tag{56}$$

such that the Lyapunov function

$$V_n=\sum_{j=1}^{n}\frac{1}{2}z_j^2+\frac{1}{2\Upsilon}\widetilde{p}^2 \tag{57}$$

satisfies

$$\begin{aligned}\dot{V}_n\le&-(1-\epsilon)\,\rho\circ\underline{\alpha}(|\eta|)\,\alpha_0(|\eta|)\\&-\left(\nu_1-6-\sum_{j=3}^{n}(2j+1)\right)z_1^2\\&-\left(\nu_2-7-\sum_{j=4}^{n}(2j)\right)z_2^2-\cdots-(\nu_{n-1}-7)\,z_{n-1}^2\\&-\nu_nz_n^2+\sum_{j=1}^{n}(n-j+1)\,\phi_{j1}^2(|\eta|).\end{aligned} \tag{58}$$

This completes the controller design procedure.

4. Main Results

After the above controller design procedure, we are now ready to state the main results.

Theorem 11. *Suppose the investigated system (1) satisfies Assumptions 1, 3, and 4 together with the local conditions*

$$\phi_{i1}^2(s)=O(\alpha_0(s)),\quad i=1,\ldots,n. \tag{59}$$

Then all the signals of the closed-loop system (1) with the controller (55) and updating law (56) are bounded on $[0,+\infty)$. *Specifically, the following convergent property holds:*

$$\lim_{t\to\infty}\left(|\eta(t)|+|x(t)|+|u(t)|\right)=0. \tag{60}$$

Proof. From the local conditions (59), one can choose the smooth function $\rho(\cdot)$ such that

$$\sum_{j=1}^{n}(n-j+1)\,\phi_{j1}^2(|\eta|)\le\frac{1-\epsilon}{2}\rho\circ\underline{\alpha}(|\eta|)\,\alpha_0(|\eta|). \tag{61}$$

One can choose positive constants ν_j $(j=1,\ldots,n)$ satisfying

$$\begin{aligned}&\nu_1-6-\sum_{j=3}^{n}(2j+1)\ge1,\\&\nu_2-7-\sum_{j=4}^{n}(2j)\ge1,\\&\vdots\\&\nu_{n-1}-7\ge1,\\&\nu_n\ge1.\end{aligned} \tag{62}$$

Then, from (58), (61), and (62), it follows that

$$\dot{V}_n\le-\frac{1-\epsilon}{2}\rho\circ\underline{\alpha}(|\eta|)\,\alpha_0(|\eta|)-\sum_{i=1}^{n}z_i^2. \tag{63}$$

Now, assume that the maximal interval of existence of the solution of the closed-loop system starting from any given initial conditions is $[0,t_f)$ for some $t_f>0$. In view of $0<\epsilon<1$, from (63), it can be concluded that V_n and hence the variables $(\eta(t),z_1(t),\ldots,z_n(t),\widetilde{p}(t))$ are bounded on $[0,t_f)$. In terms of $\widetilde{p}(t)=p^*-\widehat{p}(t)$, we obtain the boundedness of $\widehat{p}(t)$. Considering (53), it can be derived that ϑ_i $(i=1,\ldots,n)$ are bounded. In view of $z_i=x_i-\vartheta_{i-1}(\vartheta_0=0)$, we further obtain that the states x_i $(i=1,\ldots,n)$ are bounded on $[0,t_f)$.

So far all the closed-loop system signals are bounded on $[0,t_f)$. This guarantees that the finite time escape will not happen. Therefore, it is natural that t_f can be maximized to $+\infty$ by means of Theorem 3.3 in [21]. Next we will prove the convergence property of (60).

Again, according to (63), considering $\underline{\alpha}$ and α_0 are K_∞-functions, a direct application of LaSalle's invariance

principle in [21] guarantees the convergence property of $(\eta(t), z_1(t), \ldots, z_n(t))$; that is,

$$\lim_{t\to\infty} \eta(t) = 0, \qquad \lim_{t\to\infty} z_i(t) = 0 \quad (i = 1, \ldots, n). \tag{64}$$

As a consequence, from (53) and (55), the following holds:

$$\lim_{t\to\infty} \vartheta_i(t) = 0 \quad (i = 1, \ldots, n). \tag{65}$$

Particularly,

$$\lim_{t\to\infty} u(t) = 0. \tag{66}$$

In terms of (64), (65), and $x_i = z_i + \vartheta_{i-1}$ $(i = 1, \ldots, n)$, we can obtain

$$\lim_{t\to\infty} x_i(t) = 0 \quad (i = 1, \ldots, n). \tag{67}$$

This completes the proof. □

It is noted that, under some stronger conditions, the designed control law can be a linear controller. In fact, we have the following statement.

Theorem 12. *Suppose that the conditions for Theorem 11 are satisfied with δ_0 known a priori and the following additional assumptions hold:*

(i) *The uncertain functions $\phi_{i1}(\cdot)$ $(i = 1, \ldots, n)$ satisfy*

$$\limsup_{s\to+\infty} \frac{\phi_{i1}^2(s)}{\alpha_0(s)} < +\infty. \tag{68}$$

(ii) *There exist known constants $p_{ij} > 0$ $(i = 1, \ldots, n; j = 1, 2)$, such that*

$$|g_i(t, \eta, x)| \le p_{i1}\phi_{i1}(|\eta|) + p_{i2}(|x_1| + \cdots + |x_i|), \quad i = 1, \ldots, n. \tag{69}$$

(iii) *$\gamma_0(|s|) = ks^2$ where k is a positive constant.*

Then, the proposed design method can result in a linear control law

$$u = -\frac{1}{\underline{d}_n}\nu_n z_n = -\frac{1}{\underline{d}_n}\nu_n(x_n + \nu_{n-1}(x_{n-1} + \cdots + \nu_1 x_1)), \tag{70}$$

where ν_i $(i = 1, \ldots, n)$ are some sufficiently large positive constants.

Proof. Under the above hypotheses (i)–(iii), it is known that the constant p^* is known, and hence the estimation $\widehat{p}$ for p^* is no longer needed. Moreover, since conditions (59) and (68) are satisfied, the function $\rho(\cdot)$ in (7) can be chosen as a constant $\rho > 0$. For $i = 1, \ldots, n$, we consider the following function:

$$\overline{V}_{0i}(\eta) = \int_0^{V_0(\eta)} \rho_i \, ds, \tag{71}$$

where ρ_i $(i = 1, \ldots, n)$ are design constants. In view of (3), we can get

$$\begin{aligned}\dot{\overline{V}}_{0i}(\eta) &\le -\rho_i\alpha_0(|\eta|) + \rho_i\delta_0\gamma_0(|y|) \\ &= -\phi_{i1}^2(|\eta|) + \widetilde{\rho}_i\delta_0\gamma_0(|y|) + \phi_{i1}^2(|\eta|) \\ &\quad - \rho_i\alpha_0(|\eta|) - \widetilde{\rho}_i\delta_0\gamma_0(|y|) + \rho_i\delta_0\gamma_0(|y|)\end{aligned} \tag{72}$$

with some positive constants $\widetilde{\rho}_i$ $(i = 1, \ldots, n)$. We will prove that if the constants ρ_i and $\widetilde{\rho}_i$ are chosen suitably, the following inequality holds:

$$\phi_{i1}^2(|\eta|) - \rho_i\alpha_0(|\eta|) - \widetilde{\rho}_i\delta_0\gamma_0(|y|) + \rho_i\delta_0\gamma_0(|y|) \le 0, \quad i = 1, \ldots, n. \tag{73}$$

In fact, because of $\phi_{i1}^2(s) = O(\alpha_0(s))$, there exist positive constants $s_i > 0$, $c_i > 0$ satisfying $\phi_{i1}^2(s) \le c_i\alpha_0(s)$, for $s \in [0, s_i]$. Take $\rho_i \ge c_i + 1$, $\widetilde{\rho}_i = \rho_i + 1$, and then $\phi_{i1}^2(s) - \rho_i\alpha_0(s) \le 0$ and $(\rho_i - \widetilde{\rho}_i)\delta_0\gamma_0(|y|) \le 0$. In view of $\limsup_{s\to\infty}(\phi_{i1}^2(s)/\alpha_0(s)) < \infty$, there exist positive constants $s_i' > 0$, $c_i' > 0$ satisfying $\phi_{i1}^2(s) \le c_i'\alpha_0(s)$ if $s \in [s_i', \infty]$. Similarly, if we take $\rho_i \ge c_i' + 1$ and $\widetilde{\rho}_i = \rho_i + 1$, then $\phi_{i1}^2(s) - \rho_i\alpha_0(s) \le 0$, $(\rho_i - \widetilde{\rho}_i)\delta_0\gamma_0(|y|) \le 0$. In the finite closed interval $[s_i, s_i']$, let $\alpha_0(\underline{s})$ be the minimum value of $\alpha_0(s)$ and let $\phi_{i1}^2(\overline{s})$ be the maximum value of $\phi_{i1}^2(s)$, respectively, and if we take $\rho_i \ge \phi_{i1}^2(\overline{s})/\alpha_0(\underline{s})$, $\widetilde{\rho}_i = \rho_i + 1$, then $\phi_{i1}^2(s) - \rho_i\alpha_0(s) \le 0$, $(\rho_i - \widetilde{\rho}_i)\delta_0\gamma_0(|y|) \le 0$. According to the previous analysis, we choose $\rho_i = c_i + 1 + c_i' + 1 + \phi_{i1}^2(\overline{s})/\alpha_0(\underline{s})$ and $\widetilde{\rho}_i = \rho_i + 1$, and then (72) holds. Therefore we get

$$\dot{\overline{V}}_{0i}(\eta) \le -\phi_{i1}^2(|\eta|) + \widetilde{\rho}_i\delta_0\gamma_0(|y|), \quad i = 1, \ldots, n. \tag{74}$$

In view of $\gamma_0(s) = ks^2$, a direct substitution in (74) results in

$$\dot{\overline{V}}_{0i}(\eta) \le -\phi_{i1}^2(|\eta|) + \widetilde{\rho}_i\delta_0 k y^2, \quad i = 1, \ldots, n. \tag{75}$$

To deal with the unmeasured dynamics η in this case, we can choose the candidate Lyapunov function as follows:

$$V_n = \sum_{i=1}^{n} \frac{1}{2} z_i^2 + \sum_{i=1}^{n} (n - i + 2)\overline{V}_{0i}(\eta). \tag{76}$$

Consequently, a modified version of the design procedure in Section 3 leads to the linear control law

$$u = -\frac{1}{\underline{d}_n}\nu_n z_n = -\frac{1}{\underline{d}_n}\nu_n(x_n + \nu_{n-1}(x_{n-1} + \cdots + \nu_1 x_1)) \tag{77}$$

with some sufficiently large positive constants ν_i $(i = 1, \ldots, n)$. □

5. Simulation Example

In this section, we provide a simulation example to illustrate the proposed method in the paper. Consider the following nonlinear systems:

$$\begin{aligned} \dot{\eta} &= -\eta + \delta(t) y^2, \\ \dot{x}_1 &= d_1(t) x_2 + \theta_1 x_1 + \theta_2 \frac{x_1 x_2}{1 + x_2^2}, \\ \dot{x}_2 &= d_2(t) u + \theta_3 \sin(t x_1 x_2) \eta, \\ y &= x_1 \end{aligned} \tag{78}$$

with $q(t,\eta,y) = -\eta + \delta(t)y^2$, $\delta(t) = 1 + e^{-t}$, $g_1(t,\eta,x) = \theta_1 x_1 + \theta_2(x_1 x_2/(1+x_2^2))$, and $g_2(t,\eta,x) = \theta_3 \sin(t x_1 x_2)\eta$. The inverse system $\dot{\eta} = -\eta + \delta(t)y^2$ is ISS, and $V_0(\eta) = (1/2)\eta^2$ is a candidate ISS-Lyapunov function with the supply pair $\alpha_0(|\eta|) = (1/2)\eta^2$, $\delta(t)\gamma_0(|x_1|) = (1+e^{-t})(1/4)x_1^4$.

Next, we use the proposed algorithm in Section 3 to design the partial-state feedback controller.

Step 1. We consider the function $V_1 = (1/2)z_1^2 + (1/2)\eta^2$, whose time derivative satisfies

$$\begin{aligned} \dot{V}_1 &= z_1 \left(d_1(t) x_2 + \theta_1 x_1 + \theta_2 \frac{x_1 x_2}{1 + x_2^2} \right) \\ &\quad + \eta \left(-\eta + \delta(t) y^2 \right). \end{aligned} \tag{79}$$

Like the calculations in (14), we have

$$\begin{aligned} z_1 \theta_1 x_1 &\le |\theta_1| z_1^2, \\ z_1 \theta_2 \frac{x_1 x_2}{1 + x_2^2} &\le \frac{1}{2} |\theta_2| z_1^2, \\ \eta \delta(t) y^2 &\le \frac{1}{4}\eta^2 + \delta^2(t) z_1^4 \le \frac{1}{4}\eta^2 + \delta_0 z_1^4, \end{aligned} \tag{80}$$

where $\delta_0 > 0$ satisfying $\delta^2(t) \le \delta_0$.

Define $p^* = \max\{\delta_0, |\theta_i|, \theta_i^2 \mid i = 1,2,3\}$, and we get

$$\begin{aligned} & z_1 \theta_1 x_1 + z_1 \theta_2 \frac{x_1 x_2}{1 + x_2^2} + \eta \delta(t) y^2 \\ &\quad \le \frac{1}{4}\eta^2 + p^* z_1^2 \left(1 + \frac{1}{2} + z_1^2 \right) \le \frac{1}{4}\eta^2 + z_1^2 \widehat{\phi}_1(x_1) p^* \end{aligned} \tag{81}$$

with a new smooth function $\widehat{\phi}_1(x_1) \ge 1 + 1/2 + z_1^2 > 0$.

Similar to (17), we augment V_1 as follows:

$$\overline{V}_1 = V_1 + \frac{1}{2\Upsilon} \widetilde{p}^2, \tag{82}$$

where $\Upsilon > 0$ is the design parameter. In view of (79) and $p^* = \widehat{p} + \widetilde{p}$, a direct substitution leads to

$$\begin{aligned} \dot{\overline{V}}_1 &\le -\left(1 - \frac{1}{4}\right)\eta^2 + d_1(t) z_1 \vartheta_1 + d_1(t) z_1 z_2 \\ &\quad + z_1^2 \widehat{\phi}_1(x_1) \widehat{p} + \frac{1}{\Upsilon} \widetilde{p} \left(\Upsilon \widehat{\phi}_1(x_1) z_1^2 - \dot{\widehat{p}} \right). \end{aligned} \tag{83}$$

We take the virtual control and the tuning function

$$\begin{aligned} \vartheta_1 &= -\frac{1}{\underline{d}_1} \left(\nu_1 + \widehat{\phi}_1(x_1) \widehat{p} \right) z_1, \quad \nu_1 > 0, \\ \tau_1 &= \Upsilon \widehat{\phi}_1(x_1) z_1^2, \end{aligned} \tag{84}$$

and then we get

$$\begin{aligned} \dot{\overline{V}}_1 &\le -\left(1 - \frac{1}{4}\right)\eta^2 - \nu_1 z_1^2 + d_1(t) z_1 z_2 \\ &\quad + \frac{1}{\Upsilon} \widetilde{p} \left(\tau_1 - \dot{\widehat{p}} \right). \end{aligned} \tag{85}$$

Step 2. To find the actual control law u, we consider the Lyapunov function

$$V_2 = \overline{V}_1 + \frac{1}{2} z_2^2. \tag{86}$$

In view of (78) and (85), we have

$$\begin{aligned} \dot{V}_2 &\le -\left(1 - \frac{1}{4}\right)\eta^2 - \nu_1 z_1^2 + d_1(t) z_1 z_2 + \frac{1}{\Upsilon} \widetilde{p} \left(\tau_1 \right. \\ &\quad \left. - \dot{\widehat{p}} \right) + z_2 \left(d_2(t) u + \theta_3 \sin(t x_1 x_2) \eta \right. \\ &\quad \left. - \frac{\partial \vartheta_1}{\partial x_1} \left(d_1(t) x_2 + \theta_1 x_1 + \theta_2 \frac{x_1 x_2}{1 + x_2^2} \right) - \frac{\partial \vartheta_1}{\partial \widehat{p}} \dot{\widehat{p}} \right). \end{aligned} \tag{87}$$

As in (80), we have

$$\begin{aligned} & d_1(t) z_1 z_2 \le z_1^2 + \frac{\overline{d}_1^2}{4} z_2^2, \\ & z_2 \theta_3 \sin(t x_1 x_2) \eta \le \left| z_2 \theta_3 \sin(t x_1 x_2) \eta \right| \\ & \quad \le \frac{1}{4}\eta^2 + \theta_3^2 z_2^2, \\ & -z_2 \frac{\partial \vartheta_1}{\partial x_1} \theta_1 x_1 \le z_1^2 + \frac{1}{4}\theta_1^2 z_2^2 \left(\frac{\partial \vartheta_1}{\partial x_1} \right)^2, \\ & -z_2 \frac{\partial \vartheta_1}{\partial x_1} \theta_2 \frac{x_1 x_2}{1 + x_2^2} \le z_1^2 + \frac{1}{4}\theta_2^2 z_2^2 \left(\frac{\partial \vartheta_1}{\partial x_1} \right)^2 \frac{1}{4}, \\ & -z_2 \frac{\partial \vartheta_1}{\partial x_1} d_1(t) x_2 \\ & \quad \le z_2^2 \frac{\overline{d}_1}{2} \left(1 + \left(\frac{\partial \vartheta_1}{\partial x_1} \right)^2 \right) + z_1^2 \\ & \quad + \frac{1}{4} z_2^2 \left(\frac{\partial \vartheta_1}{\partial x_1} \right)^2 \frac{\overline{d}_1^2}{\underline{d}_1^2} \left(\nu_1 + \widehat{\phi}_1(x_1) \widehat{p} \right)^2. \end{aligned} \tag{88}$$

Consequently, in view of the definition of p^*, the following holds:

$$\begin{aligned} & z_2 \theta_3 \sin(t x_1 x_2) \eta - z_2 \frac{\partial \vartheta_1}{\partial x_1} \theta_1 x_1 - z_2 \frac{\partial \vartheta_1}{\partial x_1} \theta_2 \frac{x_1 x_2}{1 + x_2^2} \\ & \quad \le \frac{1}{4}\eta^2 + 2 z_1^2 + z_2^2 \widehat{\phi}_2(x_1, \widehat{p}) p^*, \end{aligned} \tag{89}$$

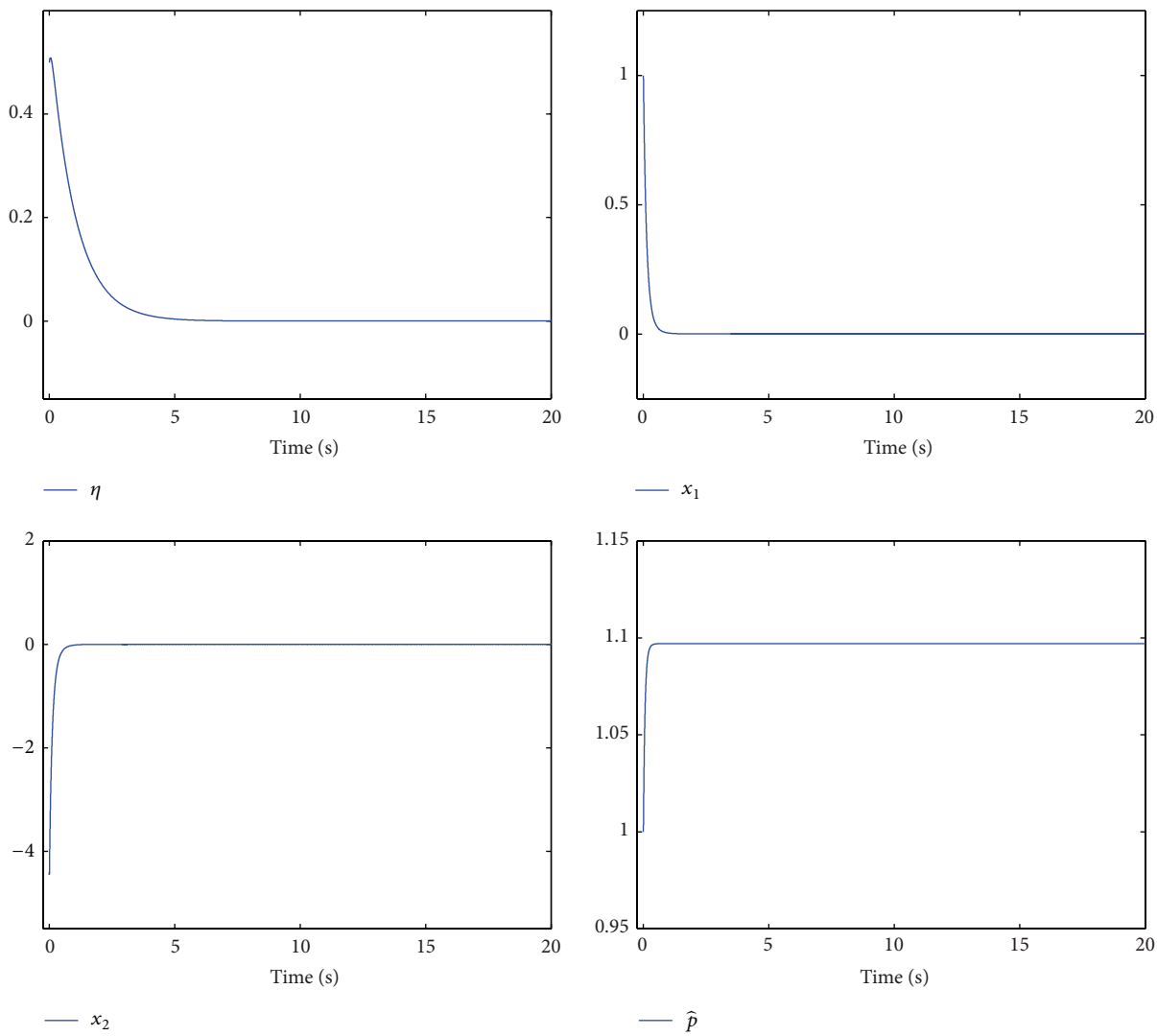


FIGURE 1: The response of the closed-loop system.

with $\widehat{\phi}_2(x_1, \widehat{p}) = 1 + (1/4)(\partial\vartheta_1/\partial x_1)^2 + (1/4)(\partial\vartheta_1/\partial x_1)^2(1/4)$. Define $\overline{\phi}_{21}(x_1, \widehat{p}) = \overline{d}_1^2/4 + (\overline{d}_1/2)(1 + (\partial\vartheta_1/\partial x_1)^2) + (1/4)(\overline{d}_1^2/\underline{d}_1^2)(\partial\vartheta_1/\partial x_1)^2(\nu_1 + \widehat{\phi}_1(x_1)\widehat{p})^2$, and one gets

$$z_2\left(d_1(t)z_1 - \frac{\partial\vartheta_1}{\partial x_1}d_1(t)x_2\right) \le 2z_1^2 + z_2^2\overline{\phi}_{21}(x_1, \widehat{p}). \quad (90)$$

As a result, with $\tau_2 = \tau_1 + \Upsilon z_2^2\widehat{\phi}_2(x_1, \widehat{p})$, we have

$$\begin{aligned}\dot{V}_2 \le &-\left(1 - \frac{2}{4}\right)\eta^2 - (\nu_1 - 4)z_1^2 + z_2^2\widehat{\phi}_2(x_1, \widehat{p})\widehat{p} \\ &+ z_2^2\overline{\phi}_{21}(x_1, \widehat{p}) + \left(\frac{1}{\Upsilon}\widetilde{p} + z_2\frac{\partial\vartheta_1}{\partial\widehat{p}}\right)\left(\tau_2 - \dot{\widehat{p}}\right) \\ &+ z_2 d_2(t)u - z_2\frac{\partial\vartheta_1}{\partial\widehat{p}}\tau_2.\end{aligned} \quad (91)$$

For the term of $-z_2(\partial\vartheta_1/\partial\widehat{p})\tau_2$, according to (36), it can be verified that

$$\begin{aligned}-z_2\frac{\partial\vartheta_1}{\partial\widehat{p}}\tau_2 \le z_1^2 + z_2^2\Bigg(&\frac{1}{4}\Upsilon^2\left(\frac{\partial\vartheta_1}{\partial\widehat{p}}\right)^2 z_1^2\widehat{\phi}_1^2(x_1) \\ &+ \frac{1}{4}\Upsilon\left(1 + z_2^2\right)\widehat{\phi}_2(x_1, \widehat{p})\left(1 + \left(\frac{\partial\vartheta_1}{\partial\widehat{p}}\right)^2\right)\Bigg).\end{aligned} \quad (92)$$

Define $\overline{\phi}_{22}(\overline{x}_2, \widehat{p}) = (1/4)\Upsilon^2 z_1^2(\partial\vartheta_1/\partial\widehat{p})^2\widehat{\phi}_1^2(x_1) + (1/4)\Upsilon(1 + z_2^2)\widehat{\phi}_2(x_1, \widehat{p})(1 + (\partial\vartheta_1/\partial\widehat{p})^2)$, and we get

$$-z_2\frac{\partial\vartheta_1}{\partial\widehat{p}}\tau_2 \le z_1^2 + z_2^2\overline{\phi}_{22}(\overline{x}_2, \widehat{p}). \quad (93)$$

Denote $\overline{\phi}_2(\overline{x}_2, \widehat{p}) = \overline{\phi}_{21}(x_1, \widehat{p}) + \overline{\phi}_{22}(\overline{x}_2, \widehat{p}) > 0$, and a direct substitution leads to

$$\dot{V}_2 \le -\left(1-\frac{2}{4}\right)\eta^2 - (\nu_1 - 5)z_1^2 + z_2\left(d_2(t)u + z_2\widehat{\phi}_2(\overline{x}_2,\widehat{p})\widehat{p} + z_2\overline{\phi}_2(\overline{x}_2,\widehat{p})\right) + \left(\frac{1}{\Upsilon}\widetilde{p} + z_2\frac{\partial\vartheta_1}{\partial\widehat{p}}\right)\left(\tau_2 - \dot{\widehat{p}}\right). \quad (94)$$

As in Step 1, we design the following partial-state feedback controller:

$$u = -\frac{1}{\underline{d}_2}\left(\nu_2 + \widehat{\phi}_2(\overline{x}_2,\widehat{p})\widehat{p} + \overline{\phi}_2(\overline{x}_2,\widehat{p})\right)z_2 \quad (95)$$

with the updating law for unknown parameter p^*

$$\dot{\widehat{p}} = \Upsilon\widehat{\phi}_1(x_1)z_1^2 + \Upsilon\widehat{\phi}_2(x_1,x_2,\widehat{p})z_2^2, \quad (96)$$

which is such that

$$\dot{V}_2 \le -\frac{1}{2}\eta^2 - (\nu_1 - 5)z_1^2 - \nu_2 z_2^2. \quad (97)$$

The Lyapunov function V_2 can be made $\dot{V}_2 \le 0$ by choosing $\nu_1 > 5$, $\nu_2 > 0$, and the stability analysis can be done in the similar way to Theorem 11. The simulation plots shown in Figures 1 and 2 are performed by MATLAB with the following parameters: $d_1(t) = 2 - \sin(t)$, $d_2(t) = 2 + \sin(t)$, $\underline{d}_1 = 1$, $\overline{d}_1 = 3$, $\underline{d}_2 = 1$, $\overline{d}_2 = 3$, $\nu_1 = 6$, $\nu_2 = 1$, $\Upsilon = 1$, the derived functions: $\phi_{11}(|\eta|) = 0$, $\phi_{21}(|\eta|) = |\eta|$, $\phi_{12}(x_1) = |x_1| + (1/2)|x_1|$, $\phi_{22}(x_1,x_2) = 0$, $\widehat{\phi}_1(x_1) = 1 + 1/2 + x_1^2$, $\widehat{\phi}_2(x_1,x_2,\widehat{p}) = 2 + (3/2)(\partial\vartheta_1/\partial x_1)^2(\nu_1 + \widehat{p})^2$, and the initial conditions: $\eta(0) = 0.5$, $x_1(0) = 1$, $z_2(0) = 0.1$, $\widehat{p}(0) = 1$.

According to our results reported in Theorem 11, the states (η, x_1, x_2) must asymptotically converge to the origin and the parameter estimate $\widehat{p}$ is bounded on $[0,\infty)$. This fact can be verified from Figure 1, which plots the trajectories of these dynamic signals $(\eta(t), x_1(t), x_2(t), \widehat{p}(t))$. It can be seen that, at about $t = 4.17$ s, $t = 0.68$ s, and $t = 0.59$ s, the states approach the origin, and at $t = 0.36$ s, the parameter estimate $\widehat{p}(t)$ is bounded near $p^* = 1.1$. In addition, according to Theorem 11, the control input u is convergent to the origin. Figure 2 demonstrates this result, and it can be shown that, at about $t = 0.51$ s, the input signal u approaches the origin. As can be seen from Figures 1 and 2, our control scheme provides a fairly good asymptotic stabilization performance.

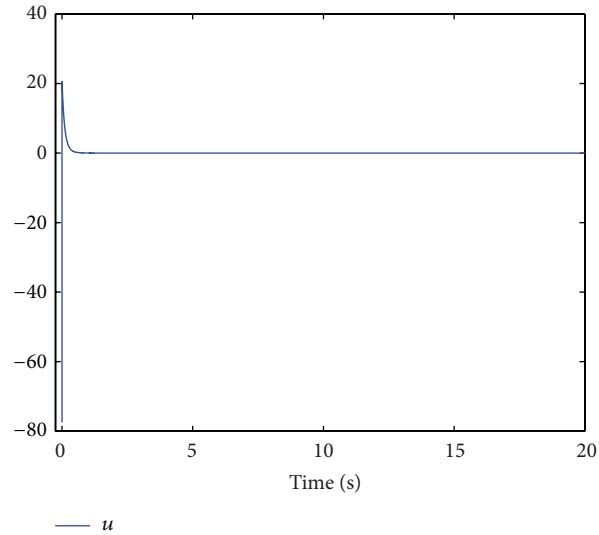


Figure 2: The control input of the closed-loop system.

6. Conclusion

The state feedback stabilization problem is investigated for a class of nonlinear systems with dynamic uncertainties and uncertain control coefficients in this paper. The dynamic uncertainty is characterized by the uncertain ISS supply rates. A global asymptotic stabilization control scheme is proposed using the backstepping design scheme. The tuning function technique is applied in this procedure, which avoids the disadvantage of overparameterization. It is shown that, under some more restrictive conditions, a linear state feedback controller can be designed by the presented algorithm. The simulation example demonstrates the effectiveness of the proposed method.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is supported in part by the National Natural Science Foundation of China under Grant no. 61304008, the Shandong Provincial Natural Science Foundation of China under Grant no. ZR2013FQ033, and the Doctor Research Foundation of Shandong Jianzhu University under Grant no. XNBS1272.

References

[1] I. R. Petersen and R. Tempo, "Robust control of uncertain systems: classical results and recent developments," *Automatica*, vol. 50, no. 5, pp. 1315–1335, 2014.

[2] E. D. Sontag, "Smooth stabilization implies coprime factorization," *IEEE Transactions on Automatic Control*, vol. 34, no. 4, pp. 435–443, 1989.

[3] E. D. Sontag, "Comments on integral variants of ISS," *Systems & Control Letters*, vol. 34, no. 1-2, pp. 93–100, 1998.

[4] P. Kokotovic and M. Arcak, "Constructive nonlinear control: a historical perspective," *Automatica*, vol. 37, no. 5, pp. 637–662, 2001.

[5] Z.-P. Jiang, A. R. Teel, and L. Praly, "Small-gain theorem for ISS systems and applications," *Mathematics of Control, Signals, and Systems*, vol. 7, no. 2, pp. 95–120, 1994.

[6] J. Tian, W. Feng, and Y. Wang, "High-order stochastic adaptive controller design with application to mechanical system," *Mathematical Problems in Engineering*, vol. 2012, Article ID 718913, 17 pages, 2012.

[7] N. Duan and H.-K. Liu, "Adaptive output feedback control for a class of stochastic nonlinear systems with SiISS inverse dynamics," *Mathematical Problems in Engineering*, vol. 2012, Article ID 673878, 15 pages, 2012.

[8] Z.-P. Jiang, I. Mareels, D. J. Hill, and J. Huang, "A unifying framework for global regulation via nonlinear output feedback: from ISS to iISS," *IEEE Transactions on Automatic Control*, vol. 49, no. 4, pp. 549–562, 2004.

[9] Y. G. Liu, "Global stabilization by output feedback for a class of nonlinear systems with uncertain control coefficients and unmeasured states dependent growth," *Science in China Series F: Information Sciences*, vol. 51, no. 10, pp. 1508–1520, 2008.

[10] X. Yan and Y. Liu, "Global practical tracking for high-order uncertain nonlinear systems with unknown control directions," *SIAM Journal on Control and Optimization*, vol. 48, no. 7, pp. 4453–4473, 2010.

[11] Y.-Q. Wu, J.-B. Yu, and Y. Zhao, "Further results on global asymptotic regulation control for a class of nonlinear systems with iISS inverse dynamics," *IEEE Transactions on Automatic Control*, vol. 56, no. 4, pp. 941–946, 2011.

[12] D. Xu and J. Huang, "Output regulation for output feedback systems with iISS inverse dynamics," *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, vol. 133, no. 4, Article ID 044503, 2011.

[13] X. Yu, Y.-Q. Wu, and X.-J. Xie, "Reduced-order observer-based output feedback regulation for a class of nonlinear systems with iISS inverse dynamics," *International Journal of Control*, vol. 85, no. 12, pp. 1942–1951, 2012.

[14] X. Yu and G. Liu, "Output feedback control of nonlinear systems with uncertain ISS/iISS supply rates and noises," *Nonlinear Analysis. Modelling and Control*, vol. 19, no. 2, pp. 286–299, 2014.

[15] J. Yu and Y. Wu, "Global set-point tracking control for a class of non-linear systems and its application in continuously stirred tank reactor systems," *IET Control Theory & Applications*, vol. 6, no. 12, pp. 1965–1971, 2012.

[16] Y.-G. Liu, "Output-feedback adaptive control for a class of nonlinear systems with unknown control directions," *Acta Automatica Sinica*, vol. 33, no. 12, pp. 1306–1312, 2007.

[17] F. Hong, S. S. Ge, B. Ren, and T. H. Lee, "Robust adaptive control for a class of uncertain strict-feedback nonlinear systems," *International Journal of Robust and Nonlinear Control*, vol. 19, no. 7, pp. 746–767, 2009.

[18] M. Krstić, I. Kanellakopoulos, and P. V. Kokotović, *Nonlinear and Adaptive Control Design*, Wiley-Interscience, New York, NY, USA, 1995.

[19] E. Sontag and A. Teel, "Changing supply functions in input/state stable systems," *IEEE Transactions on Automatic Control*, vol. 40, no. 8, pp. 1476–1478, 1995.

[20] W. Qiang-de, W. Chun-ling, and W. Yu-qiang, "Adaptive control of stochastic nonlinear systems with uncontrollable linearization," *International Journal of Adaptive Control and Signal Processing*, vol. 23, no. 7, pp. 667–678, 2009.

[21] H. K. Khalil, *Nonlinear Systems*, Prentice Hall, Upper Saddle River, NJ, USA, 2002.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 724795, 11 pages
http://dx.doi.org/10.1155/2015/724795

Research Article

Observer Based Robust Position Control of a Hydraulic Servo System Using Variable Structure Control

E. Kolsi-Gdoura, M. Feki, and N. Derbel

Research Group CEMLab, National Engineering School of Sfax, University of Sfax, 1073 Sfax, Tunisia

Correspondence should be addressed to M. Feki; moez.feki@enig.rnu.tn

Received 8 May 2015; Revised 9 July 2015; Accepted 12 July 2015

Academic Editor: Rongwei Guo

This paper deals with the position control of a hydraulic servo system rod. Our approach considers the surface design as a case of virtual controller design using the backstepping method. We first prove that a linear surface does not yield to a robust controller with respect to the unmatched uncertainty and perturbation. Next, to remedy this deficiency, a sliding controller based on the second-order sliding mode is proposed which outperforms the first controller in terms of chattering attenuation and robustness with respect to parameter uncertainty only. Next, based on backstepping a nested variable structure design method is proposed which ensures the robustness with respect to both unmatched uncertainty and perturbation. Finally, a robust sliding mode observer is appended to the closed loop control system to achieve output feedback control. The stability and convergence to reference position with zero steady state error are proven when the controller is constructed using the estimated states. To illustrate the efficiency of the proposed methods, numerical simulation results are shown.

1. Introduction

Actually, the hydraulic servo systems are very popular in several industrial applications such as robotics, aerospace flight-control actuators, heavy machinery, aircrafts, automotive industry, and a variety of automated manufacturing systems. This is mainly due to their ability to produce high power and accurate and fast responses. However, these systems have a high nonlinear behavior due to the pressure flow characteristics [1] and the leakage model inside the servovalves [2]. This fact makes the control design for precise output tracking a very challenging task.

Owing to their simplicity, linear controllers of PID type [3, 4], input/output linearization controllers [5–9], and also sliding mode controllers (SMC) [10–13] have been used to control the hydraulic servo systems. However, such controllers were designed based on the plant physical model and, therefore, the plant parameters knowledge is required. Consequently, they were shown to be highly sensitive to mismatched perturbation and uncertainties, thus resulting in performance degradation.

To improve the controller performances, several strategies have been adopted such as using the self-tuned PID controller [14, 15] and nonlinear adaptive controllers [16–18]. SMC appended with some improvements have also been used. In [19, 20] SMC method has been combined with an adaptive controller, which can compensate for the system uncertain nonlinearities, for linear uncertain parameters, and especially for the nonlinear uncertain parameters to construct an asymptotically stable tracking. In [21], SMC has been used with the PID controller to achieve control of asymmetrical hydraulic cylinder trajectory tracking. To drive electrohydraulic actuators, various robust control techniques, such as $\mathscr{H}_2$ and $\mathscr{H}_\infty$ controls, were applied [22–24]. This approach enabled the compensation for the inherent nonlinearities of the actuator and rejects matched external disturbances and attenuates mismatched external disturbances. To cope with mismatched disturbances authors used the integral SMC and to remedy the slow response due to windup phenomenon a realizable reference compensation has been used to achieve fast position tracking [25, 26]. Since it has been proposed by Levant [27, 28], higher-order SMC (HOSMC) has been widely used to control electrical drives [29, 30], electropneumatic actuators [31], and electrohydraulic actuators [32, 33].

In the present paper, we are interested in controlling the position of the rod in a hydraulic servo system that consists of a four-way spool valve supplying a double effect linear cylinder with a double-rodded piston. The piston is driving a load modeled by a mass, a spring, and a sliding viscous friction. Our work aims to design a controller that may achieve the reference position in presence of mismatched parameter uncertainty and perturbation in addition to actuator saturation. To realize this objective, we start in the second section by formulating the problem and presenting the effects of using first- and second-order SMC with a linear surface. In Section 3, we present the design of a sliding surface obtained using backstepping method and variable structure controller, leading hence to a nonlinear surface that allows achieving the reference output despite the presence of uncertainties and perturbations. Numerical simulation results are presented to illustrate the efficiency of the proposed control design. In Section 4, we present a sliding observer and prove the convergence of the observer as well as the exact position tracking using the nonlinear surface SMC issued from the observer states. Finally, the conclusion and some remarks are presented in Section 5.

2. Problem Statement

The electrohydraulic system that we will deal with in this paper is depicted in Figure 1 and modeled by the dynamical system (1). It has been shown in [34] that, for the symmetrical piston with equal surfaces S_1 and S_2 and assuming equal volume flow passing through (geometrically) identical ports, we can describe the system by three variable states where a differential pressure state substitutes the pressure of each chamber. This decrease in the system dimension ensures the observability when the system output is the piston position. Consider

$$\dot{x}_1 = \frac{4B}{V_t}\left(ku\sqrt{P_d - \operatorname{sign}(u)\, x_1} - \frac{\alpha x_1}{1+\gamma|u|} - Sx_2\right), \quad (1)$$

$$\dot{x}_2 = \frac{1}{m_t}\left(Sx_1 - bx_2 - (k_l + \Delta k_l)\, x_3\right), \quad (2)$$

$$\dot{x}_3 = x_2 + d(t), \quad (3)$$

where $x_1 = P_1 - P_2$ denotes the difference in pressure inside the two chambers of the cylinder, x_2 and x_3, respectively, denote the velocity and the position of the rod, and $|d(t)| < d_{\max}$ is a bounded constant or slowly varying external perturbation. In fact, from Newton's law, a constant force perturbation leads to a constant acceleration. Thus, the velocity perturbation may be interpreted as the result of an impulsive force that acts abruptly on the system. $m_t = m + m_0$ is the total mass of the rod and the load, $V_t = V_1 + V_2$ is the total volume of the cylinder, and $P_d = P_s - P_r$ is the pressure difference between the supply pressure P_s (pressure of the pump) and the return pressure P_r (atmospheric pressure). k_l is the spring stiffness constant with an uncertainty Δk_l and b is the friction coefficient. The system parameters used for simulations are as follows: $B = 2.2\times10^9$ Pa, $P_s = 300\times10^5$ Pa, $P_r = 10^5$ Pa, $m_0 = 50$ kg, $S = 1.5\times10^{-3}$ m^2, $V_t = 9\times10^{-4}$ m^3, $m = 20$ kg, $b = 590$ kg/s, $k_l = 125000$ N/m, $k = 3.62\times10^{-5}$ m^3 s^{-1} A^{-1} Pa$^{-1/2}$, $\alpha = 4.1816\times10^{-12}$ m^3 s^{-1} Pa^{-1}, and $\gamma = 8571$ with α and γ being intrinsic constants modeling the leakage within the servovalve [2].

The most difficult aspect in this model is the existence of a mismatched perturbation as well as a mismatched parameter uncertainty. In addition, the leakage model includes nonlinearity with respect to the control signal u. To deal with this problem, we neglect the leakage term in the design but we consider it in simulation and consider the system as a switching system between two models. The switching aspect makes the use of the sliding mode approach a good candidate to design a controller u that can drive the rod position to a constant reference position $x_{3\text{ref}}$.

2.1. First-Order SMC. The SMC design consists of two phases. In the first phase the sliding surface is designed such that the system is asymptotically stable when it is confined to it and in the second phase a switching controller is designed to ensure the existence of the sliding mode. Our idea consists in viewing the sliding surface design as a special case of backstepping design. Therefore, at sliding mode, $\sigma(x) = 0$ means $x_1 = p(x)$ and x_1 can be viewed as a virtual controller to subsystem ((2)-(3)) which describes the system behavior on the sliding surface.

Thus, should we choose a linear virtual controller,

$$p(x) = \frac{1}{S}\left(k_l x_{3\text{ref}} - C_2 x_2 - C_3\left(x_3 - x_{3\text{ref}}\right)\right), \quad (4)$$

we get the sliding surface

$$\sigma(x) = Sx_1 + C_2 x_2 + C_3\left(x_3 - x_{3\text{ref}}\right) - k_l x_{3\text{ref}}. \quad (5)$$

In sliding mode and if uncertainty and perturbations are neglected, the system is a second dimensional linear system with the characteristic equation

$$s^2 + \frac{C_2 + b}{m_t}s + \frac{C_3 + k_l}{m_t} = 0. \quad (6)$$

Using the pole placement method and imposing a stable multiple pole at $s = -\lambda$, we can determine the control parameters C_2 and C_3:

$$\begin{aligned} C_2 &= 2\lambda m_t - b, \\ C_3 &= \lambda^2 m_t - k_l. \end{aligned} \quad (7)$$

Eventually, we obtain an exponentially stable system if we can guarantee the attractivity of the sliding surface $\sigma(x) = 0$. Indeed, the attractivity condition $\sigma(x)\dot{\sigma}(x) < 0$ will be satisfied if we choose $\dot{\sigma}(x) = -W\operatorname{sign}(\sigma(x))$, where $W > 0$ is the sliding gain that should be chosen large enough in order to ensure the attractivity of the surface in presence of perturbation and uncertainty. Using the system model and (5), we can show that $\sigma(x)\dot{\sigma}(x) < 0$ is attained if the sliding gain satisfies

$$W > \left|C_3 d_{\max}\right| + \frac{4BS}{V_t}\alpha P_d + \left|\frac{V_t C_2}{S m_t}\Delta k_l\right|, \quad (8)$$

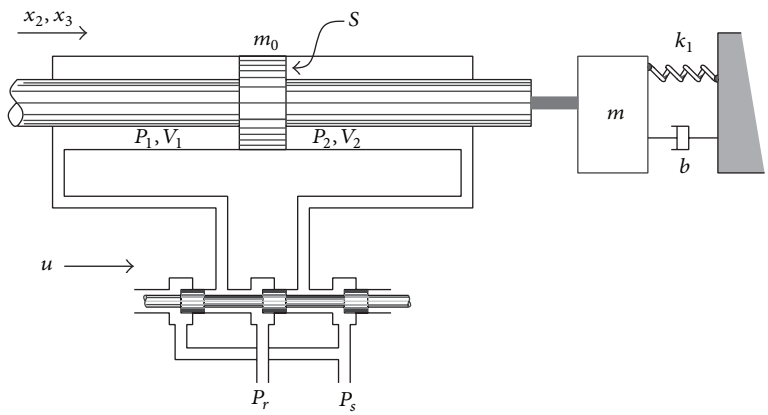


Figure 1: Hydraulic servo system controlled using a servovalve.

where $d_{\max} = \max_{t\geq 0} d(t)$. Consequently, by solving $\dot{\sigma}(x) = -W\operatorname{sign}(\sigma(x))$, the control law is expressed as follows:

$$u(x) = \begin{cases} \dfrac{N(x)}{(4BSk/V_t)\sqrt{P_d - x_1}}, & \text{if } N(x) \leq 0 \\ \dfrac{N(x)}{(4BSk/V_t)\sqrt{P_d + x_1}}, & \text{if } N(x) < 0, \end{cases} \tag{9}$$

where

$$\begin{aligned} N(x) = &-W\operatorname{sign}(\sigma(x)) - \frac{C_2}{m_t}(Sx_1 - bx_2 - k_l x_3) \\ &- C_3 x_2 + \frac{4BS^2}{V_t} x_2. \end{aligned} \tag{10}$$

Despite the perturbation $d(t)$ and the uncertainty Δk_l, the variable structure controller defined by (9) with the sliding gain W in (8) ensures the existence of sliding mode; however, we can easily deduce that we do not achieve the reference output because when the system behavior is confined to the sliding surface (5) the linear virtual controller does not guarantee any robustness with respect to the perturbation and the uncertainty. Indeed, it can be easily shown that as far as the sliding motion is preserved the system is asymptotically stable if the closed loop eigenvalue is chosen such as

$$\lambda > \sqrt{\left|\frac{\Delta k_l}{m_t}\right|} \tag{11}$$

and in that case the steady state error due to the uncertainty is given by

$$e_{k_l} = \frac{\Delta k_l}{m_t} \frac{1}{\lambda^2 + \Delta k_l / m_t} x_{3\text{ref}}. \tag{12}$$

We similarly can show that the steady state error due to the constant perturbation is given by

$$e_d = -\frac{2}{\lambda} d. \tag{13}$$

Figure 2 shows the system behavior with $\lambda = 50$, $d(t) = 0.1$, $\Delta k_l = 25000$, and $x_{3\text{ref}} = 20$ cm. We clearly notice that there is a steady state error of more than 2 cm at the position output. So the first-order sliding mode is not robust to the mismatched uncertainty and perturbation.

2.2. Second-Order SMC. Since the system is third-order single input, then we may think of designing the higher-order SMC up to second order. Let σ_h denote the sliding variable defined as

$$\sigma_h = h(x_3 - x_{3\text{ref}}) + x_2, \tag{14}$$

where $h > 0$ is a strictly positive scalar. We may verify that the relative degree of system ((1)–(3)) together with (14) with respect to the sliding variable σ_h is constant and equal to two. Thus we have

$$\ddot{\sigma}_h = a_1(x,t) + \Delta a_1(x,t) + a_2(x)u, \tag{15}$$

where $a_1(x,t)$ and $a_2(x)$ are known functions and $\Delta a_1(x,t)$ is an unknown bounded function in terms of $d(t)$ and Δk_l:

$$\begin{aligned} a_1(x,t) &= -\frac{4BS}{m_t V_t}(\alpha x_1 + Sx_2) \\ &\quad + \left(\frac{h}{m_t} - \frac{b}{m_t^2}\right)(Sx_1 - bx_2 - k_l x_3) \\ &\quad - \frac{k_l}{m_t} x_2, \\ \Delta a_1(x,t) &= -\left(\frac{h}{m_t} - \frac{b}{m_t^2}\right)\Delta k_l x_3 - \frac{\Delta k_l}{m_t} x_2 \\ &\quad - d(t)\frac{k_l + \Delta k_l}{m_t}, \\ a_2(x) &= \begin{cases} \dfrac{4BS}{m_t V_t} k\sqrt{P_d - x_1} & \text{if } u \geq 0 \\ \dfrac{4BS}{m_t V_t} k\sqrt{P_d + x_1} & \text{if } u < 0. \end{cases} \end{aligned} \tag{16}$$

The solution should be understood in the Filippov sense [35], and the trajectories of the system are supposed to be extendible infinitely in time for any bounded measurable input.

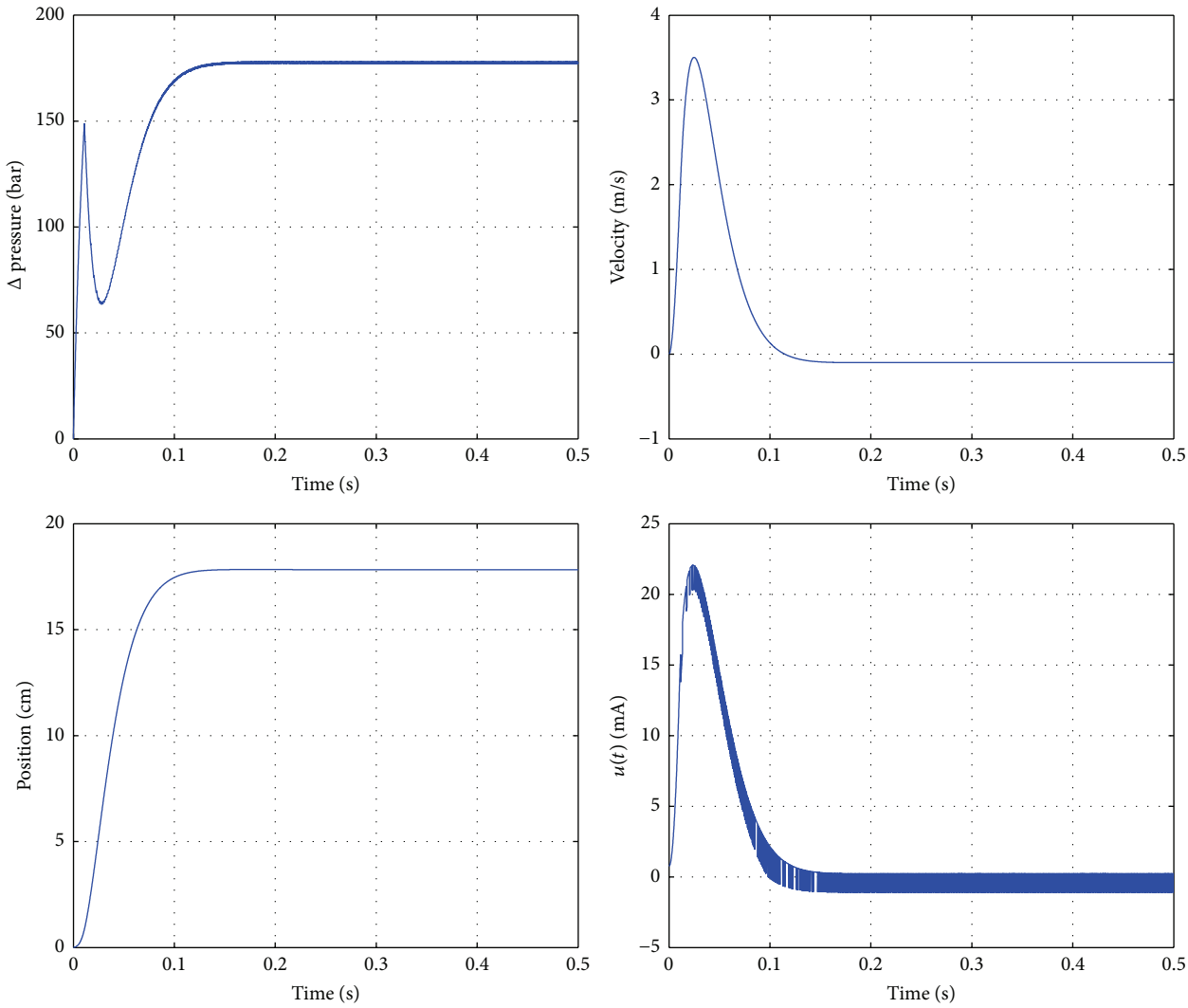


FIGURE 2: The behavior of the system under sliding mode controller defined by (9); $\lambda = 50$.

By defining $z_1 = \sigma_h$ and $z_2 = \dot{\sigma}_h$, achieving sliding mode $\sigma_h = 0$ is equivalent to the finite stabilization of the system:

$$\begin{aligned} \dot{z}_1 &= z_2, \\ \dot{z}_2 &= a_1(x,t) + \Delta a_1(x,t) + a_2(x)\,u, \end{aligned} \tag{17}$$

which hence ideally yields to the sliding set $\sigma_h = 0$ and $\dot{\sigma}_h = 0$. Taking into account the practical considerations, the sliding set is defined as follows [27].

Definition 1. Given the sliding variable $\sigma_h(x,t)$, the "real second-order sliding set" associated with ((1)–(3)) is defined as

$$s_{T_e} = \left\{ x \in \frac{\chi}{|\sigma_h|} \le k_1 T_e^2, |\dot{\sigma}_h| \le k_2 T_e \right\}, \tag{18}$$

where T_e is the finite sampling time (fixed at $T_e = 20\,\mu$s in the sequel) and k_1 and k_2 are positive constants.

Definition 2. Consider the nonempty real second-order sliding set s_{T_e} given in (18), and assume that it is locally an integral set in the Filippov sense. The corresponding behavior of system ((1)–(3)) satisfying (18) is called "real second-order sliding mode" with respect to $\sigma_h(x,t)$ [27].

The variable structure control law u can be chosen as follows:

$$\begin{aligned} u = \frac{1}{a_2(x)} &\left(-a_1(x,t) - K_1 z_1 - K_2 z_2 - W_1 \operatorname{sign}(z_1)\right. \\ &\left. - W_2 \operatorname{sign}(z_2)\right), \end{aligned} \tag{19}$$

where all the controller gains K_1, K_2, W_1, and W_2 are strictly positive. Applying controller (19) yields to

$$\begin{aligned} \begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ -K_1 & -K_2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \\ &+ \begin{bmatrix} 0 & 0 \\ -W_1 & -W_2 \end{bmatrix} \begin{bmatrix} \operatorname{sign}(z_1) \\ \operatorname{sign}(z_2) \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Delta a_1(x,t) \end{aligned} \tag{20}$$

which can be written in the following compact form:

$$\dot{Z} = A_K Z + A_W \operatorname{sign}(z) - B\Delta a(x,t), \tag{21}$$

where $A_K = \begin{bmatrix} 0 & 1 \\ -K_1 & -K_2 \end{bmatrix}$, $A_W = \begin{bmatrix} 0 & 0 \\ -W_1 & -W_2 \end{bmatrix}$, and $B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

We can determine the control parameters K_1 and K_2 such that the characteristic equation of A_K has two equal eigenvalues at $s = -\lambda$:

$$\begin{aligned} K_1 &= \lambda^2, \\ K_2 &= 2\lambda, \\ \lambda &> 0. \end{aligned} \tag{22}$$

Since A_K is a Hurwitz matrix, then it satisfies the Lyapunov function $A_K^T P + PA_K = -Q$, where P and Q are some positive definite matrices. To determine the gains W_1 and W_2 we define the following Lyapunov function candidate for system (20):

$$V(Z) = Z^T P Z. \tag{23}$$

Its time derivative in the direction of system (20) trajectories is given by

$$\begin{aligned} \dot{V} &= \dot{Z}^T P Z + Z^T P \dot{Z} \\ &= -Z^T Q Z + 2Z^T P A_W \operatorname{sign}(Z) - 2Z^T P B \Delta a(x,t). \end{aligned} \tag{24}$$

If we choose $P_{22}W_1 = P_{12}W_2$ then PA_W is a negative definite matrix and we have

$$\begin{aligned} \dot{V} < &-Z^T Q Z + 2\lambda_{\max}(PA_W)\|Z\|_1 \\ &+ 2\|Z\|_1 \lambda_{\min}(P)\|B\|_1 |\Delta a(x,t)| < 0 \end{aligned} \tag{25}$$

provided that W_1 and W_2 are also chosen such that

$$-\lambda_{\max}(PA_W) > \lambda_{\min}(P)\left|\Delta a_1(x,t)\right|, \tag{26}$$

where $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ are, respectively, the largest and least eigenvalues of the matrix.

From Figure 3, we can notice that the second-order SMC achieves a better performance than the first-order SMC. However, the robustness with respect to the constant perturbation is not guaranteed. Indeed, in sliding mode, we have $\sigma_h = 0$; thus $x_2 = -h(x_3 - x_{3\text{ref}}) = -he_3$ and hence on the sliding set $\sigma_h = 0$ we get

$$\dot{e}_3 = -he_3 + d(t) \tag{27}$$

which is a first-order nonautonomous system with a steady state value equal to $d(t)/h$. Therefore, for the constant perturbation $d(t) = 0.1$, the steady state error is as expected and delineated in Figure 3; $x_3 - x_{3\text{ref}} = 0.1$ cm. When increasing the value of h, the variable structure gains W_1 and W_2 should be increased, thus accentuating the chattering phenomenon which was attenuated by the use of the second-order sliding mode.

The boundedness of x_1 and x_2 is ensured since $\dot{V} < 0$ and $V(Z)$ depends on z_1 and z_2 which in turn depend on x_1, x_2, and x_3.

3. Sliding Mode Controller with Nonlinear Surface

To overcome the problem of mismatched perturbation and uncertainty, we suggest in this section designing a sliding surface $\sigma(x)$ based on the backstepping method and using robust variable structure virtual controller.

In Section 2.1, the sliding surface $\sigma(x) = 0$ was obtained from the design of the linear controller $x_1 = p(x)$ for subsystem ((2)-(3)). To ensure robustness with respect to the parametric uncertainty, a variable structure virtual controller $x_1 = p(x)$ is designed by choosing a sliding surface $\sigma_1(x_2, x_3)$ and requiring its attractivity by imposing

$$\dot{\sigma}_1(x_2, x_3) = -W_1 \operatorname{sign}(\sigma_1(x_2, x_3)). \tag{28}$$

Therefore, the virtual controller $x_1 = p(x)$ is obtained by solving (28) and the sliding surface $\sigma(x) = x_1 - p(x)$. However, when sliding on $\sigma_1(x_2, x_3)$ is achieved, we get $\sigma_1(x_2, x_3) = 0$ which means $x_2 = \nu(x_3)$ and thus in sliding mode we have

$$\dot{x}_3 = \nu(x_3) + d(t). \tag{29}$$

Again to ensure asymptotic convergence of $x_3(t)$ to $x_{3\text{ref}}$ a simple linear proportional term $\nu(x_3) = -C(x_3 - x_{3\text{ref}})$ is not satisfactory, but a variable structure term can guarantee the required convergence. Indeed, with

$$\nu(x_3) = -W_3(x_3 - x_{3\text{ref}}) - W_2 \operatorname{sign}(x_3 - x_{3\text{ref}}), \tag{30}$$

where the gains are chosen such that $W_2 > d_{\max}$ and $W_3 > 0$, we may easily show the convergence of $x_3(t)$ to $x_{3\text{ref}}$ as far as the sliding surface

$$\begin{aligned} \sigma_1(x_2, x_3) = x_2 &+ W_3(x_3 - x_{3\text{ref}}) \\ &+ W_2 \operatorname{sign}(x_3 - x_{3\text{ref}}) \end{aligned} \tag{31}$$

is attractive.

When attempting to achieve the attractivity of $\sigma_1(x_2, x_3)$ by imposing (28), we face the attempt to differentiate a discontinuous function. Although this is mathematically impossible, we can overcome the problem from an engineering point of view by either replacing the discontinuous signum function with a smoother saturation function or directly considering that the derivative of the signum function is the Dirac impulse ($\delta(x)$) which is zero everywhere except at an isolated single point. Indeed, in engineering realization using a processor and discretization process, the isolated discontinuities, especially when they are few, can easily be avoided and would not cause any problems. Therefore, from (28), we have

$$\begin{aligned} &\frac{1}{m_t}(Sx_1 - bx_2 - k_l x_3) + W_3 x_2 + W_2 x_2 \delta(x_3 - x_{3\text{ref}}) \\ &\quad = -W_1 \operatorname{sign}(\sigma_1(x_2, x_3)) \end{aligned} \tag{32}$$

and the sliding surface $\sigma(x)$ becomes

$$\begin{aligned} \sigma(x) = Sx_1 &+ (m_t W_3 + m_t W_2 \delta(x_3 - x_{3\text{ref}}) - b)x_2 \\ &- k_l x_3 + m_t W_1 \operatorname{sign}(\sigma_1(x_2, x_3)). \end{aligned} \tag{33}$$

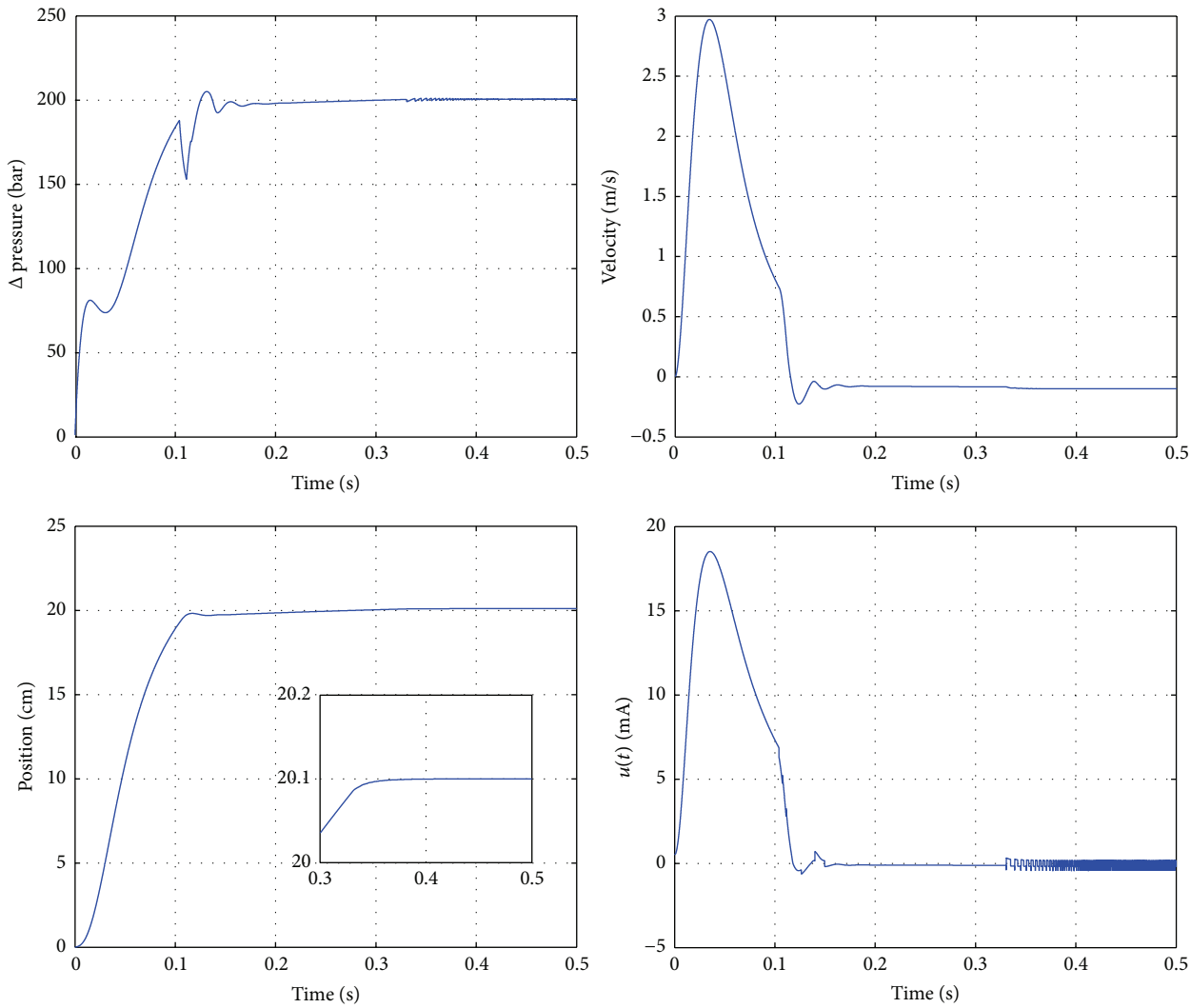


FIGURE 3: Behavior of system under second-order sliding mode controller (19); $\lambda = 50$ and $h = 100$.

Finally, by imposing $\dot{\sigma}(x) = -W\text{sign}(\sigma(x))$, the attractivity condition $\sigma(x)\dot{\sigma}(x) < 0$ is satisfied. The choice of the surface $\sigma(x)$ and its imposed derivative lead to the following controller:

$$u(x) = \begin{cases} \dfrac{N_d(x)}{(4BSk/V_t)\sqrt{P_d - x_1}}, & \text{if } N_d(x) \geq 0 \\ \dfrac{N_d(x)}{(4BSk/V_t)\sqrt{P_d + x_1}}, & \text{if } N_d(x) < 0 \end{cases} \tag{34}$$

with

$$\begin{aligned} N_d(x) = &-W\,\text{sign}(\sigma(x)) - m_t W_1 \delta(\sigma_1(x)) \\ &+ \left(\frac{4BS\alpha}{Vt} + \frac{bS}{m_t} - SW_4\right) x_1 \\ &+ \left(\frac{4BS^2}{Vt} - \frac{b^2}{m_t} + k_l + bW_4\right) x_2 \\ &+ \left(k_l W_4 - \frac{bk_l}{m_t}\right) x_3, \end{aligned} \tag{35}$$

where the derivative of the Dirac impulse is considered as zero and the sliding gain W should be chosen such that the attractivity occurs in presence of uncertainty and perturbation:

$$W > \left|\frac{bV_t}{Sm_t}\Delta k_l\right| + \left|\frac{V_t}{S}W_3\Delta k_l\right| + \left|k_l d_{\max}\right|. \tag{36}$$

From Figure 4, we can notice clearly that we have attained our aim to drive the hydraulic servo system to the reference position. Nevertheless, due to the nested sliding modes, the chattering phenomenon was emphasized.

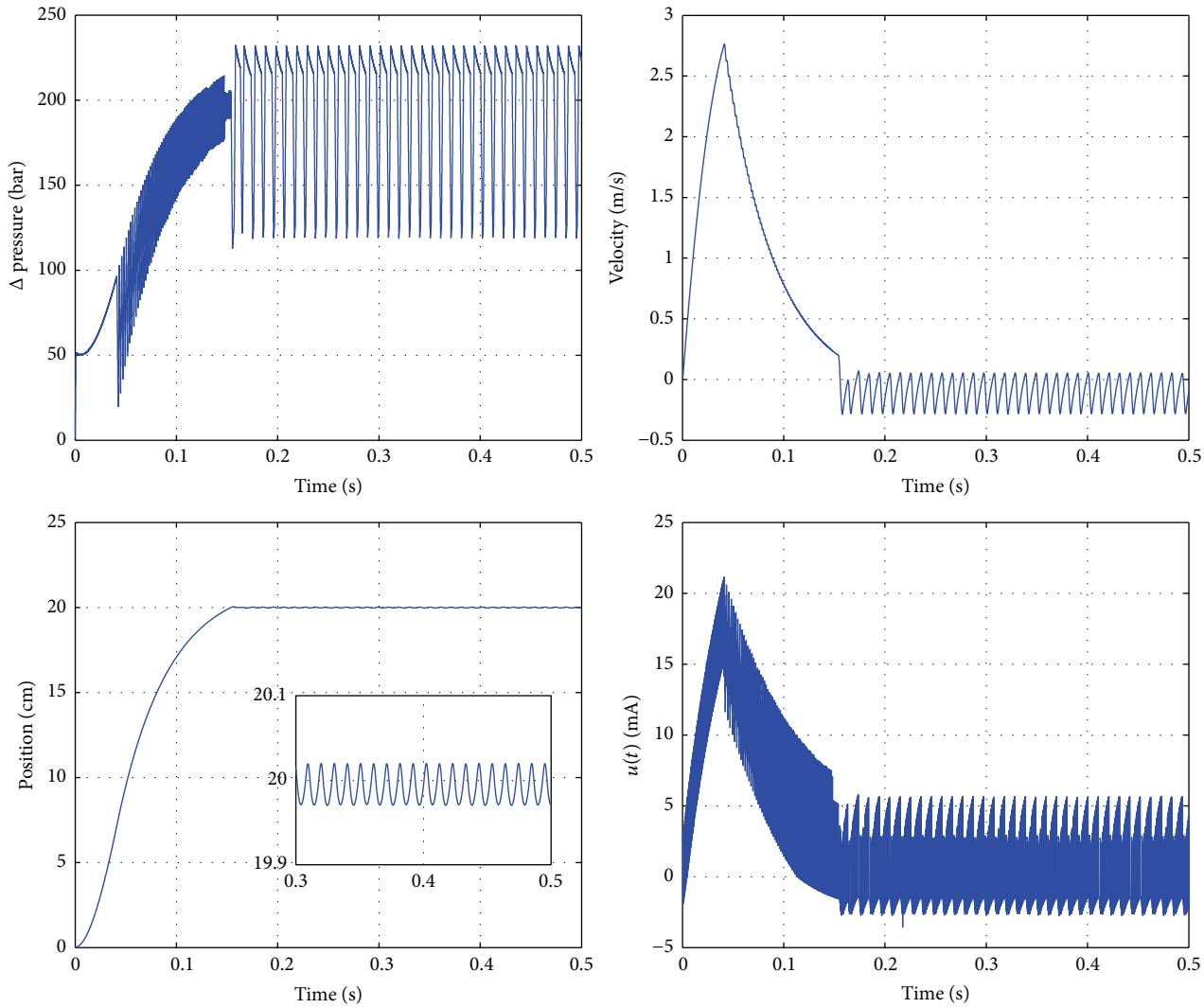


Figure 4: Behavior of system under sliding mode controller using nonlinear surface defined by (33).

In order to attenuate it and to avoid the mathematical problem of taking the derivative of the signum function we have used a smooth saturation function $\operatorname{sign}(x) \simeq \tanh(\mu x)$ for a sufficiently high positive value $\mu > 0$. The results depicted in Figure 5 show that the use of the smooth function helped to achieve the reference value in presence of mismatched uncertainty and perturbation with attenuated chattering.

4. Sliding Mode Observer Design

As we can notice, the controller conceived in the foregoing section uses all three state variables. However, measuring the differential pressure $\Delta P = x_1$ is a costly task and requires high technology procedure to avoid additional leakage. To remedy this situation, we propose in this section designing a sliding mode observer that may estimate the required states that are then used to construct the controller.

Let us consider the observer model given in (37) and inferred from the step-by-step observer presented in [36, 37]:

$$\dot{z}_1 = \frac{4B}{V_t}\left(ku\sqrt{P_d - \operatorname{sign}(u)\,\tilde{z}_1} - \frac{\alpha\tilde{z}_1}{1+\gamma|u|} - S\tilde{z}_2\right) + L_1 \operatorname{sign}(\tilde{z}_1 - z_1), \tag{37}$$

$$\dot{z}_2 = \frac{1}{m_t}\left(Sz_1 - b\tilde{z}_2 - k_l z_3\right) + L_2 \operatorname{sign}(\tilde{z}_2 - z_2), \tag{38}$$

$$\dot{z}_3 = z_2 + L_3 \operatorname{sign}(x_3 - z_3), \tag{39}$$

where L_1, L_2, and L_3 are the observer gain and $\tilde{z}_1$ and $\tilde{z}_2$ are defined as

$$\begin{aligned} \tilde{z}_1 &= z_1 + \frac{k_l}{S}(x_3 - z_3) + \frac{m_t}{S}L_2 \operatorname{sign}(\tilde{z}_2 - z_2), \\ \tilde{z}_2 &= z_2 + L_3 \operatorname{sign}(x_3 - z_3). \end{aligned} \tag{40}$$

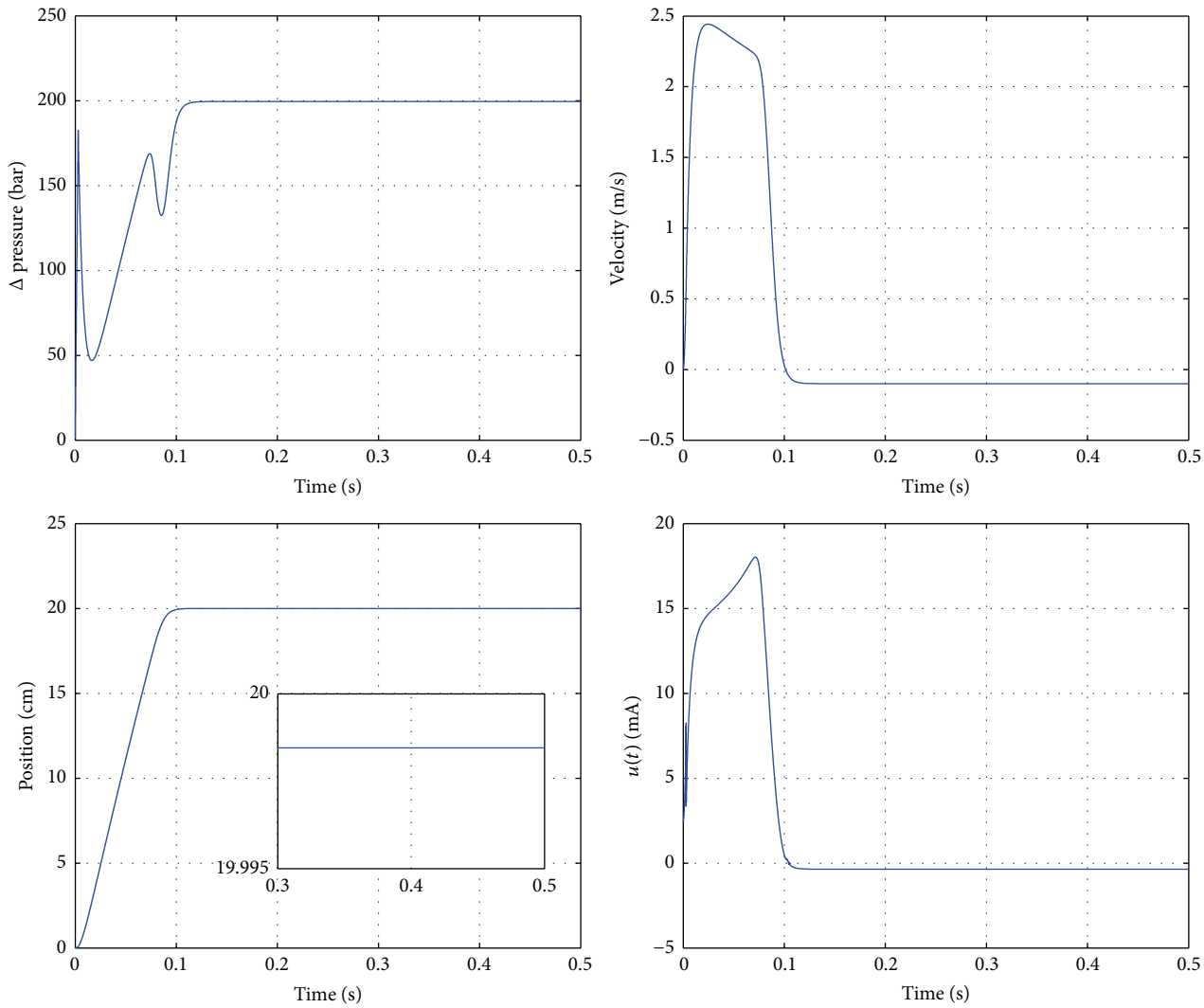


FIGURE 5: Behavior of system under sliding mode controller using nonlinear surface defined by (33) and a smooth saturation function tanh(·).

To prove the efficiency of the observer and the fact that the estimated states based controller can also achieve accurate positioning in presence of perturbation and uncertainty, we will proceed by a step-by-step proof.

Step 1. Let $e_3 = x_3 - z_3$ and let $e_2 = x_2 - z_2$; from (3) and (39) the error dynamics are expressed as

$$\dot{e}_3 = e_2 + d(t) - L_3 \operatorname{sign}(e_3). \tag{41}$$

Thus, if L_3 is chosen such that

$$L_3 > \sup_{t>0} \{e_2(t) + d(t)\} \tag{42}$$

then a sliding mode is established at the observer sliding surface $e_3 = 0$ within a finite time. Moreover, at sliding mode, we obtain $\dot{e}_3 = 0$ and thus from (41) we have

$$0 = x_2 - z_2 + d(t) - L_3 \operatorname{sign}(e_3); \tag{43}$$

that is,

$$x_2 + d(t) = z_2 + L_3 \operatorname{sign}(e_3) = \tilde{z}_2. \tag{44}$$

Now, if the sliding surface σ_1 of the nonlinear surface based SMC is expressed in terms of the estimated state $\tilde{z}_2$,

$$\begin{aligned}\sigma_1(\tilde{z}_2, x_3) = \tilde{z}_2 + W_3(x_3 - x_{3\text{ref}}) \\ + W_2 \operatorname{sign}(x_3 - x_{3\text{ref}}),\end{aligned} \tag{45}$$

then, by using (3), (44), and (45), the position dynamics are given by

$$\dot{x}_3 = -W_3(x_3 - x_{3\text{ref}}) - W_2 \operatorname{sign}(x_3 - x_{3\text{ref}}). \tag{46}$$

Therefore, within a finite time we obtain $x_3 = x_{3\text{ref}}$.

Step 2. Let $e_1 = x_1 - z_1$; from (2) and (38) the error dynamics are expressed as

$$\begin{aligned}\dot{e}_2 = \frac{1}{m_t}(Se_1 - be_2 - bL_3 \operatorname{sign}(e_3) - k_l e_3) - \frac{\Delta k_l}{m_t} x_3 \\ - L_2 \operatorname{sign}(\tilde{z}_2 - z_2).\end{aligned} \tag{47}$$

Thus, if L_2 is chosen such that

$$L_2 > \frac{1}{m_t}\left|\Delta k_l x_3 + S e_1 + b d_{\max}\right| \tag{48}$$

then we can deduce that $e_2 + d$ tend to zero and hence z_2 tend to $x_2 + d$. That is, the observer will estimate the velocity with the constant perturbation. Moreover, when the error dynamics are sliding on $e_3 = 0$ and $e_2 + d = 0$, then $\dot{e}_2 = 0$ and we get

$$\tilde{z}_1 = x_1 + \frac{b}{S} e_2 - \frac{\Delta k_l}{S} x_3. \tag{49}$$

That is, the observer will estimate the differential pressure with a constant difference proportional to the perturbation and the uncertainty. Now, if the controller sliding surface σ is expressed using the observer state variables,

$$\begin{aligned}\sigma(z) = {} & S\tilde{z}_1 + \left(m_t W_3 + m_t W_2 \delta\left(x_3 - x_{3\mathrm{ref}}\right) - b\right)\tilde{z}_2 \\ & - k_l z_3 + m_t W_1 \operatorname{sign}\left(\sigma_1\left(\tilde{z}_2, x_3\right)\right),\end{aligned} \tag{50}$$

then, at sliding mode of the controller $\dot{\sigma}(z) = 0$, the velocity dynamics are given by

$$\dot{x}_2 = -W_3\left(x_2 + d\right) - W_1 \operatorname{sign}\left(\sigma_1\left(\tilde{z}_2, x_3\right)\right). \tag{51}$$

Hence the velocity of the system rod will tend to $-d$.

Step 3. Finally, the differential pressure error dynamics can be roughly expressed as

$$\dot{e}_1 = f_1(x) - f_1(z) - L_1 \operatorname{sign}\left(e_1 + \frac{b}{S} e_2 - \frac{\Delta k_l}{S} x_3\right); \tag{52}$$

if the system is already sliding on the surfaces $x_3 = x_{3\mathrm{ref}}$ and $e_2 = -d$ then by choosing $L_1 > |f_1(x) - f_1(z)|$ the differential pressure is estimated with the constant difference $-(b/S)d + \Delta V_t$.

Figure 6 shows the convergence of the observer state z_3 to x_3 although they are starting from different initial conditions; indeed, the observer is starting at rod position 10 cm whereas the system is starting at the origin. As expected from the above analysis, z_2 tends to $x_2 + d$ and z_1 tends to x_1 with a constant difference of 33 bar. The observer gains are chosen as $L_1 = 1000$, $L_2 = 100$, and $L_3 = 1$. In Figure 7, the observer and system behaviors are shown with the controller being calculated using the estimated states. The observer initial position is −10 cm.

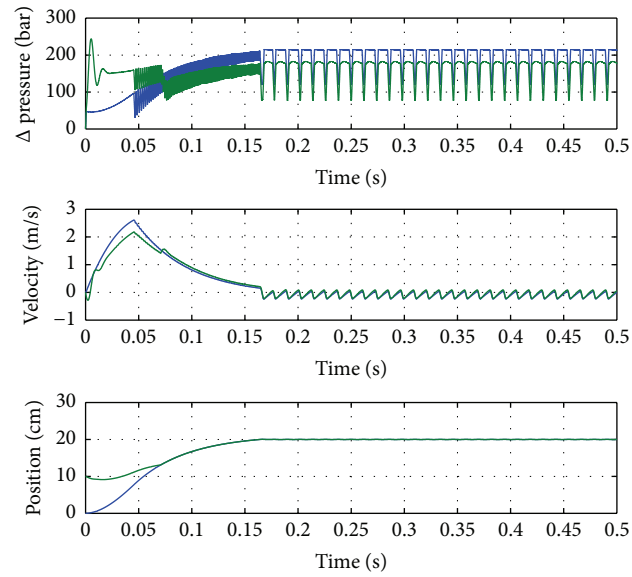


Figure 6: Behavior of the observer and the hydraulic servo system controlled with sliding mode controller with states feedback.

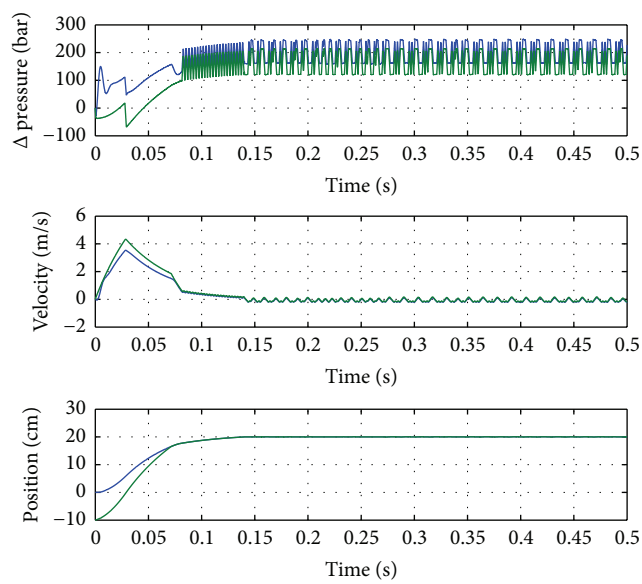


Figure 7: Behavior of the observer and the hydraulic servo system controlled with sliding mode controller with estimated states feedback.

5. Simulation Results Analysis

The controllers designed in this paper used the sliding mode theory which is the most used approach to deal with systems running under uncertainty conditions. However, we have seen in the second section that the first-order sliding mode controller with a linear sliding surface is not robust with respect to perturbation and mismatched uncertainty. In fact, by using the fact that on the sliding surface the system behaves in a similar way to a linear second-order system, it can be easily shown that as far as the sliding motion is preserved, the system is asymptotically stable if the closed loop eigenvalues are chosen as in (11) and thus the steady state error due to the uncertainty is expressed by (12). Also, the steady state error due to the constant perturbation is given by (13).

Therefore, the steady state error gets smaller as the closed loop eigenvalues have a larger amplitude. This is illustrated by the simulation results delineated in Figure 2 where $\lambda = 50$ and the total error is 2.46 cm.

To circumvent the problem, we have suggested a second-order sliding mode controller. Applying this controller, we

can notice that it achieves a better performance than the first SMC. In fact, the second-order SMC outperforms the first-order SMC in sense of robustness with respect to mismatched perturbation but it cannot guarantee the robustness with respect to the constant perturbation. The steady state error due to the perturbation is equal to $d(t)/h$ as demonstrated previously. When the perturbation $d(t) = 0.1$ the error is equal to 0.1 cm as delineated in Figure 3. If we increase the value of h, we should also increase the value of the variable structure gains W_1 and W_2. So we obtain a chattering phenomenon.

To overcome the problem of mismatched perturbation and uncertainty, we have suggested a sliding mode controller with nonlinear surface. The idea is based on the backstepping method and using a robust variable structure virtual controller. The obtained results achieved robustness with respect to parameter uncertainty and perturbation. The zoomed curve on Figure 4 shows that we have attained our aim to drive the hydraulic servo system to a reference position but with a chattering phenomenon. So to attenuate this problem we suggest substituting the discontinuous function $\text{sign}(x)$ with a smooth saturation function $\tanh(\mu x)$ for sufficiently high positive value $\mu > 0$. Owing to this smooth function, we can achieve the reference value in presence of mismatched uncertainty and perturbation with attenuated chattering.

Finally, in Section 4, we designed a sliding mode observer in order to estimate the required states. The efficiency of this observer is proved mathematically and also by simulations presented in Figures 6 and 7.

Compared to other methods such as that presented in [17], which is based on adaptive approach to achieve position control, we notice that our method provides faster response since the adaptive controller proposed therein attempts to estimate the mismatched nonlinear uncertainty. Indeed, the rod makes a displacement of 30 cm within more than 0.8 s, whereas with our method a maximum of 0.2 s will be needed to make the same displacement. In [21], the authors presented a controller based on variable structure PID to drive the hydraulic system position. Although fast response has been achieved, which is comparable to our results, the good robustness was obtained since the uncertainties and the perturbation were matched with the controller. In our case, the robustness against the mismatched character of the perturbation and uncertainty presented the main challenge to take. In [26], authors obtained comparable results to those presented in this paper using integral sliding mode controller with realizable reference compensation applied to an asymmetric piston.

6. Conclusion

In this paper, we developed several controllers based on the sliding mode theory. Our aim was to control the position of a hydraulic servo system piston in presence of mismatched uncertainty and perturbation. We have shown that a first-order sliding mode controller did not achieve any robustness. Next we have developed a second-order sliding mode controller that has shown robustness with respect to parametric uncertainty but was not robust to the perturbation. Finally, we have suggested a sliding mode controller based on a nonlinear sliding surface. The design is based on the backstepping method where on each step a variable structure virtual controller design leads to the design of the sliding surface. This controller emphasized the chattering phenomenon due to the nested sliding modes. As a remedy, we suggested substituting the discontinuous function with a smooth saturation function. Eventually, we have designed a robust sliding observer in order to substitute the unmeasured states with their estimates. A step-by-step proof has shown that the controller issued from the estimated states achieved the position tracking.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] H. E. Merritt, *Hydraulic Control Systems*, John Wiley & Sons, New York, NY, USA, 1967.

[2] M. Feki and E. Richard, "Including leakage flow in the servovalve static model," *International Journal of Modelling and Simulation*, vol. 25, no. 1, pp. 4004–4009, 2005.

[3] A. Alleyne, R. Liu, and H. Wright, "On the limitations of force tracking control for hydraulic active suspensions," in *Proceedings of the American Control Conference (ACC '98)*, vol. 1, pp. 43–47, June 1998.

[4] Y. Cheng and B. L. R. De Moor, "Robustness analysis and control system design for a hydraulic servo system," *IEEE Transactions on Control Systems Technology*, vol. 2, no. 3, pp. 183–197, 1994.

[5] M. Feki, E. Richard, and F. Gomes Almeida, "Commande en effort d'un vèrin hydraulique par linèarization entrèe/sortie," in *Journèes Doctorales d'Automatiques (JDA '99)*, pp. 181–184, Nancy-France, 1999.

[6] H. Hahn, A. Piepenbrink, and K.-D. Leimbach, "Input/output linearization control of an electro servo-hydraulic actuator," in *Proceedings of the 3rd IEEE Conference on Control Applications*, pp. 995–1000, August 1994.

[7] M. Jovanovic, "Nonlinear control of an electrohydraulic velocity servosystem," in *Proceedings of the 2002 American Control Conference*, vol. 1, pp. 588–593, 2002.

[8] G. A. Sohl and J. E. Bobrow, "Experiments and simulations on the nonlinear control of a hydraulic servosystem," *IEEE Transactions on Control Systems Technology*, vol. 7, no. 2, pp. 238–247, 1999.

[9] A. G. Alleyne and R. Liu, "Systematic control of a class of nonlinear systems with application to electrohydraulic cylinder pressure control," *IEEE Transactions on Control Systems Technology*, vol. 8, no. 4, pp. 623–634, 2000.

[10] M. A. Avila, A. G. Loukianov, and E. N. Sanchez, "Electrohydraulic actuator trajectory tracking," in *Proceedings of the American Control Conference (AAC '04)*, vol. 3, pp. 2603–2608, IEEE, Boston, Mass, USA, July 2004.

[11] A. Bonchis, P. I. Corke, D. C. Rye, and Q. P. Ha, "Variable structure methods in hydraulic servo systems control," *Automatica*, vol. 37, no. 4, pp. 589–595, 2001.

[12] M. Jerouane and F. Lamnabhi-Lagarrigue, "A new robust sliding mode controller for a hydraulic actuator," in *Proceedings of the*

40th IEEE Conference on Decision and Control (CDC '01), vol. 1, pp. 908–913, December 2001.

[13] H.-M. Chen, J.-C. Renn, and J.-P. Su, "Sliding mode control with varying boundary layers for an electro-hydraulic position servo system," *The International Journal of Advanced Manufacturing Technology*, vol. 26, no. 1-2, pp. 117–123, 2005.

[14] J. M. Zheng, S. D. Zhao, and S. G. Wei, "Application of self-tuning fuzzy PID controller for a SRM direct drive volume control hydraulic press," *Control Engineering Practice*, vol. 17, no. 12, pp. 1398–1404, 2009, Special Section: The 2007 IFAC Symposium on Advances in Automotive Control.

[15] A. Aly, "Pid parameters optimization using genetic algorithm technique for electrohydraulic servo control system," *Intelligent Control and Automation*, vol. 2, no. 2, pp. 69–76, 2011.

[16] B. Yao, F. Bu, and G. T. C. Chiu, "Nonlinear adaptive robust control of electro-hydraulic servo systems with discontinuous projections," in *Proceedings of the 37th IEEE Conference on Decision and Control (CDC '98)*, pp. 2265–2270, December 1998.

[17] B. Yao, F. Bu, J. Reedy, and G. T.-C. Chiu, "Adaptive robust motion control of single-rod hydraulic actuators: theory and experiments," *IEEE/ASME Transactions on Mechatronics*, vol. 5, no. 1, pp. 79–91, 2000.

[18] K. K. Ahn and Q. T. Dinh, "Self-tuning of quantitative feedback theory for force control of an electro-hydraulic test machine," *Control Engineering Practice*, vol. 17, no. 11, pp. 1291–1306, 2009.

[19] E. Kolsi Gdoura, M. Feki, and N. Derbel, "Sliding mode control of a hydraulic servo system position using adaptive sliding surface and adaptive gain," *International Journal of Modelling, Identification and Control*, vol. 23, pp. 211–219, 2015.

[20] C. Guan and S. Pan, "Adaptive sliding mode control of electro-hydraulic system with nonlinear unknown parameters," *Control Engineering Practice*, vol. 16, no. 11, pp. 1275–1284, 2008.

[21] D. Liu, Z. Tang, and Z. Pei, "Variable structure compensation PID control of asymmetrical hydraulic cylinder trajectory tracking," *Mathematical Problems in Engineering*, vol. 2015, Article ID 890704, 9 pages, 2015.

[22] A. G. Loukianov, J. Rivera, Y. V. Orlov, and E. Y. Morales Teraoka, "Robust trajectory tracking for an electrohydraulic actuator," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 9, pp. 3523–3531, 2009.

[23] M. Ye, Q. Wang, and S. Jiao, "Robust H_2/H_∞ control for the electrohydraulic steering system of a four-wheel vehicle," *Mathematical Problems in Engineering*, vol. 2014, Article ID 208019, 12 pages, 2014.

[24] W. Shen, J. Jiang, H. R. Karimi, and X. Su, "Observer-based robust control for hydraulic velocity control system," *Mathematical Problems in Engineering*, vol. 2013, Article ID 689132, 9 pages, 2013.

[25] E. Kolsi-Gdoura, M. Feki, and N. Derbel, "Sliding mode-based robust position control of an electrohydraulic system," in *Proceedings of the 10th International Multi-Conference on Systems, Signals and Devices (SSD '13)*, pp. 1–5, March 2013.

[26] E. Kolsi Gdoura, M. Feki, and N. Derbel, "Control of a hydraulic servo system using sliding mode with an integral and realizable reference compensation," *Control Engineering and Applied Informatics*, vol. 17, pp. 111–119, 2015.

[27] A. Levant, "Sliding order and sliding accuracy in sliding mode control," *International Journal of Control*, vol. 58, no. 6, pp. 1247–1263, 1993.

[28] A. Levant, "Higher-order sliding modes, differentiation and output-feedback control," *International Journal of Control*, vol. 76, no. 9-10, pp. 924–941, 2003.

[29] M. Djemai, K. Busawon, K. Benmansour, and A. Marouf, "High-order sliding mode control of a DC motor drive via a switched controlled multi-cellular converter," *International Journal of Systems Science*, vol. 42, no. 11, pp. 1869–1882, 2011.

[30] M. Lavanya, R. M. Brisilla, and V. Sankaranarayanan, "Higher order sliding mode control of permanent magnet dc motor," in *Proceedings of the 12th International Workshop on Variable Structure Systems (VSS '12)*, pp. 226–230, January 2012.

[31] S. Laghrouche, M. Smaoui, F. Plestan, and X. Brun, "Higher order sliding mode control based on optimal approach of an electropneumatic actuator," *International Journal of Control*, vol. 79, no. 2, pp. 119–131, 2006.

[32] A. G. Loukianov, E. Sanchez, and C. Lizalde, "Force tracking neural block control for an electro-hydraulic actuator via second-order sliding mode," *International Journal of Robust and Nonlinear Control*, vol. 18, no. 3, pp. 319–332, 2008.

[33] J. Komsta, N. van Oijen, and P. Antoszkiewicz, "Integral sliding mode compensator for load pressure control of die-cushion cylinder drive," *Control Engineering Practice*, vol. 21, no. 5, pp. 708–718, 2013.

[34] M. Feki, *Synthèse de commandes et d'observateurs pour les systèmes non-linèaires: application aux systèmes hydrauliques [Ph.D. thesis]*, Universitè de Metz, Metz, France, 2001.

[35] V. V. Filippov, "On a scalar equation with discontinuous right-hand side and the uniqueness theorem," *Differential Equations*, vol. 38, no. 10, pp. 1435–1445, 2002.

[36] T. Boukhobza, M. Djemai, and J. P. Barbot, "Nonlinear sliding observer for systems in output and output derivative injection form," in *Proceedings of the IFAC World Congress*, pp. 299–305, San Antonio, Tex, USA, 1996.

[37] H. Saadaoui, N. Manamanni, M. Djemaï, J. P. Barbot, and T. Floquet, "Exact differentiation and sliding mode observers for switched Lagrangian systems," *Nonlinear Analysis, Theory, Methods & Applications*, vol. 65, no. 5, pp. 1050–1069, 2006.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 143739, 12 pages
http://dx.doi.org/10.1155/2015/143739

Research Article

Optimal Limited Stop-Loss Reinsurance under VaR, TVaR, and CTE Risk Measures

Xianhua Zhou,[1] Huadong Zhang,[2] and Qingquan Fan[3]

[1] *China Institute for Actuarial Science, Central University of Finance and Economics, Beijing 100081, China*
[2] *Research Institute of Applied Mathematics, Anhua Agricultural Insurance Co., Ltd., Beijing 100037, China*
[3] *School of Economics and Management, Tsinghua University, Beijing 100084, China*

Correspondence should be addressed to Huadong Zhang; zhanghuadona@126.com

Received 27 April 2015; Revised 14 July 2015; Accepted 15 July 2015

Academic Editor: Xinguang Zhang

This paper aims to provide a practical optimal reinsurance scheme under particular conditions, with the goal of minimizing total insurer risk. Excess of loss reinsurance is an essential part of the reinsurance market, but the concept of stop-loss reinsurance tends to be unpopular. We study the purchase arrangement of optimal reinsurance, under which the liability of reinsurers is limited by the excess of loss ratio, in order to generate a reinsurance scheme that is closer to reality. We explore the optimization of limited stop-loss reinsurance under three risk measures: value at risk (VaR), tail value at risk (TVaR), and conditional tail expectation (CTE). We analyze the topic from the following aspects: (1) finding the optimal franchise point with limited stop-loss coverage, (2) finding the optimal limited stop-loss coverage within a certain franchise point, and (3) finding the optimal franchise point with limited stop-loss coverage. We provide several numerical examples. Our results show the existence of optimal values and locations under the various constraint conditions.

1. Introduction

Reinsurance, an agreement between insurers and reinsurers that allows insurers to transfer and diversify away a certain amount of risk, is the primary risk management tool used by insurance companies. The amount that an insurer pays to transfer risk to the reinsurer is known as the reinsurance premium. The losses caused by accidents that meet the requirements in the reinsurance contract (and that are borne by the reinsurer) are known as reinsurance recoverable. The insurer aims to reduce its compensation expenses to the greatest extent possible.

However, the situation is ever-changing, and optimal reinsurance has become a popular topic for both researchers and practitioners. This has resulted in a plethora of important insights. The earliest study on optimal reinsurance focused on safety loading, Borch [1]. A later article by Gerber [2] showed that excess of loss reinsurance is optimal at the expected value of the reinsurance principle. Gajek and Zagrodny [3] found that the optimal reinsurance form is the minimum variance under the standard deviation reinsurance principle, and Kaluszka [4] also showed that change loss reinsurance is optimal through mean-variance analysis of optimal reinsurance. Other researchers, such as Bu [5], have suggested that insurers wishing to determine the optimal trade-off between retained risk and expected profits should purchase middle layer reinsurance, because the purchase of reinsurance for high loss layers is not usually economically feasible. Cai and Tan [6] determined the optimal retention rate by minimizing VaR and CTE for insurers, while Cai et al. [7] found that the optimal reinsurance form is based on the principle of minimum VaR or CTE and that results differ under the two methods. Under CTE, excess of loss reinsurance is always the best product; under VaR, the answer is more complicated. Cai et al. [7] discussed this condition further for stop-loss reinsurance, but several issues remain, such as moral risk, overly high expenses, and a lack of stop-loss reinsurance products due to the extremely high risk for some types of insurance (e.g., crop insurance). The premium rates are relatively high because of the large reinsurance coverage, which in turn affects the insurance company's profits. Weng [8] attempts to fit an optimal reinsurance model in which

reinsurance budget constraints are given, using the principles of minimum CTE and standard deviation reinsurance. He notes that, when the reinsurance premium is subject to a small budget, the best form is likely to be limited stop-loss reinsurance, not standard stop-loss reinsurance. This is because, if the purchase reinsurance is insufficient, the primary insurer may become insolvent or even go bankrupt in the aftermath of several back-to-back catastrophes, Fu and Khury [9]. Chi and Tan [10] study the class of increasing convex ceded loss functions based on VaR and CVaR without specific forms of reinsurance. Their conclusions confirm the tenets of stop-loss reinsurance, except for Theorem 3.2 in [10]. To avoid the problem of moral hazards, Chi and Weng [11] conduct research under the Vajda condition, which stipulates that both the insurer's retained loss and the proportion paid by a reinsurer are increasing in indemnity. Other researchers, such as Gajek and Zagrodny [12] and Guerra and Centeno [13], prefer the theory of stop-loss reinsurance. More recent findings have been consistent with the notion that stop-loss reinsurance is the optimal product. Weng [8] and Porth et al. [14] add the reinsurance premium budget constraint to their model. The reinsurance premium equation must be solved when subject to the budget constraint, so the condition leads to a more complex problem. Therefore, we do not consider the reinsurance premium budget constraint here. Li et al. [15] study the optimal reinsurance and investment problem by capturing both the insurer's and the reinsurer's utility. Brandtner and Kürsten [16] investigate the problem of optimal reinsurance of risk management within the regulatory framework of Solvency II, under conditional VaR and spectral risk measures as its natural extension. Li et al. [15] modeled the risk process by Brownian motion with drift and studied the optimization problem of maximizing the exponential utility of terminal wealth under the controls of reinsurance and investment. Their results showed that optimal excess of loss reinsurance is generally a better product than optimal proportional reinsurance.

Thus, limited stop-loss reinsurance, which is the focus of this study, is the most practical real-world solution. In this study, we regard stop-loss reinsurance as a special case. We aim to solve the problem and provide optimal reinsurance advice to achieve optimal risk transfer under different risk measurement models and limiting conditions. The remainder of this paper is organized as follows. Section 2 presents the concepts and formulas related to reinsurance, as well as the formula of a reinsurance risk measurement model. Sections 3 and 4 explore optimal reinsurance arrangements and possible problems under VaR, TVaR, and CTE. Section 5 provides numerical examples and analyzes reinsurance premium constraints. Section 6 concludes.

2. Basic Theory of Reinsurance and Risk Measures

2.1. Limited Stop-Loss Reinsurance Theory. To illustrate the concept behind limited stop-loss reinsurance, consider an example involving coverage for the total amount of claim X over one year. The reinsurer will pay the percentage of X that exceeds a certain amount, for example, franchise point d.

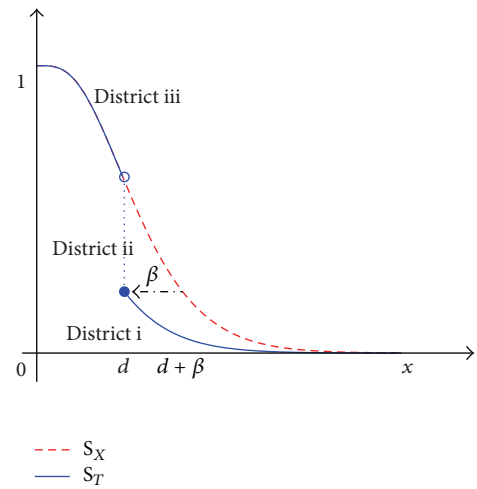


FIGURE 1: Relation between S_X and S_I.

The reinsurer's liability is limited to that amount (e.g., limited stop-loss coverage β). In general, $d \in (0, \infty)$ and $\beta \in (0, \infty)$. We let X_I be the loss random variable of the cedent in the presence of limited stop-loss reinsurance, and we then have

$$X_I = \begin{cases} X, & X < d, \\ d, & d \le X \le d + \beta, \\ X - \beta, & X > d + \beta. \end{cases} \tag{1}$$

For stop-loss reinsurance, $\beta = \infty$. In other words, the reinsurer pays the excess of loss over d.

In this paper, subscripts X, I, and T stand for the underlying loss, the retained loss, and the total loss of the cedent in the presence of limited stop-loss reinsurance. Survival function $S_X(x) = \Pr\{X > x\} = 1 - F_X(x)$, which is commonly used in actuarial science (for more details, see Cai and Tan [6]). Given that cumulative distribution function $F_X(x)$ is discontinuous, we use a strict inequality in survival function S_X and in the risk measure. According to (1), we therefore have

$$S_I = \begin{cases} S_X(x), & 0 \le x < d, \\ S_X(x + \beta), & x \ge d. \end{cases} \tag{2}$$

Figure 1 shows the relationship between S_X and S_I, where S_I is discontinuous. According to the jump of S_I, the entire plane is divided into three districts, as Figure 1 shows.

Let $\Pi(d, \beta)$ be the pure risk premium for the reinsurer as follows:

$$\begin{aligned} \Pi(d, \beta) &= \int_d^{d+\beta} (x - d) f_X(x)\, dx + \beta [1 - F(d + \beta)] \\ &= \int_d^{d+\beta} S_X(x)\, dx, \end{aligned} \tag{3}$$

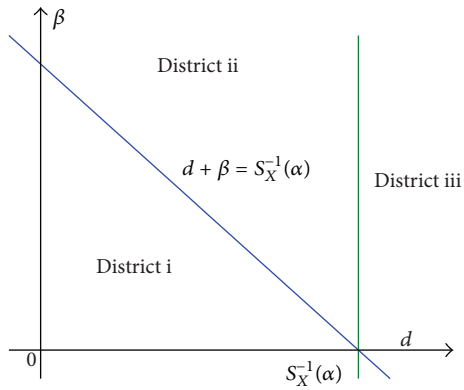


Figure 2: Plane (d, β).

where $f_X(x)$ is the probability density function of X. Then, according to the expected value premium principle, the reinsurance premium $\delta(d, \beta)$ is

$$\delta(d, \beta) = (1 + \rho)\,\Pi(d, \beta), \tag{4}$$

where $\rho > 0$ is the safety loading. Hence, total risk T is

$$X_T = X_I + \delta(d, \beta). \tag{5}$$

2.2. Risk Measure after Reinsurance. In reinsurance studies, the risk measure is used to ensure optimal decision making (see, e.g., Cai et al. [7] and Bernard and Tian [17]). In risk management, several models, such as VaR, CTE, TVaR, CVaR, and ES, are commonly used. Within a certain time period and at a confidence level of $1 - \alpha$, $0 < \alpha < 1$, the highest risk value of X does not exceed VaR; CTE is the expected value of events that occur outside the probability alpha. According to the consistency axiom of risk measure defined by Artzner et al. [18], several of these risk measures do not satisfy the consistency axiom. For example, VaR does not meet the additivity condition.

According to Hang [19] and the combined characteristic of X_I, we have

$$\begin{aligned}\mathrm{VaR}_I(d, \alpha, \beta) &= \inf\{x \mid \Pr(X_I > x) \le \alpha\}\\ &= \inf\{x \mid S_I(x) \le \alpha\}\\ &= \begin{cases} S_X^{-1}(\alpha) - \beta, & 0 < d + \beta < S_X^{-1}(\alpha),\ \text{i},\\ d, & d \le S_X^{-1}(\alpha) \le d + \beta,\ \text{ii},\\ S_X^{-1}(\alpha), & S_X^{-1}(\alpha) < d,\ \text{iii}.\end{cases}\end{aligned} \tag{6}$$

For conciseness, we denote $\mathrm{VaR}_T^{\mathrm{i}}$ as VaR_T in district i throughout the remainder of paper; the variables with no subscript stand for those in the entire district. The rest of the variables are similar. S_I is divided into three districts. Similarly, we divide plane (d, β) into three districts: i, ii, and iii (Figure 2). The boundaries of i and ii and of ii and iii are classified in ii.

From (5) and (6), we therefore have

$$\mathrm{VaR}_T(d, \alpha, \beta) = \mathrm{VaR}_I(d, \alpha, \beta) + \delta(d, \beta), \tag{7}$$

where VaR_I is continuous on the plane and $\delta(d, \beta)$ is continuous. Hence, VaR_T is also continuous on the plane.

From (5) and (7), we can easily prove $S_T(\mathrm{VaR}_T(d, \alpha, \beta)) = S_I(\mathrm{VaR}_I(d, \alpha, \beta))$. In district i, we have $0 < d + \beta < S_X^{-1}(\alpha)$, and $\mathrm{VaR}_I(d, \alpha, \beta) = S_X^{-1}(\alpha) - \beta > d$. With (2), we have

$$\begin{aligned}&S_I(\mathrm{VaR}_I(d, \alpha, \beta)) = S_I\left(S_X^{-1}(\alpha) - \beta\right) = \alpha,\\ &E\left[(X_I - \mathrm{VaR}_I(d, \alpha, \beta))^+\right]\\ &\quad = \int_{\mathrm{VaR}_I(d,\alpha,\beta)}^{\infty} (x - \beta - \mathrm{VaR}_I(d, \alpha, \beta))\, f_X(x)\, dx\\ &\quad = \int_{S_X^{-1}(\alpha)}^{\infty} S_X(x)\, dx.\end{aligned} \tag{8}$$

In district ii, we have $d \le S_X^{-1}(\alpha) \le d + \beta$ and $\mathrm{VaR}_I(d, \alpha, \beta) = d$. With (2), we therefore have

$$\begin{aligned}&S_I(\mathrm{VaR}_I(d, \alpha, \beta)) = S_I(d) = S_X(d + \beta),\\ &E\left[(X_I - \mathrm{VaR}_I(d, \alpha, \beta))^+\right]\\ &\quad = \int_{d+\beta}^{\infty} (x - \beta - d)\, f_X(x)\, dx = \int_{d+\beta}^{\infty} S_X(x)\, dx.\end{aligned} \tag{9}$$

In district iii, we have $S_X^{-1}(\alpha) < d$ and $\mathrm{VaR}_I(d, \alpha, \beta) = S_X^{-1}(\alpha) < d$. With (2), we have

$$\begin{aligned}&S_I(\mathrm{VaR}_I(d, \alpha, \beta)) = S_I\left(S_X^{-1}(\alpha)\right) = \alpha,\\ &E\left[(X_I - \mathrm{VaR}_I(d, \alpha, \beta))^+\right]\\ &\quad = \int_{S_X^{-1}(\alpha)}^{d} \left(x - S_X^{-1}(\alpha)\right) f_X(x)\, dx\\ &\qquad + \left[d - S_X^{-1}(\alpha)\right]\left[F_X(d + \beta) - F_X(d)\right]\\ &\qquad + \int_{d+\beta}^{\infty} \left(x - \beta - S_X^{-1}(\alpha)\right) f_X(x)\, dx\\ &\quad = \int_{S_X^{-1}(\alpha)}^{d} S_X(x)\, dx + \int_{d+\beta}^{\infty} S_X(x)\, dx.\end{aligned} \tag{10}$$

We obtain the expression of ES as follows:

$$\begin{aligned}\mathrm{ES}_I(d, \alpha, \beta) &= E\left[(X_I - \mathrm{VaR}_I(d, \alpha, \beta))^+\right]\\ &= \begin{cases} \displaystyle\int_{S_X^{-1}(\alpha)}^{\infty} S_X(x)\, dx, & \text{i},\\ \displaystyle\int_{S_X^{-1}(\alpha)}^{\infty} S_X(x)\, dx - \int_{S_X^{-1}(\alpha)}^{d+\beta} S_X(x)\, dx, & \text{ii},\\ \displaystyle\int_{S_X^{-1}(\alpha)}^{\infty} S_X(x)\, dx - \int_{d}^{d+\beta} S_X(x)\, dx, & \text{iii}.\end{cases}\end{aligned} \tag{11}$$

And, with (5) and (6), we obtain

$$\begin{aligned}\mathrm{ES}_T(d,\alpha,\beta) &= E\left[\left(X_T-\mathrm{VaR}_T(d,\alpha,\beta)\right)^+\right]\\ &= E\left[\left(X_I-\mathrm{VaR}_I(d,\alpha,\beta)\right)^+\right]\\ &= \mathrm{ES}_I(d,\alpha,\beta).\end{aligned} \tag{12}$$

According to the transformation among risk measure models (Hang [19] and Charpentier [20]), we have

$$\begin{aligned}\mathrm{TVaR}_I(d,\alpha,\beta) &= \inf_{a>0}\left\{a+\frac{1}{\alpha}E\left[\left(X_I-a\right)^+\right]\right\}\\ &= \mathrm{VaR}_I(d,\alpha,\beta)+\frac{1}{\alpha}\mathrm{ES}_I(d,\alpha,\beta),\\ \mathrm{CVaR}_I(d,\alpha,\beta)&\\ &= E\left[X_I-\mathrm{VaR}_I(d,\alpha,\beta)\mid X_I>\mathrm{VaR}_I(d,\alpha,\beta)\right]\\ &= \frac{1}{S_I\left(\mathrm{VaR}_I(d,\alpha,\beta)\right)}\mathrm{ES}_I(d,\alpha,\beta),\\ \mathrm{CTE}_I(d,\alpha,\beta) &= E\left[X_I\mid X_I>\mathrm{VaR}_I(d,\alpha,\beta)\right]\\ &= \mathrm{VaR}_I(d,\alpha,\beta)\\ &\quad+\frac{1}{S_I\left(\mathrm{VaR}_I(d,\alpha,\beta)\right)}\mathrm{ES}_I(d,\alpha,\beta).\end{aligned} \tag{13}$$

Obviously, TVaR, CTE, and CVaR can all be obtained by combining VaR, ES, S, and α. Equations (7) and (12) denote the transforming relationships between X_I and X_T of VaR and ES, respectively. Hence, no expressions of TVaR, CTE, or CVaR exist in X_T. For information risk measures, see Charpentier [20].

We define risk measures in this paper as per Hang [19], whose definitions are slightly different from those proposed by Wirch and Hardy [21]. According to $S_T = S_I$ and (12), ES and CVaR models do not include the information of reinsurance premium $\delta(d,\beta)$. Thus, they are not suitable for making reinsurance decisions, and no further analysis of them is needed in this context.

Under risk measure Λ and condition Θ, the aim is to solve the optimization problem

$$\min_{\theta\in\Theta}\Lambda_T(\theta). \tag{14}$$

Note that, in this study, $\theta = (d,\alpha,\beta)$ and $\Lambda \in \{\mathrm{VaR}, \mathrm{TVaR}, \mathrm{CTE}\}$, but Θ has different expressions.

3. Optimal Reinsurance under the VaR Risk Measure

We let $\Lambda = \mathrm{VaR}$, the most commonly used and concise risk measure in risk management. From (3), (4), (6), and (7), we therefore have

$$\begin{aligned}&\frac{\partial\mathrm{VaR}_T(d,\alpha,\beta)}{\partial d}\\ &\quad=\begin{cases}(1+\rho)\left(S_X(d+\beta)-S_X(d)\right)<0, & \text{i},\\ 1+(1+\rho)\left(S_X(d+\beta)-S_X(d)\right), & \text{ii},\\ (1+\rho)\left(S_X(d+\beta)-S_X(d)\right)<0, & \text{iii},\end{cases}\\ &\frac{\partial^2\mathrm{VaR}_T(d,\alpha,\beta)}{\partial d^2}=(1+\rho)\left(f_X(d)-f_X(d+\beta)\right),\\ &\qquad(d,\beta)\in\text{i, ii, iii},\\ &\frac{\partial\mathrm{VaR}_T(d,\alpha,\beta)}{\partial\beta}=\begin{cases}-1+(1+\rho)S_X(d+\beta), & \text{i},\\ (1+\rho)S_X(d+\beta)>0, & \text{ii},\\ (1+\rho)S_X(d+\beta)>0, & \text{iii},\end{cases}\\ &\frac{\partial^2\mathrm{VaR}_T(d,\alpha,\beta)}{\partial\beta^2}=-(1+\rho)f_X(d+\beta)<0,\\ &\qquad(d,\beta)\in\text{i, ii, iii},\\ &\frac{\partial^2\mathrm{VaR}_T(d,\alpha,\beta)}{\partial d\partial\beta}=-(1+\rho)f_X(d+\beta)<0,\\ &\qquad(d,\beta)\in\text{i, ii, iii}.\end{aligned} \tag{15}$$

In any district, VaR_T is monotonous for either d or β. In other words, it cannot form a stagnation point. Therefore, the optimal solution can only exist on the border.

3.1. Optimal d for Given β. At any given point in time on the reinsurance market, some reinsurance policies are not active. For example, the coverage of reinsurance in the market is generally given as $\beta = 0.2, 0.3$. For management purposes, company executives typically provide a definite coverage amount in advance, with the goal of finding the optimal franchise point for given reinsurance coverage β that will minimize total insurer risk. We let $\Theta = \{d \mid d > 0\}$. Therefore, the optimization problem must be addressed as follows:

$$\min_{d>0}\mathrm{VaR}_T(d,\alpha,\beta). \tag{16}$$

As we noted previously, VaR_T is continuous on the entire plane. From (15), we know that VaR_T is decreasing in i and iii, and

$$\mathrm{VaR}_T(\infty,\alpha,\beta)=\lim_{d\to\infty}\mathrm{VaR}_T(d,\alpha,\beta)=S_X^{-1}(\alpha), \tag{17}$$

where $\mathrm{VaR}_T(\infty, \alpha, \beta)$ represents the case where d approaches infinity and $\mathrm{VaR}_T^{\mathrm{ii}}(0, \alpha, \beta)$ represents 0. For the sake of space, we do not provide any detailed demonstrations here, however.

Next, we establish that

$$h(\beta) = (1+\rho)\int_{S_X^{-1}(\alpha)-\beta}^{S_X^{-1}(\alpha)} S_X(x)\,dx - \beta, \quad 0 \le \beta \le S_X^{-1}(\alpha), \tag{18}$$
$$h'(\beta) = (1+\rho) S_X\left(S_X^{-1}(\alpha) - \beta\right) - 1,$$
$$h''(\beta) = (1+\rho) f_X\left(S_X^{-1}(\alpha) - \beta\right) > 0.$$

Note that h is a strictly convex function in β. By letting $h'(\beta) = 0$, we can solve

$$\widetilde{\beta} = S_X^{-1}(\alpha) - S_X^{-1}\left(\frac{1}{1+\rho}\right). \tag{19}$$

We let $\widetilde{\beta} > 0$; that is,

$$\alpha * (1+\rho) < 1, \tag{20}$$

while, at the same time, $h(0) = 0$. By letting $h(\beta^*) = 0$, we can obtain the numerical solution of β^*. Thus, $\beta^* > \widetilde{\beta}$ under the condition of (20), and we have $\forall \beta \in [0, \beta^*]$, $h(\beta) \le 0$.

Note that (20) is very important, because, in general, safety loading ρ cannot be too large, or the risk distribution will be unsuitable. If $\rho = 1$, then $\alpha < 0.5$; if $\rho = 0.5$, then $\alpha < 2/3$. When the general consideration of α is relatively small and if the insurance company's risk tolerance level $\alpha = 0.5$, then it will be unsuitable for the excess of loss ratio reinsurance.

Let d^* be the solution of

$$S_X(d) - S_X(d+\beta) = \frac{1}{1+\rho}. \tag{21}$$

In this manner, if $(d^*, \beta) \in \mathrm{ii}$, then

$$\left.\frac{\partial \mathrm{VaR}_T(d, \alpha, \beta)}{\partial d}\right|_{(d^*,\beta)\in \mathrm{ii}} = 0. \tag{22}$$

The existence and uniqueness of d^* depends on the distribution of f_X and ρ. We define $\mathring{d} = \inf\{d : \forall \beta > 0, f_X(d) > f_X(d+\beta)\}$. In fact, $\mathring{d}$ is the rightmost peak of risk distribution f_X.

Lemma 1. *If (20) is satisfied, then $\forall \beta \in (0, \beta^*)$, VaR_T can achieve the optimum, and franchise d should be in district ii. Furthermore, (1) if $\beta \le S_X^{-1}(\alpha) - \mathring{d}$ and $d^* > S_X^{-1}(\alpha) - \beta$, then one can find the optimal VaR_T in $d = d^*$ and (2) if $\beta \le S_X^{-1}(\alpha) - \mathring{d}$ and $d^* \le S_X^{-1}(\alpha) - \beta$, one can find the optimal VaR_T in $d = S_X^{-1}(\alpha) - \beta$.*

Proof. As mentioned in the previous sections, districts i and iii are open sets. However, VaR_T is monotonically decreasing in d. Therefore, if an optimal VaR_T exists, it must be the optimal $\mathrm{VaR}_T^{\mathrm{ii}}$. In boundary district ii, $d = S_X^{-1}(\alpha) - \beta$, so, from (6), (7), and (20), we have

$$\mathrm{VaR}_T\left(S_X^{-1}(\alpha) - \beta, \alpha, \beta\right) - \mathrm{VaR}_T(\infty, \alpha, \beta) = h(\beta). \tag{23}$$

Moreover, as we can deduce from (20), $h(\beta) \le 0$. Thus, we know that $\mathrm{VaR}_T(S_X^{-1}(\alpha) - \beta, \alpha, \beta)$ is not larger than $\mathrm{VaR}_T(\infty, \alpha, \beta)$. In a given situation β, district ii about d is a closed set, so VaR_T in district ii will attain an optimal solution.

When $\beta \le S_X^{-1}(\alpha) - \mathring{d}$, $d \ge \mathring{d}$. With (15), we know that $\mathrm{VaR}_T^{\mathrm{ii}}$ about d is convex, which means that the optima of $\mathrm{VaR}_T^{\mathrm{ii}}$ are found at extreme points where the first derivative of $\mathrm{VaR}_T^{\mathrm{ii}}$ is 0. If $d^* \ge S_X^{-1}(\alpha)$, then $\mathrm{VaR}_T^{\mathrm{ii}}$ is monotonically decreasing in $[S_X^{-1}(\alpha) - \beta, S_X^{-1}(\alpha)]$. However, VaR_T on $d = S_X^{-1}(\alpha)$ is continuous, and it will begin to decrease from this point to $\mathrm{VaR}_T(\infty, \alpha, \beta)$. This condition is contrary to that found under $h(\beta) \le 0$. Hence, under the condition of (20), $d^* < S_X^{-1}(\alpha)$.

If $d^* > S_X^{-1}(\alpha) - \beta$ and $d^* \in (S_X^{-1}(\alpha) - \beta, S_X^{-1}(\alpha))$, VaR_T on $d = d^*$ will exist as an optimal solution. When $d^* \le S_X^{-1}(\alpha) - \beta$, $\mathrm{VaR}_T^{\mathrm{ii}}$ on d is increasing; therefore, when $d = S_X^{-1}(\alpha) - \beta$, we can find the optimal solution, which is also the optimal VaR_T. Figure 3 gives a graphical representation of $\min_{d>0} \mathrm{VaR}_T(d, \alpha, \beta)$. □

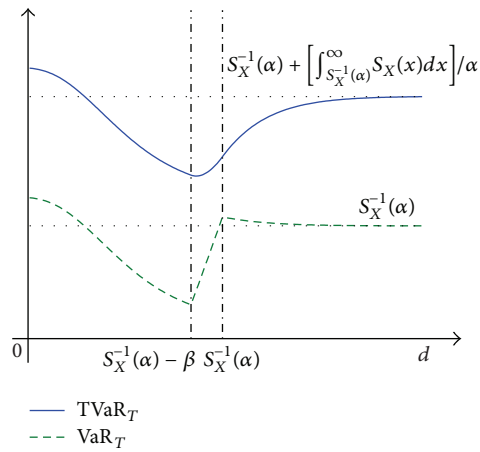


Figure 3: When β is given, VaR_T and TVaR_T satisfy (20).

The proof indicates the existence of d^* and implies that the relationship between d^* and $\mathring{d}$ only affects the position of the optimal solution, not its existence.

3.2. Optimal β for Given d. Note that insurance companies may consider a reinsurance position where the franchise point is the profit equilibrium point, because they can obtain a certain profit before the threshold. Beyond this threshold, profits will be reduced or even reach a deficit. Thus, the best option is appropriate reinsurance coverage β to achieve optimality. d is constant, and β is a variable. The optimization problem is $\Lambda = \mathrm{VaR}$, $\Theta = \{\beta \mid \beta > 0\}$.

Next, we let

$$g(d) = d + (1+\rho)\int_d^{S_X^{-1}(\alpha)} S_X(x)\,dx - S_X^{-1}(\alpha), \quad 0 \le d \le S_X^{-1}(\alpha), \tag{24}$$

$$g'(d) = 1 - (1+\rho) S_X(d),$$

$$g''(d) = (1+\rho) f_X(d) > 0,$$

where g is a strictly convex function in d. The solution of $g'(d) = 0$ is $\widetilde{d}$:

$$\widetilde{d} = S_X^{-1}\left(\frac{1}{1+\rho}\right), \quad \widetilde{d} < S_X^{-1}(\alpha) \longleftrightarrow (20). \tag{25}$$

Furthermore, $g(S_X^{-1}(\alpha)) = 0$. Hence, under (20), another solution is $\widetilde{d^*} < S_X^{-1}(1/(1+\rho))$, which can be obtained by using a numerical method. For $\forall d \in [\widetilde{d^*}, S_X^{-1}(\alpha)]$, we have $g(d) \le 0$.

Lemma 2. *If (20) is satisfied, $\forall d \in [\widetilde{d^*}, S_X^{-1}(\alpha)]$. VaR_T can then obtain the optimal value when $\beta = S_X^{-1}(\alpha) - d$:*

$$VaR_T\left(d, \alpha, S_X^{-1}(\alpha) - d\right) = d + (1+\rho)\int_d^{S_X^{-1}(\alpha)} S_X(x)\,dx. \tag{26}$$

$\forall d \notin [\widetilde{d^*}, S_X^{-1}(\alpha)]$, and no optimal solutions of β exist in VaR_T.

Proof. According to (15), when $d > S_X^{-1}(\alpha)$ (district iii, open set), then VaR_T will be increasing with regard to β; hence, no optimal value exists. However, when $\beta \ge S_X^{-1}(\alpha) - d$ (district ii, half closed), VaR_T will be increasing with regard to β, and the minimum will be reached at the lower boundary. VaR_T in district i is convex with regard to β. Thus, if the value obtained on the boundary of districts i and ii is not greater than the lower limit value, then we can obtain the optimal value, or the combination of properties of the convex function will indicate no optimal value.

$\forall d \in [\widetilde{d^*}, S_X^{-1}(\alpha)]$, and let $\beta = S_X^{-1}(\alpha) - d$:

$$\mathrm{VaR}_T\left(d, \alpha, S_X^{-1}(\alpha) - d\right) - \mathrm{VaR}_T(d, \alpha, \infty) = g(d) < 0. \tag{27}$$

The boundary of districts i and ii $d + \beta = S_X^{-1}(\alpha)$. Given that the value of VaR_T is less than the limit value of $S_X^{-1}(\alpha)$, this is optimal. However, $\forall d < \widetilde{d^*}$, $g(d) > 0$, so no optimal solution exists. Figure 4 gives a graphical representation of $\min_{\beta>0}\mathrm{VaR}_T(d, \alpha, \beta)$. □

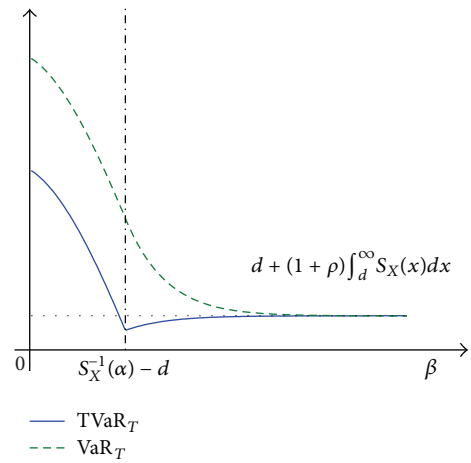


Figure 4: When d is given, VaR_T and TVaR_T satisfy (20).

3.3. Comprehensive Effect. After analyzing the optimal decision under given franchise point d or given reinsurance coverage β, we consider comprehensive effects with two variables. The optimization problem is $\Lambda = \mathrm{VaR}$, $\Theta = \{(d, \beta) \mid d > 0,\ \beta > 0\}$. The following analysis is under a function of d, and we obtain similar results if we perform it under a function of β.

Lemma 3. *Under the condition of (20), VaR_T can obtain the global optimal value on $(\widetilde{d}, S_X^{-1}(\alpha) - \widetilde{d})$.*

Proof. The optimal value of VaR_T only occurs on the border of districts i and ii. Lemma 2 ensures that $\mathrm{VaR}_T(d, \alpha, \beta) \le S_X^{-1}(\alpha)$ under the condition of (20) is not empty. We let $\forall (d, \beta) : d + \beta = S_X^{-1}(\alpha)$, and

$$\mathrm{VaR}_T(d, \alpha, \beta) = d + (1+\rho)\int_d^{S_X^{-1}(\alpha)} S_X(x)\,dx,$$
$$\frac{\partial \mathrm{VaR}_T(d, \alpha, \beta)}{\partial d} = 1 - (1+\rho) S_X(d), \tag{28}$$
$$\frac{\partial^2 \mathrm{VaR}_T(d, \alpha, \beta)}{\partial d^2} = (1+\rho) f_X(d) > 0.$$

VaR_T is convex in the border district, and we obtain the minimum value when the first derivative is 0. We do not consider $d = 0$ or ∞:

$$\frac{\partial \mathrm{VaR}_T(d, \alpha, \beta)}{\partial d} = 0 \longleftrightarrow d = S_X^{-1}\left(\frac{1}{1+\rho}\right). \tag{29}$$

The value of $\widetilde{d}$ falls within the district, as required by Lemma 2. Thus, under the condition of (20), VaR_T attains the global optimal value on $(\widetilde{d}, S_X^{-1}(\alpha) - \widetilde{d})$. □

Given a company's risk preference α and safety loading ρ, we can obtain the optimal solution on $\widetilde{d} = S_X^{-1}(1/(1+\rho))$, $\widetilde{\beta} = S_X^{-1}(\alpha) - S_X^{-1}(1/(1+\rho))$. An undesirable phenomenon thus exists: risk preference α cannot affect the optimal franchise

point. It can only affect the reinsurance coverage. Chi and Tan [10] set no constraint on $\beta = 0$ or $\beta = \infty$. Theorem 3.2 in [10] provides the same conclusion as Lemma 3.

4. Optimal Reinsurance under TVaR and CTE Risk Measures

4.1. Under TVaR Risk Measures. In this subsection, Λ = TVaR. Equation (13) provides the TVaR formula, and VaR_T and ES_T are both continuous on the entire plane. Thus, TVaR_T is also continuous. According to (13) and (15), we have

$$\begin{aligned}\frac{\partial \mathrm{TVaR}_T(d,\alpha,\beta)}{\partial d} &= \frac{\partial \mathrm{TVaR}_T(d,\alpha,\beta)}{\partial d} + \frac{1}{\alpha}\frac{\partial \mathrm{ES}_T(d,\alpha,\beta)}{\partial d} \\ &= \begin{cases} (1+\rho)(S_X(d+\beta) - S_X(d)) < 0, & \text{i}, \\ 1 - \left[\frac{1}{\alpha} - (1+\rho)\right] S_X(d+\beta) - (1+\rho) S_X(d), & \text{ii}, \\ \left[\frac{1}{\alpha} - (1+\rho)\right](S_X(d) - S_X(d+\beta)), & \text{iii}, \end{cases}\end{aligned} \tag{30}$$

$$\frac{\partial \mathrm{TVaR}_T(d,\alpha,\beta)}{\partial \beta} = \begin{cases} -1 + (1+\rho) S_X(d+\beta), & \text{i}, \\ -\left[\frac{1}{\alpha} - (1+\rho)\right] S_X(d+\beta), & \text{ii}, \\ -\left[\frac{1}{\alpha} - (1+\rho)\right] S_X(d+\beta), & \text{iii}. \end{cases} \tag{31}$$

For a given β, TVaR is monotonically decreasing in district i, and, under the condition of (20), it is monotonically increasing in district iii. Hence, an optimal solution exists in district ii.

Lemma 4. *Under the condition of (20) and (1) $\forall \beta > 0$ ($\Theta = \{d \mid d > 0\}$), the optimal solution of TVaR appears in district ii; (2) in the case of d ($\Theta = \{\beta \mid \beta > 0\}$), no optimal value of TVaR exists; and (3) no global optimal value of TVaR ($\Theta = \{(d,\beta) \mid d > 0,\ \beta > 0\}$) exists.*

Proof. Figure 3 gives a graphical representation of $\min_{d>0} \mathrm{TVaR}_T(d,\alpha,\beta)$. At a given d,

$$\frac{\partial^2 \mathrm{TVaR}_T^{\mathrm{i}}(d,\alpha,\beta)}{\partial \beta^2} = -(1+\rho) f_X(d+\beta) < 0. \tag{32}$$

Therefore, in district i, the TVaR_T of β is convex. However, under (20), TVaR_T in districts ii and iii of β is decreasing, so no optimal solution exists (Figure 3). In this case, we cannot obtain an optimal solution by only changing β. And, moreover, no optimal solution exists when we change d and β. □

4.2. Under CTE Risk Measures. In this subsection, Λ = CTE. According to (13) and so on, we have

$$\begin{aligned}\mathrm{CTE}_T(d,\alpha,\beta) &= \mathrm{VaR}_T(d,\alpha,\beta) + \mathrm{CVaR}_T(d,\alpha,\beta) \\ &= \begin{cases} S_X^{-1}(\alpha) - \beta + (1+\rho)\int_d^{d+\beta} S_X(x)\,dx + \frac{1}{\alpha}\int_{S_X^{-1}(\alpha)}^{\infty} S_X(x)\,dx, & \text{i}, \\ d + (1+\rho)\int_d^{d+\beta} S_X(x)\,dx + \frac{1}{S_X(d+\beta)}\int_{d+\beta}^{\infty} S_X(x)\,dx, & \text{ii}, \\ S_X^{-1}(\alpha) + \frac{1}{\alpha}\int_{S_X^{-1}(\alpha)}^{\infty} S_X(x)\,dx - \left[\frac{1}{\alpha} - (1+\rho)\right]\int_d^{d+\beta} S_X(x)\,dx, & \text{iii}. \end{cases}\end{aligned} \tag{33}$$

Lemma 5. *CTE_T is continuous between districts i and ii and discontinuous between districts ii and iii. One therefore has:*

$$CTE_T^{ii}\left(S_X^{-1}(\alpha),\alpha,\beta\right) > CTE_T^{iii}\left(S_X^{-1}(\alpha),\alpha,\beta\right). \tag{34}$$

Proof. VaR_T and ES_T are continuous on the plane, so we only need to prove the numerator of CTE_T. If $\forall (d_0,\beta_0)$: $d_0 + \beta_0 = S_X^{-1}(\alpha)$, then $(d_0,\beta_0) \in$ ii; it is in the boundary of districts i and ii. We have $S_X(d_0+\beta_0) = \alpha$. Hence, CTE_T is continuous between districts i and ii.

If $\forall \beta_0 > 0$, $d_0 = S_X^{-1}(\alpha)$, then $(d_0,\beta_0) \in$ ii; it is on the boundary of districts ii and iii. We have $1/S_X(d_0+\beta_0) > 1/\alpha$. Thus, CTE_T is discontinuous between districts ii and iii. Furthermore, $\mathrm{CTE}_T^{\mathrm{ii}}(S_X^{-1}(\alpha),\alpha,\beta) > \mathrm{CTE}_T^{\mathrm{iii}}(S_X^{-1}(\alpha),\alpha,\beta)$. □

From (15) and (33), we have

$$\begin{aligned}\frac{\partial \mathrm{CTE}_T}{\partial d} &= \begin{cases} -(1+\rho)(S_X(d) - S_X(d+\beta)) < 0, & \text{i}, \\ -(1+\rho)(S_X(d) - S_X(d+\beta)) + \frac{f_X(d+\beta)}{S_X(d+\beta)}\mathrm{CVaR}_T^{\mathrm{ii}}, & \text{ii}, \\ \left[\frac{1}{\alpha} - (1+\rho)\right](S_X(d) - S_X(d+\beta)), & \text{iii}, \end{cases}\end{aligned} \tag{35}$$

$$\frac{\partial \mathrm{CTE}_T}{\partial \beta} = \begin{cases} -1 + (1+\rho) S_X(d+\beta), & \text{i}, \\ (1+\rho) S_X(d+\beta) + \frac{f_X(d+\beta)}{S_X(d+\beta)}\mathrm{CVaR}_T^{\mathrm{ii}} - 1, & \text{ii}, \\ -\left[\frac{1}{\alpha} - (1+\rho)\right] S_X(d+\beta), & \text{iii}. \end{cases} \tag{36}$$

In district ii, the partial derivatives in d and β depend on risk distribution f_X, which makes it difficult to determine the existence or location of the optimal solution. We therefore let

$$\begin{aligned} H(\beta) &= \mathrm{CTE}_T^{\mathrm{ii}}\big|_{d+\beta=S_X^{-1}(\alpha)} - \mathrm{CTE}_T^{\mathrm{iii}}\big|_{d\to S_X^{-1}(\alpha)} \\ &= \left[\frac{1}{\alpha} - (1+\rho)\right]\int_{S_X^{-1}(\alpha)}^{S_X^{-1}(\alpha)+\beta} S_X(x)\,dx \\ &\quad + (1+\rho)\int_{S_X^{-1}(\alpha)-\beta}^{S_X^{-1}(\alpha)} S_X(x)\,dx - \beta. \end{aligned} \tag{37}$$

We can obtain the solution d^{CTE} of $\partial\mathrm{CTE}_T^{\mathrm{ii}}/\partial d = 0$ with a numerical method (but it may not be in district ii). With (35), determining the position and uniqueness of d^{CTE} is difficult.

Lemma 6. *For a given β ($\Theta = \{d \mid d > 0\}$) under the condition of (20), CTE can obtain the optimal value in district ii if one of the following conditions is reached:*

(1) $H(\beta) \le 0$;

(2) $d^{CTE} \in ii$ *and* $CTE(d^{CTE}, \alpha, \beta) \le CTE_T^{iii}(S_X^{-1}(\alpha), \alpha, \beta)$.

Proof. $\partial\mathrm{CTE}_T^{\mathrm{i}}/\partial d < 0$, so no optimal solution exists in district i. From (35), we know no optimal solution exists under the condition of (20) because $\mathrm{CTE}_T^{\mathrm{iii}}$ is monotonically increasing in d. Thus, the optimal solution can only appear in district ii. For $H(\beta) \le 0$, the left border of district ii is not greater than the left boundary of district iii; hence, no optimal value exists in district ii. As (35) shows, $\partial\mathrm{CTE}_T^{\mathrm{ii}}/\partial d$ is continuous. Under the condition of statement (2) in Lemma 6, we can ensure the existence of the optimal solution. □

Lemma 7. *Under the condition of (20), CTE_T cannot attain the optimal value on given d ($\Theta = \{\beta \mid \beta > 0\}$).*

Proof. According to (20), $\partial\mathrm{CTE}_T^{\mathrm{iii}}/\partial\beta < 0$; no optimal value exists in the open set district iii. According to (36), we have

$$\frac{\partial^2\mathrm{CTE}_T^{\mathrm{i}}}{\partial\beta^2} = -(1+\rho) f_X(d+\beta) < 0, \tag{38}$$

where $\mathrm{CTE}_T^{\mathrm{i}}$ is strictly convex and the minimum value can only be obtained at the border. However, district i is an open set, so no optimal value exists. When $\beta = \infty$ and $0 < d < S_X^{-1}(\alpha)$, this is as described by Cai et al. [7]. Therefore,

$$\mathrm{CTE}_T^{\mathrm{ii}}(d, \alpha, \infty) = d + (1+\rho)\int_d^{\infty} S_X(x)\,dx. \tag{39}$$

Moreover, according to (33), we have

$$\begin{aligned} &\mathrm{CTE}_T^{\mathrm{ii}}(d, \alpha, \beta) \\ &\quad = \mathrm{CTE}_T^{\mathrm{ii}}(d, \alpha, \infty) \\ &\qquad + \left[\frac{1}{S_X(d+\beta)} - (1+\rho)\right]\int_{d+\beta}^{\infty} S_X(x)\,dx. \end{aligned} \tag{40}$$

Under (20), $\mathrm{CTE}_T^{\mathrm{ii}}(d, \alpha, \beta) > \mathrm{CTE}_T^{\mathrm{ii}}(d, \alpha, \infty)$, and no optimal solution exists. □

d is increasing within district iii, and the minimum can only appear when $d \to S_X^{-1}(\alpha)$. Thus, no optimal solution exists in the open set district iii. When $\beta = \infty$, Cai et al. [7] note that CTE_T is continuous. Under the conditions of (20), the minimum is obtained at $\tilde{d}$:

$$\mathrm{CTE}_T^{\mathrm{ii}}(\tilde{d}, \alpha, \infty) = \tilde{d} + (1+\rho)\int_{\tilde{d}}^{\infty} S_X(x)\,dx. \tag{41}$$

$\mathrm{CTE}_T^{\mathrm{i}}(0, \alpha, \beta)$ is convex on $\beta \in (0, S_X^{-1}(\alpha))$, so the minimum appears at both ends:

$$\begin{aligned} &\mathrm{CTE}_T^{\mathrm{i}}\left(0, \alpha, S_X^{-1}(\alpha)\right) \\ &\quad = (1+\rho)\int_0^{S_X^{-1}(\alpha)} S_X(x)\,dx + \frac{1}{\alpha}\int_{S_X^{-1}(\alpha)}^{\infty} S_X(x)\,dx \\ &\quad = \mathrm{CTE}_T^{\mathrm{ii}}(0, \alpha, \infty) \\ &\qquad + \left[\frac{1}{\alpha} - (1+\rho)\right]\int_{S_X^{-1}(\alpha)}^{\infty} S_X(x)\,dx. \end{aligned} \tag{42}$$

Therefore, under the condition of (20), $\mathrm{CTE}_T^{\mathrm{i}}(0, \alpha, S_X^{-1}(\alpha)) < \mathrm{CTE}_T^{\mathrm{ii}}(0, \alpha, \infty)$. We let

$$\begin{aligned} G(d) &= \mathrm{CTE}_T^{\mathrm{ii}}(d, \alpha, \infty) - \mathrm{CTE}_T^{\mathrm{i}}(d, \alpha, 0), \\ G'(d) &= 1 - (1+\rho) S_X(d), \\ G''(d) &= (1+\rho) f_x(d) > 0, \end{aligned} \tag{43}$$

where G is convex and $G'(\tilde{d}) = 0$. $G(S_X^{-1}(\alpha)) = 0$, so $G(\tilde{d}) < 0$. $\mathrm{CTE}_T(d, \alpha, 0)$ is a constant, so $\mathrm{CTE}_T^{\mathrm{ii}}(\tilde{d}, \alpha, \infty) < \mathrm{CTE}_T^{\mathrm{i}}(d, \alpha, 0)$. $\mathrm{CTE}_T^{\mathrm{ii}}(\tilde{d}, \alpha, \infty)$ is the minimum boundary value; if a smaller internal value exists in the flat areas, we can obtain the global optimal solution.

Lemma 8. *No global optimum CTE_T exists under the condition of (20), $\Theta = \{(d, \beta) \mid d > 0,\ \beta > 0\}$.*

Proof. Assume that the optimal solution of CTE_T exists. According to our previous analysis, the global optimal solution can only appear in district ii, and it cannot be the right boundary. We must therefore prove that it is impossible to obtain the optimal value in the lower boundary of district ii:

$$\begin{aligned} &\mathrm{CTE}_T^{\mathrm{ii}}\left(d, \alpha, S_X^{-1}(\alpha) - d\right) \\ &\quad = d + (1+\rho)\int_d^{S_X^{-1}(\alpha)} S_X(x)\,dx \\ &\qquad + \frac{1}{\alpha}\int_{S_X^{-1}(\alpha)}^{\infty} S_X(x)\,dx, \\ &\mathrm{CTE}_T^{\mathrm{ii}'}\left(d, \alpha, S_X^{-1}(\alpha) - d\right) = 1 - (1+\rho) S_X(d), \\ &\mathrm{CTE}_T^{\mathrm{ii}''}\left(d, \alpha, S_X^{-1}(\alpha) - d\right) = (1+\rho) f_X(d) > 0. \end{aligned} \tag{44}$$

In the lower boundary, we obtain the optimum at $\tilde{d}$. This is equivalent to having d given; according to Lemma 7, the above conclusions are reached.

Assuming that the optimal solution is in the interior of district ii, it must be a saddle point according to the extreme value theory of multivariate functions. Moreover,

$$\frac{\partial \mathrm{CTE}_T^{\mathrm{ii}}}{\partial \beta} = 0,$$
$$\frac{\partial \mathrm{CTE}_T^{\mathrm{ii}}}{\partial d} = \frac{\partial \mathrm{CTE}_T^{\mathrm{ii}}}{\partial \beta} + 1 - (1+\rho) S_X(d) = 0. \tag{45}$$

The solution to the equations is $d = \widetilde{d}$, assuming that $\beta = \beta'$. This is in the interior of district ii, so $\beta' > S_X^{-1}(\alpha) - \widetilde{d}$. CTE_T can only obtain the optimal value in $(\widetilde{d}, \beta')$. However,

$$\begin{aligned}\mathrm{CTE}_T^{\mathrm{ii}}&\left(\widetilde{d}, \alpha, \beta'\right)\\ &= \mathrm{CTE}_T^{\mathrm{ii}}\left(\widetilde{d}, \alpha, \infty\right)\\ &\quad + \left[\frac{1}{S_X\left(\widetilde{d}+\beta'\right)} - (1+\rho)\right]\int_{\widetilde{d}+\beta'}^{\infty} S_X(x)\,dx.\end{aligned} \tag{46}$$

In addition, $\widetilde{d} + \beta' > S_X^{-1}(\alpha)$, so $1/S_X(\widetilde{d} + \beta') > (1 + \rho)$; that is, $\mathrm{CTE}_T^{\mathrm{ii}}(\widetilde{d}, \alpha, \beta') > \mathrm{CTE}_T^{\mathrm{ii}}(\widetilde{d}, \alpha, \infty)$. No global optimum exists. □

5. Numerical Examples

5.1. VaR, TVaR, and CTE. We assume that the claim ratio of some insurance is gamma (4.1405, 0.1796) and $\alpha = 0.1$ and $\rho = 0.3$. Figures 5, 6, and 7 show the calculation results of VaR, TVaR, and CTE, respectively.

Note in Figure 5 that the contour shows a circumstance where the optimal trajectories of a given d and a given β cross inside the plane; the crossing point is the global optimal point. In this case, $\widetilde{d^*} < 0$; this appears when d is fixed on a line of $d + \beta = S_X^{-1}(\alpha)$.

Then, as shown in the contour map in Figure 6 and consistent with the theoretical analysis for a given d, TVaR decreases with an increase in β; in other words, no global optimal value exists. For a given β, a local optimum exists in district ii.

Figure 7 shows that a jump occurs when $d = S_X^{-1}(\alpha)$, forming a cross section. However, based on Lemma 6, the graphs do not show the optimal trajectories to a given β.

5.2. Difference between CTE and TVaR. Equations (8)–(13) show several differences in the denominators in district ii in the function of CTE and TVaR. Even small differences can lead to issues. However, no differences exist between CTE and TVaR when the variable of risk X is continuous; hence, many other studies use them interchangeably. Cai et al. [7] prove that CTE_T is continuous of total loss after reinsurance under stop-loss reinsurance. Thus, they can also be used similarly as substitutes for each other. We prove that CTE_T is discontinuous of total loss after reinsurance under limited stop-loss reinsurance. Obviously, therefore, they cannot be used interchangeably.

We assume that the claim ratio of some insurance is gamma (4, 0.125). We have $\rho = 0.2$, $\alpha = 0.01$, given reinsurance coverage $\beta = 1$. The results imply that TVaR is optimal when $d = 0.4317$ ($d < S_X^{-1}(\alpha) - \beta$, in i), and CTE is optimal when $d = S_X^{-1}(\alpha)$, as shown in Figure 8. $\beta = 1$ and the franchise point are large under CTE and can almost be approximated by stop-loss reinsurance.

For the risk distribution discussed above, note that (20) remains valid when $\rho = 0.5$, $\alpha = 0.2$, and $\beta = 0.6$. However, $H(\beta) > 0$, $\mathrm{CTE}(d, \alpha, \beta) > \mathrm{CTE}_T^{\mathrm{iii}}(S_X^{-1}(\alpha), \alpha, \beta)$, and $\forall d \in$ ii. Hence, we cannot obtain the optimal solution under CTE. From Lemma 4, we can obtain it under TVaR, $d = 0.4152$ ($\in$ ii), which is shown in Figure 8.

The model definition indicates that the CTE risk measures are more accurate than those of TVaR. However, the analysis and examples above show that discontinuous CTE can hinder optimal decision making to some extent.

5.3. Reinsurance Premium Budget Constraints. We mentioned the optimization problem regarding reinsurance premium budget constraints briefly during the introduction. Porth et al. [14] avoid solving the reinsurance premium equation (4), but they used the stochastic simulation method to solve the optimization problem. The previous proof shows that, in some cases, no possible optimal solution exists. Therefore, we first need to determine whether that is the case before applying stochastic simulation.

The CTE in Porth et al. [14] is consistent with TVaR used in this paper. Under financial constraints, whether a local optimum exists depends on the existence of solutions to the equations of d and β:

$$\begin{aligned}1 - \left[\frac{1}{\alpha} - (1+\rho)\right] S_X(d+\beta) - (1+\rho) S_X(d) &= 0,\\ (1+\rho)\int_d^{d+\beta} S_X(x)\,dx &= \pi.\end{aligned} \tag{47}$$

The first equation derives from (31); as per Lemma 4, we know that this is the local optimal trajectory. The second equation is the financial constraint equation (strict expression is $\delta(d, \beta) \le \pi$),

$$\begin{aligned}\frac{dd}{d\beta} &= -\frac{[1/\alpha - (1+\rho)]\, f_X(d+\beta)}{(1+\rho) f_X(d) + [1/\alpha - (1+\rho)]\, f_X(d+\beta)}\\ &\quad < 0,\\ \frac{dd}{d\beta} &= \frac{S_X(d+\beta)}{S_X(d) - S_X(d+\beta)} > 0.\end{aligned} \tag{48}$$

For the first equation, we have

$$\beta \longrightarrow \infty \Longrightarrow S_X(d+\beta) = 0 \Longrightarrow d_1 = S_X^{-1}\left(\frac{1}{1+\rho}\right). \tag{49}$$

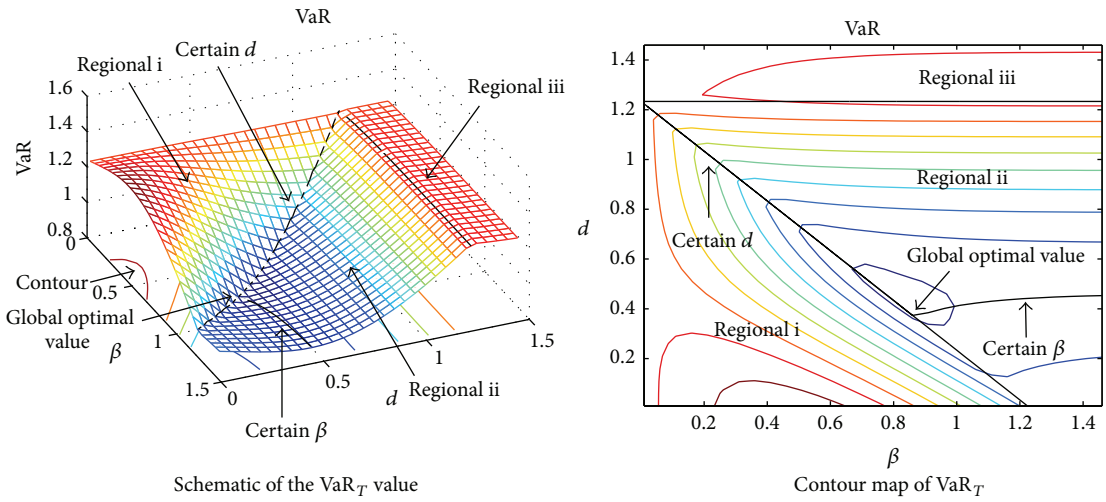


FIGURE 5: VaR_T that satisfies (20).

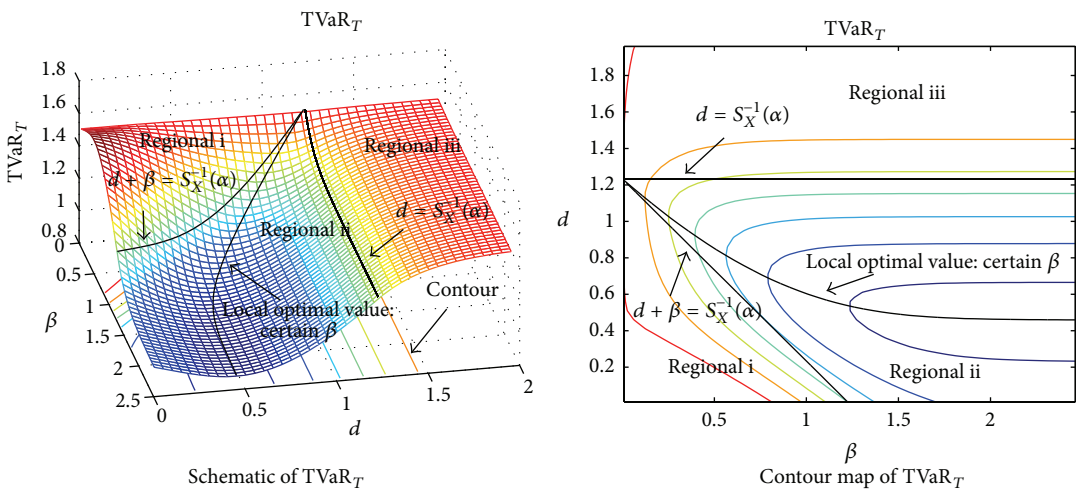


FIGURE 6: TVaR_T that satisfies (20).

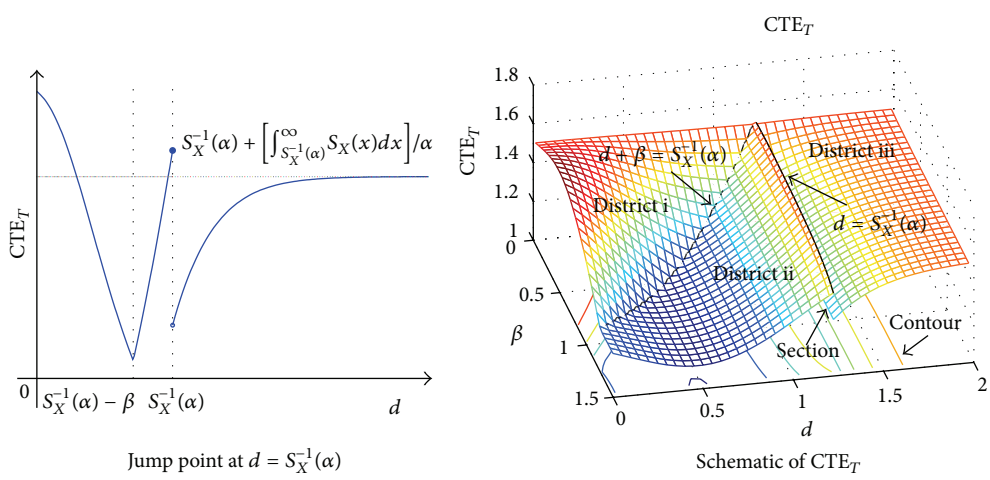


FIGURE 7: CTE_T that satisfies (20).

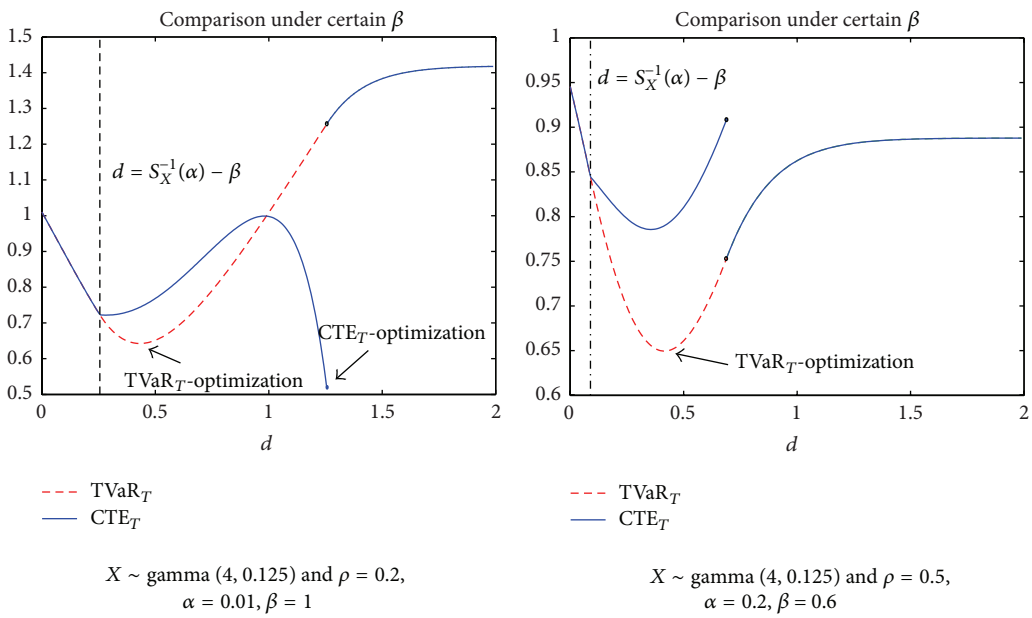


Figure 8: Optimal d for given β.

Under financial constraints, the existence of a local optimal solution depends on whether the following inequality is true:

$$\delta\left(S_X^{-1}\left(\frac{1}{1+\rho}\right), \infty\right) = (1+\rho)\int_{S_X^{-1}(1/(1+\rho))}^{\infty} S_X(x)\,dd \le \pi. \tag{50}$$

Similarly, we assume that the claim ratio of some insurance is gamma (4.1405, 0.1796) and $\alpha = 0.1$ and $\rho = 0.3$. However, financial budget $\pi = 0.1$ and 0.3, respectively. The results are shown in Figure 9. "$\pi = 0.1$" is the corresponding financial constraint line. When $\pi = 0.1$, the financial constraint line and the local optimal trajectories will effectively intersect (namely, the efficient solution of (47)). When $\pi = 0.3$, no intersection point will exist with local optimal trajectories. Therefore, using a numerical solution or stochastic simulation when $\pi = 0.3$ will lead to unpredictable results. Although $\pi = 0.3$ is therefore not feasible in general in businesses, we cannot rule out its existence in special businesses.

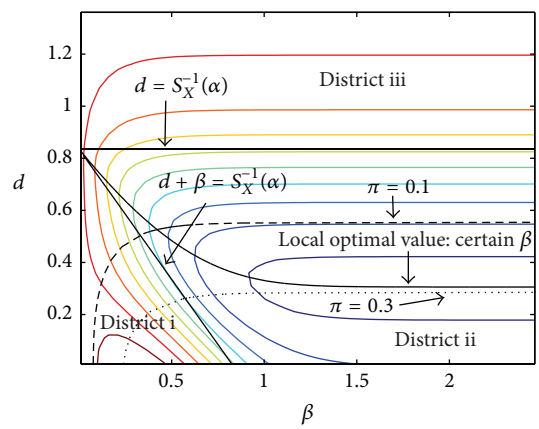


Figure 9: Under different financial constraints of TVaR_T.

6. Conclusion

We have mentioned the numerical solution of the equation several times throughout this paper. But it is ultimately unnecessary. For example, for a given β, β^* is the numerical solution of $h(\beta) = 0$ under VaR. To calculate it, we need only to determine whether (18) is positive or negative with β rather than with β^*. If we apply the optimization method directly to obtain an optimal solution, then the algorithm discussed here may create some confusion as follows. (1) Is the solution optimal? (2) If not, is that because of the algorithm or is it because no optimal solution exists? On the basis of the numerical method, we posit that the existence of the solution and considerable useful location information are theoretically provided.

Compared with stop-loss reinsurance ($\beta = \infty$), limited stop-loss coverage is more suitable for, for example, agricultural insurance. The latter's limited coverage β results in the discontinuity of risk probability functions and risk measures, such as the jump point between S_I and the CTE section. This discontinuous district determines the existence and location of an optimal solution.

Table 1 illustrates these results. All discussions are based on (20). Other conditions are also listed in Table 1: "No" means the solution only needs to satisfy (20); "—" means no optimal solution exists (no rigorous solution occurs; e.g., $\beta = \infty$). Note that VaR does not meet the consistency requirement. However, when we make an optimal reinsurance decision under VaR, it provides clear advice (i.e., global optimal reinsurance arrangements). VaR is limited in that

Table 1: Existence condition of optimal solution.

Risk measure	Certain β	Certain d	Comprehensive effect
VaR	$\forall\beta \in (0, \beta^*)$	$\forall d \in [\widetilde{d}^*, S_X^{-1}(\alpha)]$	No
TVaR	$\forall\beta \in (0, \beta^*)$	—	—
CTE	Lemma 6	—	—

the appetite for risk α cannot affect the optimal franchise point, which is determined solely by risk distribution and safety loading. Compared with VaR and TVaR, CTE is more reasonable. However, optimal reinsurance decisions can be difficult to be designed under CTE, and they can only be provided within a limited range. When β is given, optimal reinsurance purchase advice can be provided under TVaR and CTE. If risk preference α is known or if the maximum acceptable loss is d, then VaR should be the first choice. If the goal of reinsurance coverage is known, CTE would be recommended. In contrast, ES and CVaR are inappropriate choices for making reinsurance decisions.

In this study, note that safety loading ρ is a constant. An increase in the reinsurance compensation point will concurrently increase model uncertainty. Thus, we should set ρ as an increasing function of compensation point d. Moreover, we considered no financial constraints in this study. However, if we can effectively solve the reinsurance premium equation, then we can also effectively solve the problem of financial constraints and obtain the corresponding analytical solution. Restricted to research methods, we used several numerical methods to determine the existence of an optimal value. Fixed point theory has a strong advantage in dealing with optimization problems. Thus, introducing a fixed point method would avoid the numerical solutions used here.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This paper was funded by the Youth Project of the National Natural Science Foundation of China (71102125) and the MOE Project of the Key Research Institute of Humanities and Social Sciences at Universities (13JJD790041). The authors are grateful for the support provided by the Beijing Education Committee through the Young Talents Plan Project.

References

[1] K. Borch, "The safety loading of reinsurance premiums," *Scandinavian Actuarial Journal*, no. 3-4, pp. 163–184, 1960.

[2] H. U. Gerber, *An Introduction to Mathematical Risk Theory*, vol. 8, SS Huebner Foundation for Insurance Education, Wharton School, University of Pennsylvania, Philadelphia, Pa, USA, 1979.

[3] L. Gajek and D. Zagrodny, "Insurer's optimal reinsurance strategies," *Insurance: Mathematics & Economics*, vol. 27, no. 1, pp. 105–112, 2000.

[4] M. Kaluszka, "Mean-variance optimal reinsurance arrangements," *Scandinavian Actuarial Journal*, no. 1, pp. 28–41, 2004.

[5] Y. Bu, "On optimal reinsurance arrangement," in *Casualty Actuarial Society Forum*, pp. 1–20, 2005.

[6] J. Cai and K. S. Tan, "Optimal retention for a stop-loss reinsurance under the VaR and CTE risk measures," *Astin Bulletin*, vol. 37, no. 1, pp. 93–112, 2007.

[7] J. Cai, K. S. Tan, C. Weng, and Y. Zhang, "Optimal reinsurance under VaR and CTE risk measures," *Insurance: Mathematics & Economics*, vol. 43, no. 1, pp. 185–196, 2008.

[8] C. Weng, *Optimal reinsurance designs: from an insurer's perspective [Ph.D. thesis]*, University of Waterloo, 2009.

[9] L. Fu and C. Khury, "Optimal layers for catastrophe reinsurance," *Variance*, vol. 4, no. 2, pp. 191–208, 2010.

[10] Y. Chi and K. S. Tan, "Optimal reinsurance under VaR and CVaR Risk measures: a simplified approach," *Astin Bulletin*, vol. 41, no. 2, pp. 487–509, 2011.

[11] Y. Chi and C. Weng, "Optimal reinsurance subject to vajda condition," *Insurance: Mathematics and Economics*, vol. 53, no. 1, pp. 179–189, 2013.

[12] L. Gajek and D. Zagrodny, "Optimal reinsurance under general risk measures," *Insurance: Mathematics and Economics*, vol. 34, no. 2, pp. 227–240, 2004.

[13] M. Guerra and M. L. Centeno, "Optimal reinsurance policy: the adjustment coefficient and the expected utility criteria," *Insurance: Mathematics & Economics*, vol. 42, no. 2, pp. 529–539, 2008.

[14] L. Porth, K. S. Tan, and C. Weng, "Optimal reinsurance analysis from a crop insurer's perspective," *Agricultural Finance Review*, vol. 73, no. 2, pp. 310–328, 2013.

[15] Q. Li, M. Gu, and Z. Liang, "Optimal excess-of-loss reinsurance and investment polices under the CEV model," *Annals of Operations Research*, vol. 223, no. 1, pp. 273–290, 2014.

[16] M. Brandtner and W. Kürsten, "Solvency ii, regulatory capital, and optimal reinsurance: how good are conditional value-at-risk and spectral risk measures?" *Insurance: Mathematics and Economics*, vol. 59, pp. 156–167, 2014.

[17] C. Bernard and W. Tian, "Optimal reinsurance arrangements under tail risk measures," *Journal of Risk and Insurance*, vol. 76, no. 3, pp. 709–725, 2009.

[18] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.

[19] T. Hang, *Actuarial Aspects of Non-Life Insurance*, China Financial & Economic Publishing House, Beijing, China, 2010.

[20] A. Charpentier, *Mesures de risque*, Université Rennes 1, Rennes, France, 2010.

[21] J. L. Wirch and M. R. Hardy, "A synthesis of risk measures for capital adequacy," *Insurance: Mathematics and Economics*, vol. 25, no. 3, pp. 337–347, 1999.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 431608, 7 pages
http://dx.doi.org/10.1155/2015/431608

Research Article

Multilingual Text Detection with Nonlinear Neural Network

Lin Li,[1,2] Shengsheng Yu,[2] Luo Zhong,[1] and Xiaozhen Li[2]

[1] *School of Computer Science and Technology, Wuhan University of Technology, Wuhan 430070, China*
[2] *School of Computer Science, Huazhong University of Science and Technology, Wuhan 430070, China*

Correspondence should be addressed to Lin Li; lilin.wzy@gmail.com

Received 11 July 2015; Accepted 2 September 2015

Academic Editor: Xinguang Zhang

Multilingual text detection in natural scenes is still a challenging task in computer vision. In this paper, we apply an unsupervised learning algorithm to learn language-independent stroke feature and combine unsupervised stroke feature learning and automatically multilayer feature extraction to improve the representational power of text feature. We also develop a novel nonlinear network based on traditional Convolutional Neural Network that is able to detect multilingual text regions in the images. The proposed method is evaluated on standard benchmarks and multilingual dataset and demonstrates improvement over the previous work.

1. Introduction

Texts within images contain rich semantic information, which can be very beneficial for visual information retrieval and image understanding. Driven by a variety of real-world applications, scene text detection and recognition have become active research topics in computer vision. To efficiently read text from photography, the majority of methods follow the intuitive two-step process: text detection followed by text recognition [1]. To a large extent, the performance of text detection affects the accuracy of text recognition. Extracting textual information from natural scenes is a critical prerequisite for further text recognition and other image analysis tasks.

Text detection has been considered in many studies and considerable progress has been achieved in recent years [2–14]. However, most of the text detection methods have focused on English; few investigations have been done on the problem of the multilingual text detection. In our daily lives, multilingual texts coexist everywhere; many environments contain two or more scripts text in a single image and, for example, product tags, street signs, license plates, billboards, and guide information. More and more applications need to achieve text detection regardless of language type.

For different languages, the characters take many different forms and have inconsistent heights, strokes, and writing format. There are thousands of languages in the world, and the representative and universal features of multilingual text are still unknown. In addition, text embedded in images can be in variation of font style and size, different alignment and orientation, unknown colors, and varying lighting condition. Due to these factors, multilingual text detection in natural scenes is a challenging and important problem.

Our study is focused on learning the general stroke feature representations and detecting text from image even in a multiscript environment. Unlike traditional methods, which mainly relied on the combination of a number of hand-engineered features, we aim to test the feasibility of proposing a common text detector only using automatically learning text feature, by improving discriminative clustering algorithm, to obtain language-independent stroke features. The learned stroke features incorporating with nonlinear neural network provide an alternative way to effectively increase the character representational power. To use deep learning text feature, we are able to use simple nonmaximal suppression to locate text.

In the following, we first reviewed the recent published literature followed by the proposed multilingual text detection method from Section 3 to Section 4. In Section 5, the experimental evaluation is presented. The paper is concluded in Section 6.

2. Related Work

Existing methods proposed for text detection in natural scenes can be broadly categorized into two groups: connected component methods and sliding window methods.

Connected component methods separate text and non-text information at pixel-level, group text pixels to construct character candidates from images by connected component analysis. Epshtein et al. [2] leveraged the idea of the recovery of stroke width and proposed using the CCs in a stroke width transformed image. Yao et al. [3] extract regions in the Stroke Width Transform (SWT) domain. Neumann and Matas [4] posed the character detection problem as an efficient sequential selection from the set of Extremal Regions (ERs). Chen et al. [5] determined the stroke width using the distance transform of edge-enhanced Maximally Stable Extremal Regions (MSER). Using MSERs as CCs representing characters has become the focus of several recent works [6–9].

Sliding window-based methods, also known as region-based methods, scan a sliding subwindow through the image to search for possible texts and then use machine learning techniques to locate text. Wang et al. [10], extending their previous work [11], have built an end-to-end scene text recognition system based on a sliding window character classifier using Random Ferns. Wang et al. [12] use multi-layer neural networks for text detection. Jaderberg et al. [13] achieve state-of-the-art performance by implementing sliding window detection as a byproduct of the Convolutional Neural Network (CNN).

In the task of multilingual text detection, previous studies are mostly sliding window-based method. In [14, 16], authors have proposed similar methods using hand-engineered features to describe the text. Subwindow scanned on different scales and positions on the image pyramid in order to classify texts in images. Therefore, to achieve text detection which is invariance to language type, the feature representation is very important. However, little research attempted to apply deep learning to learn multilingual text feature. CNN is a special kind of neural network, and its deep nonlinear network structure shows the strong ability of learning discriminative features of datasets from observation samples. Therefore, we alternatively investigate the problem of multilingual text detection based on the framework of CNN.

3. Stroke Feature Learning

According to the study of linguistics, the basic feature of text is stroke, such as the Latin alphabet and Chinese basic stokes. And different languages share the same characteristics in appearance. Inspired by these ideas, it is possible that language-independent stroke features can be designed.

In order to cope with multilingual scenes, we seek to learn a bank of universal low-level stroke features directly from raw images. The learning stroke features should be able to capture the essential substructures of strokes. At the same time, they are of the most representative and discriminative stroke features. Many unsupervised learning algorithms can be used for learning the hidden data prototypes from dataset,

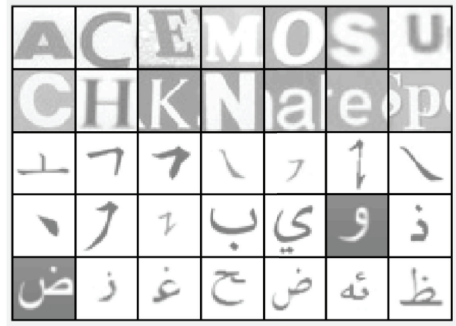


Figure 1: Training samples for stroke feature learning.

such as K-means clustering and sparse coding. The goal of sparse coding is to construct a dictionary D and minimize the error in reconstruction $\min_{D,s} \sum_i \|Ds^{(i)} - x^{(i)}\|_2^2 + \lambda \|s^{(i)}\|_1$, so that a data vector $x^{(i)} \in \mathbb{R}^n$ $(i = 1, \ldots, m)$ can be mapped to a code vector $s^{(i)}$. For every $s^{(i)}$, sparse coding algorithm is required to repeatedly solve a convex optimization problem. When applied to large scale image data, the optimization problem during the sparse coding procedure is very expensive. Relatively speaking, the optimal $s^{(i)}$ in classic K-means algorithm is simply as follows:

$$s_j^{(i)} = \begin{cases} D^{(j)\top} x^{(i)} & \text{if } j = \operatorname{argmax}_j \left| D^{(j)\top} x^{(i)} \right| \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

In addition, K-means has been identified as a fast and effective method to learn feature from images by computer vision researchers. Therefore, we improve the variant K-means clustering method proposed by Coates et al. [18] and use it to learn stroke feature representations, since it learns representative stroke features from large collections while much faster.

In particular, we first collect a set of training images, which are 32×32 gray scale images extracted from ICDAR 2003, ICDAR 2011, and ICDAR 2013 dataset, multilingual dataset, and Google. It contains a character in the middle of each image. Characters in training images include 26 uppercase letters, 26 lowercase letters, 10 digits, 20 Chinese basic strokes, and 28 Arabic alphabets. Some training images used for stroke feature learning are illustrated in Figure 1. We randomly extract m 8×8 pixel patches from images. Before training the cropped patches, we apply contrast normalized preprocessing for each patch. In order to avoid generating many highly correlated stroke features, ZAC whitening is used for the patches to yield vectors $x^{(i)} \in \mathbb{R}^{64}, i \in \{1, \ldots, m\}$.

Because K-means algorithm is highly dependent on the initialization process, the different initial guess of centroids affects the clustering result seriously. In order to lead to desirous clustering result, we propose a novel initialization method to choose suitable initial stroke features. We introduce the dispersion metric in the local information of data, guaranteeing the selection of initial centroids from the local spatial data-intensive region and the centroids apart from

Input: the patches x ($x^{(i)} \in \mathbb{R}^{64}, i = 1, \ldots, m$)
Output: initial features $F_0 \in \mathbb{R}^{64 \times n_1}$
(1) $C \leftarrow \emptyset$
(2) construct w_{ij} based on (2)
(3) computer dispersion metric $d^{(i)} = \sum_{j=1}^{m} w_{ij}$ and threshold $q = \text{median}(d)$
(4) **for** all data in x
(5) **if** data $x^{(j)}$ with dispersion metric $d^{(j)} > q$
(6) $C \leftarrow C \cup x^{(j)}$
(7) **end if**
(8) **end for**
(9) $F_0^{(1)} = x^{(i)}$, $x^{(i)}$ is random selected from C
(10) $F_0^{(2)} = \underset{x^{(j)}}{\text{argmax}} \left\| x^{(j)} - F_0^{(1)} \right\|, \quad \forall x^{(j)} \in C$
(11) set $k = 2$
(12) **repeat**
(13) $k = k + 1$
(14) $F_0^{(k)} = \underset{x^{(j)}}{\text{argmax}} \left\{ \min \left\| x^{(j)} - F_0^{(t)} \right\|, \ \forall x^{(j)} \in C, \ t = 1, \ldots, k-1 \right\}$
(15) **until** $k = n_1$

Algorithm 1: Stroke feature initialization method.

each other with a certain distance. Our initialization framework includes three steps: (1) estimating local dispersion metric for each set of data, (2) selecting the data which have higher metric than a threshold as candidates for initial features, and (3) determining initial stroke features from candidates. The implements of the proposed initialization method are as follows. We firstly construct an adjacency graph and Gram matrix; Gram matrix is computed according to the following:

$$w_{ij} = \begin{cases} 1 & \text{if } \left\| x^{(i)} - x^{(j)} \right\| < \varepsilon \\ 0 & \text{otherwise} \end{cases}$$
$$\varepsilon = \frac{\sum_{i=1}^{m} \sum_{j=1}^{m} \text{dis}\left(x^{(i)}, x^{(j)}\right)}{m \times (m-1)} \times 0.5, \tag{2}$$

where $\text{dis}(x^{(i)}, x^{(j)})$ is the distance of the patches $x^{(i)}$ and $x^{(j)}$. Secondly, we introduce the dispersion metric d with m components whose entries are given by $d^{(i)} = \sum_{j=1}^{m} w_{ij}$ ($i = 1, \ldots, m$). Set a threshold $q = \text{median}(d)$; if the value of $d^{(i)}$ associated with the data $x^{(i)}$ is larger than threshold; $x^{(i)}$ is marked as a candidate of initial features. Then, we use an algorithm similar to [19] to select initial stroke features from candidates. Because we use the stroke features as the first layer convolution kernels of our proposed CNN, n_1 is the number of first layer convolutional filters. The detailed steps of the proposed initialization method are presented in Algorithm 1.

After initializing F_0, we learn stroke features $F \in \mathbb{R}^{64 \times n_1}$ according to the following:

$$\min_{F, s^{(i)}} \sum_{i} \left\| F s^{(i)} - x^{(i)} \right\|^2. \tag{3}$$

For all $j \in \{1, \ldots, n_1\}$, compute inner products $F^{(j)\top} x^{(i)}$. Set the value of k which equals the value of j which maximizes the inner products. If $j = k$, then $s_j^{(i)} = F^{(j)\top} x^{(i)}$; or else $s_j^{(i)} = 0$. Then, fix $s^{(i)}$, minimizing (3) to obtain F. The optimization is done by alternating minimization over F and $s^{(i)}$. The full stroke feature learning algorithm with K-means is summarized in Algorithm 2.

For general clustering algorithm, the number of clusters is known in advance or set by prior knowledge. In our method, the learned stroke features incorporate with Convolutional Neural Network classifier for text detection. Therefore, we further study how to choose the appropriate number of features to achieve the highest text/no text classification accuracy. In order to analyze the impact of learned stroke feature number, we learned four stroke feature sets with different number. Given that the first layer convolution kernels of our CNNs have $n_1 = 96, 128, 256$ and 320, we train detector with different stroke feature sets. Evaluate the performance of the detection model at the subset of ICDAR 2003 test images. As shown in Figure 2, the F-measure increases as n_1 gets larger. Once n_1 equals 256, the recall is at maximum value, and about 80% of detected text matches ground truth. While n_1 is greater than 256, F-measure is not increased and even slightly reduced. Based on our detailed analysis, in our method, we select $n_1 = 256$.

4. Multilingual Text Detection

The idea of our text detection is to design "feature learning" pipeline that can lead to representative text features and use these features for detecting multilingual text. Two main components in this pipeline are as follows: (1) use the unsupervised clustering algorithm to generate a set of stroke features F; (2) build a hierarchy network and combine it with stroke features F to learn a high-level text feature. The first component has been described in detail in Section 3. How to

Input: m 8×8 input patches $x^{(i)} \in \mathbb{R}^{64}$
Output: learning stroke features $F \in \mathbb{R}^{64 \times n_1}$
Procedure:
(1) Normalize input

$$x^{(i)} = \frac{x^{(i)} - \text{mean}\left(x^{(i)}\right)}{\sqrt{10 + var\left(x^{(i)}\right)}}$$

(2) ZAC whiten input

$$VDV^{\top} = \text{cov}(x)$$
$$x^{(i)} = V(D + 0.1 \times I)^{-1/2} V^{\top} x^{(i)}$$

(3) Initialize F, follow the steps in Algorithm 1
(4) Repeat
Set $s_k^{(i)} = F^{(j)\top} x^{(i)}$ for $k = \text{argmax}_j |F^{(j)\top} x^{(i)}|$
Set $s_j^{(i)} = 0$ for all other $j \neq k$
Fix $s^{(i)}$, $\min_{F,s^{(i)}} \sum_i \|Fs^{(i)} - x^{(i)}\|^2$ s.t. $\|s^{(i)}\|_1 = \|s^{(i)}\|_\infty$ and $\|F^{(j)}\|_2 = 1$
Until convergence or reach iteration limitation

Algorithm 2: Stroke feature learning algorithm.

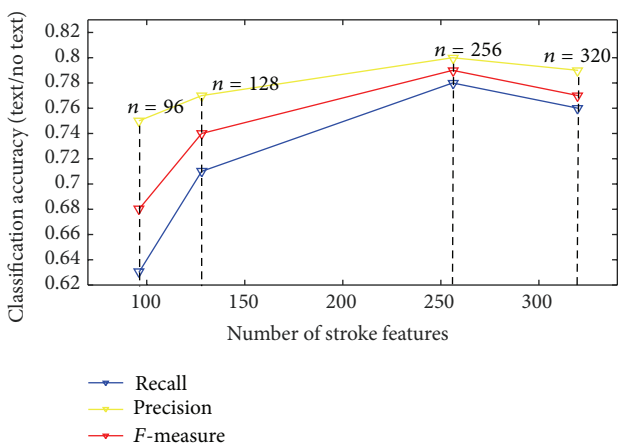


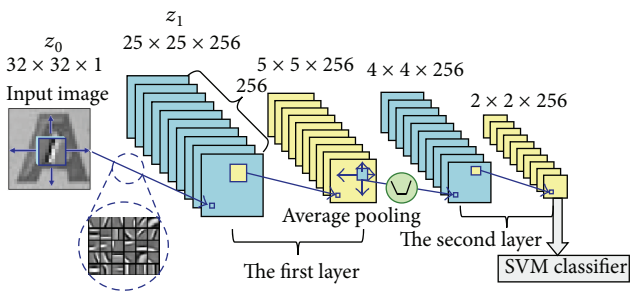
Figure 2: The accuracy analysis on different stroke feature number.

build and train the multilayer neural network is presented in Section 4.

By making several technical changes over traditional Convolutional Neural Network architectures [20], we develop a novel classifier for multilingual text detection. We have two major improvements: (1) different from traditional method that convolution kernel of CNN is randomly generated, we select the unsupervised learning stroke features F as the first layer convolution kernels of our network; (2) the intermediate features obtained from the learning process, which function in the second layers convolution kernels, can be used to more efficiently extract text features.

Our network has two convolutional layers with n_1 and n_2 filters, respectively. We fix the filters in the first convolution layer which are stroke features learned in Section 3; so low-level filters are $F \in \mathbb{R}^{64 \times n_1}$ and $n_1 = 256$. We build a set of labeled training datasets; all training images are 32×32 fixed-sized images (8877 positive, 9000 negative). Starting from the first layer, given an input image, the input is a grayscale cropped training image; that is, $z_0 = x$. The input convolves with 256 filters of size 8×8, resulting in a map of size 25×25 (to avoid boundary effects) and 256 channels. The first convolutional layer output z_1 is a new feature map computed by a nonlinear response function $z_1 = \max\{0, |F^{\top} x| - \alpha\}$, where $\alpha = 0.5$. Convolutional layers can be intertwined with pooling layers that simplify system parameters by statistical aggregation of features. We average pool over the first convolutional layer response map to reduce the size down to 5×5. The sequence continues by another convolutional and pooling layers, resulting in feature maps with 256 channels and size of 2×2; this size is the same as the dimension of the second layer convolutional filters. The second layer outputs are fully connected to the classification layer. The SVM classifier is used as a binary classifier that aims to estimate whether a 32×32 image contains text. We train the network using stochastic gradient descent and back-propagation. Classification error function includes loss term and regularization term. Loss term is a squared hinge loss and the norm used in the penalization is L2. We also use dropout in the second convolutional layer to help prevent over fitting. The structure of the proposed neural network is presented in Figure 3.

After our network has been trained, the detection process starts from a large, raw pixel input image and leverages the convolutional structure of the CNN to process the entire image. We slide a 32×32 pixels' window across an input image and put these sliding windows to the learned classifier. Use the intermediate hidden layers as features to classify text/no text and generate text bounding boxes. We set 12 different scales in our detection method. At a certain scale s, the input image's scale changes; the sliding window scans through the scaled image. At each point (x, y), if windows contain single centered characters, produce positive detector response $R_s[x, y]$. In each row r of the scaled image, check whether there are $R_s[x, y] > 0$. If there exists positive detector response, then form a line-level bounding box L_s^r with



F: first layer convolution kernels

Figure 3: The proposed network for multilingual text detection.

the same height as the sliding window at scale s. And $\max(x)$ and $\min(x)$ are defined as the left and right boundaries of L_s^r. At each scale, the input image is resizing and a set of candidate text bounding boxes are generated independently. The above procedure was repeated 12 times and yields groups of possibly overlapping bounding boxes. We then apply nonmaximal suppression (NMS) to score each box and remove all boxes that overlaps by more than 50% with lower score and obtain the final text bounding boxes L.

5. Experiments

5.1. Dataset. To evaluate the effectiveness and robustness of the proposed text detection algorithm, we have conducted experiments on standard benchmarks, including the challenging datasets ICDAR 2003 [21], MSRA-TD500 [3], and KAIST [17].

The ICDAR 2003 Robust Reading and Text Locating database is a widely used public dataset for scene text detection algorithm. The database contains 258 training images and 251 testing images. It contains the path to each image and text bounding box annotations for each image.

MSRA-TD500 dataset contains images with text in English and Chinese. The dataset contains 500 images in total, with varying resolutions from 1296 × 864 to 1920 × 1280. These images are taken from indoor (office and mall) and outdoor (street) scenes using a packet camera.

KAIST provides a scene text dataset consisting of 3000 images of indoor and outdoor scenes. Word and character bounding boxes are provided as well as segmentation maps of characters. Texts in KAIST images are English, Korean, and a mixture of English and Korean.

We also created a new multilingual dataset that is composed of three representative languages: English, Chinese, and Arabic. These three languages stand for three types of writing systems: English standing for alphabet, Chinese standing for ideograph, and Arabic standing for abjad. Each group corresponding to the one language contains 80 images.

To learn the stroke feature, train samples include 5980 English text samples, 800 Chinese text samples, and 1100 Arabic text samples. Then, 3000 nontext samples are extracted from 200 negative images using bootstrap method. All these samples are normalized to 32 × 32, which is consistent with the detected window.

Table 1: Performance of the proposed method.

Language	Total image	Precision	Recall	F-measure
English	800	0.76	0.88	0.8
Korean	500	0.73	0.84	0.78
Chinese	200	0.68	0.96	0.79
Arabic	200	0.70	0.72	0.66

Figure 4: Text detection samples on different language images.

5.2. Results. The proposed algorithm is implemented using Intel Core i5 processor at 2.9 GHz 8 GB RAM and MATLAB R2014b.

To validate the performance of our proposed algorithm, we use the definitions in ICDAR 2003 competition [21] for text detection precision, recall, and F-measure calculation. Therefore, $P = \sum_{r_e \in E} m(r_e, T)/|E|$ and $R = \sum_{r_t \in T} m(r_t, E)/|T|$, where $m(r, R)$ is the best match for a rectangle r in a set of rectangles R, E, and T which are our estimated rectangles and the ground truth rectangles, respectively. We adopt the F-measure to combine the precision and recall figures into a single measure of quality, $F = 1/(\alpha/P + \alpha/R)$, where $\alpha = 0.5$.

Experiments are carried out on a set of images containing text in four different languages, namely, English, Chinese, Arabic, and Korean. English text images are selected from ICDAR 2003, ICDAR 2011, and ICDAR 2013, Korean images from KAIST, some Chinese images from MSRA-TD500 and the other from multilingual dataset, and Arabic images from multilingual dataset and Google. The results of these evaluations are summarized in Table 1. As can be seen from Table 1, F-measures on different language are close to each other, except Arabic, because the Arabic special nature of continuous writing style makes the recall of this script lower. The experiment result indicates that our method is not tuned to any particular language and performs approximately equally good on all the scripts.

Figure 4 shows some texts successful detection by our system on images containing different language text. Although the texts contained in training samples are only in English, Chinese, and Arabic, our method can detect the text not only in three representative languages, but also in a number of other languages, such as French, German, Korean, and Japanese. This shows that our method has some robustness.

Table 2: Performance comparison on different benchmarks.

Method	Dataset	Precision	Recall	F-measure
Pan [15]	ICDAR 2003	0.645	0.659	0.652
Zhou [16]	ICDAR 2003	0.37	0.79	0.53
Our method	ICDAR 2003	0.45	**0.80**	0.57
Lee [17]	KAIST	0.69	0.60	0.64
Our method	KAIST	0.59	**0.79**	**0.67**

Figure 5: Text detection samples on images containing two different languages text.

We also picked methods proposed by Zhou et al. [16], Pan et al. [15], and Lee et al. [17] for further consideration. These algorithms have good results on the standard benchmarks and use different approaches to detect text. The performance comparison analysis can be seen in Table 2. Our method has achieved high recall at different benchmarks. It also reflects the representative of our learned feature is strong, which can successfully detect all the information associated with the text in images.

But our test results are not good, with $P/R/F$-means of 0.30/0.32/0.31 on the MSRA-TD500 dataset. Shi et al. [6] have achieved the state-of-the-art text detection performance with 0.52/0.53/0.5 on the same dataset. The main reason is that the MSRA-TD500 dataset is created for the purpose of study of multiorientation text detection, which has a lot of images containing no horizontal text lines. But our method gives the text bounding boxes based on the horizontal direction.

Figure 5 shows some other test samples. The results reflect our method is efficient on the circumstance that a single image contains two or more different languages texts and numbers. The bottom row in Figure 5 shows some fail samples; some of these problems are miss detection for part of Arabic text, because Arabic words mostly are linked by continuous line. In this case, use of the stroke feature to detect text is not sufficient. Stroke width of the implementation is essential for such languages as Arabic. There are other problems caused by the interference terms which have the appearance similar to the text.

6. Conclusion

The aim of the study is to propose a multilingual text detection method. Traditional methods in this area mainly rely on large amounts of hand-engineered features or prior knowledge. Our work is distinct in two ways: (1) we use primitive stroke feature learned by unsupervised learning algorithm as network convolutional kernels; (2) we leverage the trained multilayer neural network to learn high-level abstract text features used for detector. Experiments on the public benchmark and multilingual dataset show our method is able to localize text regions of different scripts in natural scene images. The experiment results demonstrate the robustness of the proposed method.

From the failed samples in the experiments, we analyze the limitations of our technology for further improvement. On the one hand, some languages have continuous writing style, like Arabic; automatically learning features are not enough for detection; the connected components analysis will be added into our method to improve the precision of final results. On the other hand, multiorientation text problem will be considered.

Conflict of Interests

The authors declared that they have no conflict of interests regarding this work.

References

[1] X. Chen and A. L. Yuille, "Detecting and reading text in natural scenes," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04)*, vol. 2, pp. II366–II373, July 2004.

[2] B. Epshtein, E. Ofek, and Y. Wexler, "Detecting text in natural scenes with stroke width transform," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 2963–2970, June 2010.

[3] C. Yao, X. Bai, W. Liu, Y. Ma, and Z. Tu, "Detecting texts of arbitrary orientations in natural images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 1083–1090, June 2012.

[4] L. Neumann and J. Matas, "Real-time scene text localization and recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 3538–3545, Providence, RI, USA, June 2012.

[5] H. Chen, S. S. Tsai, G. Schroth, D. M. Chen, R. Grzeszczuk, and B. Girod, "Robust text detection in natural images with edge-enhanced maximally stable extremal regions," in *Proceedings of the 18th IEEE International Conference on Image Processing (ICIP '11)*, pp. 2609–2612, September 2011.

[6] C. Shi, C. Wang, B. Xiao, Y. Zhang, and S. Gao, "Scene text detection using graph model built upon maximally stable extremal regions," *Pattern Recognition Letters*, vol. 34, no. 2, pp. 107–116, 2013.

[7] A. Shahab, F. Shafait, and A. Dengel, "ICDAR 2011 robust reading competition challenge 2: reading text in scene images," in *Proceedings of the 11th International Conference on Document Analysis and Recognition (ICDAR '11)*, pp. 1491–1496, Beijing, China, September 2011.

[8] X.-C. Yin, X. Yin, K. Huang, and H.-W. Hao, "Robust text detection in natural scene images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 5, pp. 970–983, 2014.

[9] L. Neumann and J. Matas, "Scene text localization and recognition with oriented stroke detection," in *Proceedings of the 14th IEEE International Conference on Computer Vision (ICCV '13)*, pp. 97–104, December 2013.

[10] K. Wang, B. Babenko, and S. Belongie, "End-to-end scene text recognition," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV '11)*, pp. 1457–1464, IEEE, Barcelona, Spain, November 2011.

[11] K. Wang and S. Belongie, "Word spotting in the wild," in *Proceedings of the 11th European Conference on Computer Vision (ECCV '10)*, pp. 591–604, 2010.

[12] T. Wang, D. J. Wu, A. Coates, and A. Y. Ng, "End-to-end text recognition with convolutional neural networks," in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR '12)*, pp. 3304–3308, November 2012.

[13] M. Jaderberg, A. Vedaldi, and A. Zisserman, "Deep features for text spotting," in *Proceedings of the 22nd International Conference on Pattern Recognition (ICPR '14)*, 2014.

[14] A. Raza, I. Siddiqi, C. Djeddi, and A. Ennaji, "Multilingual artificial text detection using a cascade of transforms," in *Proceedings of the 12th International Conference on Document Analysis and Recognition (ICDAR '13)*, pp. 309–313, Washington, DC, USA, August 2013.

[15] Y.-F. Pan, X. Hou, and C.-L. Liu, "A hybrid approach to detect and localize texts in natural scene images," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 800–813, 2011.

[16] G. Zhou, Y. Liu, Q. Meng, and Y. Zhang, "Detecting multilingual text in natural scene," in *Proceedings of the 1st International Symposium on Access Spaces (ISAS '11)*, pp. 116–120, IEEE, Yokohama, Japan, June 2011.

[17] S. Lee, M. S. Cho, K. Jung, and J. H. Kim, "Scene text extraction with edge constraint and text collinearity," in *Proceedings of the 20th International Conference on Pattern Recognition (ICPR '10)*, pp. 3983–3986, August 2010.

[18] A. Coates, B. Carpenter, C. Case et al., "Text detection and character recognition in scene images with unsupervised feature learning," in *Proceedings of the 11th International Conference on Document Analysis and Recognition (ICDAR '11)*, pp. 440–445, September 2011.

[19] M. E. Celebi and H. A. Kingravi, "Deterministic initialization of the k-means algorithm using hierarchical clustering," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 26, no. 7, Article ID 1250018, 2012.

[20] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[21] S. Lucas, P. Panaretos, L. Sosa, A. Tang, S. Wong, and R. Young, "ICDAR 2003 robust reading competitions," in *Proceedings of the 7th International Conference on Document Analysis and Recognition (ICDAR '03)*, pp. 682–687, Edinburgh, UK, August 2003.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 134246, 8 pages
http://dx.doi.org/10.1155/2015/134246

Research Article

A Dependent Insurance Risk Model with Surrender and Investment under the Thinning Process

Wenguang Yu[1] and Yujuan Huang[2]

[1] *School of Insurance, Shandong University of Finance and Economics, Jinan 250014, China*
[2] *School of Science, Shandong Jiaotong University, Jinan 250023, China*

Correspondence should be addressed to Yujuan Huang; yujuanh518@163.com

Received 26 August 2015; Accepted 17 September 2015

Academic Editor: Xinguang Zhang

A dependent insurance risk model with surrender and investment under the thinning process is discussed, where the arrival of the policies follows a compound Poisson-Geometric process, and the occurrences of the claim and surrender happen as the p-thinning process and the q-thinning process of the arrival process, respectively. By the martingale theory, the properties of the surplus process, adjustment coefficient equation, the upper bound of ruin probability, and explicit expression of ruin probability are obtained. Moreover, we also get the Laplace transformation, the expectation, and the variance of the time when the surplus reaches a given level for the first time. Finally, various trends of the upper bound of ruin probability and the expectation and the variance of the time when the surplus reaches a given level for the first time are simulated analytically along with changing the investment size, investment interest rates, claim rate, and surrender rate.

1. Introduction

In the classical ruin theory, compound Poisson risk model,

$$U(t) = u + ct - \sum_{i=1}^{N(t)} X_i \quad \text{for } t \geq 0, \tag{1}$$

is the main research object [1, 2], where $u \geq 0$ is the initial reserve, c is the premium rate, and $\{N(t),\ t \geq 0\}$ is a Poisson process with intensity $\lambda > 0$, representing the number of claims up to time t. The individual claim sizes $X_1, X_2, \ldots$, independent of $\{N(t),\ t \geq 0\}$, are i.i.d. positive random variables with distribution function $F(x)$ and density function $f(x)$ with mean μ. In the model, the premium income process is a linear function of time; it does not matter to claim. But in actual life, the arrival of policy of insurance company is usually associated with occurrence of claim; for example, the more the number of policies sold, the more the number of claims happened. Therefore, many studies in literature discuss the dependent relationship among the premium income, interclaim arrivals, and the claim size. See, for example, Liu et al. [3] considering a Markov-dependent risk model with a constant dividend barrier. Shi et al. [4] explore methods that allow for the correlation among frequency and severity components for microlevel insurance data. Jiang et al. [5] investigate some uniform asymptotic estimates for finite-time ruin probabilities when the claim size vector and its interarrival time are subject to certain general dependence structure. Zhang and Yang [6], Shi et al. [7], and Zou et al. [8] consider a compound Poisson risk model and a dependence structure of the claim size and interclaim time modeled by a Farlie-Gumbel-Morgenstern copula.

The above papers always assume the claim number follows a Poisson distribution, but in fact the claim number does not fully comply with the rule of Poisson distribution and its variance is often greater than the mean. Except the natural environment, an important reason for this phenomenon is that insurance companies have adopted risk aversion mechanism, such as franchise system and no-claim discount system [9]. This makes the policy holder weighs the interests which may not claim for compensation in the event of an accident; it will cause the claim number to be less than the number of accidents. In addition, on the one hand, the insurance company will have huge funds and various kinds of reserves in the operation process, which formed the huge amount of available funds. On the other hand, in order to

protect the interests of the insured, the insurance company must use the fund rationally and effectively. In fact, the insurance industry is very active in the financial markets. In the financial markets of western developed countries, the total amount of funds provided by the insurance industry is close to commercial banks. So considering the risk model with investment income has greater practical value and realistic significance [10–12].

In view of the above problems, this paper will promote the premium income process of insurance companies to follow the compound Poisson-Geometric process [13–15], while the counting processes of claim and surrender are the p-thinning process and the q-thinning process of premium income process and further consideration of the investment interest rate. For the new improved model, we study the properties of surplus process, adjustment coefficient equation, ruin probability, and the expectation and variance of the first time to reach a given level. Finally, numerical analysis is also given.

The contents of this paper are organized as follows: Section 2 introduces the risk model. In Section 3, we give the main results of the paper. Finally, we provide the numerical examples in Section 4.

2. The Risk Model

Definition 1. Let $u \geq 0$ and (Ω, F, P) be a probability space; $t \geq 0$; then, the surplus process with initial surplus u is defined as follows:

$$U(t) = u + A\alpha t + \sum_{k=1}^{N(t)} X_k - \sum_{k=1}^{N(t,p)} Y_k - \sum_{k=1}^{N(t,q)} Z_k + \beta W(t), \tag{2}$$

where u represents the initial capital and A $(A < u)$ represents the investment capital, which is based on the size of initial capital, premium income per unit of time, and the predicted claim sizes. α represents the investment income per unit of time. $\{N(t);\ t \geq 0\}$ is a Poisson-Geometric process with parameters λ $(\lambda > 0)$ and ρ $(0 < \rho < 1)$ denoting the number of premiums up to time t; namely, $N(t) \sim PG(\lambda t, \rho)$. $\{X_k \geq 0;\ k \geq 1\}$ is a sequence of i.i.d. random variables representing the amount of the kth premium and $E[X_k] = \mu_X$ $\mathrm{Var}[X_k] = \sigma_X^2$. $\{N(t,p);\ t \geq 0\}$ is the p-thinning process of $\{N(t);\ t \geq 0\}$ denoting the number of claims up to time t; namely, $\{N(t,p);\ t \geq 0\} \sim PG(\lambda pt, \rho)$. The individual claims sizes $\{Y_k \geq 0;\ k \geq 1\}$ are a sequence of i.i.d. random variables and $E[Y_k] = \mu_Y$ $\mathrm{Var}[Y_k] = \sigma_Y^2$. $\{N(t,q);\ t \geq 0\}$ is the q-thinning process of $\{N(t);\ t \geq 0\}$ denoting the number of surrenders up to time t; namely, $\{N(t,q);\ t \geq 0\} \sim PG(\lambda qt, \rho)$ and $0 < q < 1$ and $0 < p+q < 1$. The sequence of i.i.d. random variables $\{Z_k \geq 0;\ k \geq 1\}$ represents the amount of the kth payment of insurance policy and $E[Z_k] = \mu_Z$, $\mathrm{Var}[Z_k] = \sigma_Z^2$, and $\mu_Z < \mu_X$. $\{W(t);\ t \geq 0\}$ is a standard Brownian motion denoting the uncertain benefits and payments of insurance companies. $\beta > 0$ is a constant, representing the diffusion volatility parameter. In addition, we suppose that $\{X_k,\ k \geq 1\}$, $\{Y_k,\ k \geq 1\}$, $\{Z_k,\ k \geq 1\}$, $\{W(t),\ t \geq 0\}$, and $\{N(t);\ t \geq 0\}$ are mutually independent. From the theory of point process, $\{N(t,p);\ t \geq 0\}$ and $\{N(t,q);\ t \geq 0\}$ are also mutually independent.

Let $S(t) = A\alpha t + \sum_{k=1}^{N(t)} X_k - \sum_{k=1}^{N(t,p)} Y_k - \sum_{k=1}^{N(t,q)} Z_k + \beta W(t)$ be profits process. In order to ensure the insurance company's steady business, we assume $E[S(t)] > 0$, and the relative security loading factor θ is defined as follows:

$$\theta = \frac{\lambda\mu_X + A\alpha(1-\rho)}{\lambda p\mu_Y + \lambda q\mu_Z} - 1 > 0. \tag{3}$$

3. Main Results

Lemma 2. *The profits process $\{S(t),\ t \geq 0\}$ has the following properties:*

(i) *$\{S(t),\ t \geq 0\}$ has stationary and independent increments.*

(ii) *$E[S(t)] = (A\alpha + \lambda\mu_X/(1-\rho) - \lambda p\mu_Y/(1-\rho) - \lambda q\mu_Z/(1-\rho))t$.*

Lemma 3. *For the profits process $\{S(t),\ t \geq 0\}$, when $E[S(t)] \geq 0$, one has the following:*

$$\lim_{t\to\infty} U(t) = \infty, \quad a.s. \tag{4}$$

Lemma 4. *For the profits process $\{S(t),\ t \geq 0\}$, suppose $E[e^{-rS(t)}] < 0$ for some $r > 0$; then, there is a function $g(r)$ such that*

$$E\left[e^{-rS(t)}\right] = e^{tg(r)}. \tag{5}$$

Proof. Consider

$$\begin{aligned} E[\exp(-rS(t))] &= E[\exp(-rA\alpha t)] \\ &\cdot E\left[\exp\left(-r\sum_{K=1}^{N(t)} X_k\right)\right] \cdot E\left[\exp\left(r\sum_{K=1}^{N(t,p)} Y_k\right)\right] \\ &\cdot E\left[\exp\left(r\sum_{K=1}^{N(t,q)} Z_k\right)\right] \cdot E[\exp(-r\beta W(t))] \\ &= \exp\left\{t\left[-rA\alpha + \frac{\lambda(M_X(-r)-1)}{1-\rho M_X(-r)}\right.\right. \\ &\left.\left. + \frac{\lambda p(M_Y(r)-1)}{1-\rho M_Y(r)} + \frac{\lambda q(M_Z(r)-1)}{1-\rho M_Z(r)} + \frac{1}{2}\beta^2 r^2\right]\right\}. \end{aligned} \tag{6}$$

Let

$$\begin{aligned} g(r) &= -rA\alpha + \frac{\lambda(M_X(-r)-1)}{1-\rho M_X(-r)} + \frac{\lambda p(M_Y(r)-1)}{1-\rho M_Y(r)} \\ &+ \frac{\lambda q(M_Z(r)-1)}{1-\rho M_Z(r)} + \frac{1}{2}\beta^2 r^2, \end{aligned} \tag{7}$$

where $M_X(r) = E[e^{rX}]$ is the moment generating function of X. Similarly, we can define $M_Y(r)$ and $M_Z(r)$. □

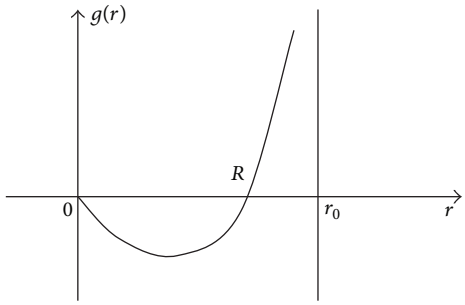


Figure 1: Adjustment coefficient R.

The following discussions are adjustment coefficient and the adjustment coefficient equation. Since the ruin probability as a number of indicators can evaluate insurance company solvency, it attracts attention. The research goal is to obtain specific expression of ruin probability. However, it is very difficult to directly obtain the expression of this function, but Lundberg found an indirect expression way by introducing a parameter which can play the intermediary role, namely, Lundberg coefficient or adjustment coefficient. Its principle is that the ruin probability is expressed as a function of adjustment coefficient and then seeks the calculation for adjustment coefficient. Thus, the adjustment coefficient plays a very important role in the study of ruin probability.

Lemma 5. *Equation $g(r) = 0$ is said to be an adjustment coefficient equation of the risk model (2), and it has a unique positive solution $r = R$, which is called an adjustment coefficient (see Figure 1).*

Proof. We only need to prove that it has the following four properties:

(1) $g(0) = 0$.

(2) $g'(0) < 0$.

(3) $g(r) \to +\infty$ $(r \to +\infty)$.

(4) $g''(r) > 0, \forall r \in (0, +\infty)$.

Obviously, $g(0) = 0$.

Since

$$\begin{aligned} g'(r) = -A\alpha &+ \frac{\lambda(1-\rho)E[-X\exp(-rX)]}{(1-\rho M_X(-r))^2} \\ &+ \frac{\lambda p(1-\rho)E[Y\exp(rY)]}{(1-\rho M_Y(r))^2} \\ &+ \frac{\lambda q(1-\rho)E[Z\exp(rZ)]}{(1-\rho M_Z(r))^2} + \beta^2 r \end{aligned} \tag{8}$$

and $M_X(0) = M_Y(0) = M_Z(0) = 1$, then we have

$$\begin{aligned} g'(0) &= -A\alpha - \frac{\lambda\mu_X}{1-\rho} + \frac{\lambda p\mu_Y}{1-\rho} + \frac{\lambda q\mu_Z}{1-\rho} \\ &= -\theta\left(\frac{\lambda p\mu_Y}{1-\rho} + \frac{\lambda q\mu_Z}{1-\rho}\right) < 0. \end{aligned} \tag{9}$$

Further,

$$\begin{aligned} g''(r) = &\frac{\lambda(1-\rho)(1-\rho M_X(-r))}{(1-\rho M_X(-r))^4} \\ &\cdot\left\{(1-\rho M_X(-r))E\left[X^2e^{-rX}\right]\right. \\ &\left.+ 2\rho\left(E\left[-Xe^{-rX}\right]\right)^2\right\} + \frac{\lambda p(1-\rho)(1-\rho M_Y(r))}{(1-\rho M_Y(r))^4} \\ &\cdot\left\{(1-\rho M_Y(r))E\left[Y^2e^{-rY}\right] + 2\rho\left(E\left[Ye^{-rY}\right]\right)^2\right\} \\ &+ \frac{\lambda q(1-\rho)(1-\rho M_Z(r))}{(1-\rho M_Z(r))^4} \\ &\cdot\left\{(1-\rho M_Z(r))E\left[Z^2e^{-rZ}\right] + 2\rho\left(E\left[Ze^{-rZ}\right]\right)^2\right\} \\ &+ \beta^2. \end{aligned} \tag{10}$$

It is easy to see the moment generating functions $M_Y(r)$, $M_Z(r)$ are increasing function, so there exists an $r_1 > 0$ such that $M_Y(r_1) = 1/\rho$ due to $0 < \rho < 1$. Similarly, there exists an $r_2 > 0$ such that $M_Y(r_2) = 1/\rho$. Let $r_0 = \min\{r_1, r_2\}$; then, when $0 < r < r_0$, we have $1 < M_Y(r), M_Z(r), < 1/\rho$; that is, $1 - \rho M_Y(r) > 0$ and $1 - \rho M_Z(r) > 0$. When $r > 0$, we have $M_X(-r) < 1$; that is, $1 - \rho M_X(-r) > 0$. So $g''(r) > 0$ $(0 < r < r_0)$ and $g(r)$ is a lower convex function. And because $\lim_{r\to+\infty} g(r) = +\infty$, then $g(r) = 0$ has a unique positive solution R. □

Theorem 6. *The adjustment coefficient R satisfies the following inequality:*

$$\begin{aligned} &\frac{2[A\alpha + \lambda\mu_X]}{\lambda(\mu_X^2 + \sigma_X^2) + \beta^2} \le R \\ &\le \frac{2[(1-\rho)A\alpha + \lambda\mu_X - \lambda p\mu_Y - \lambda q\mu_Z]}{\lambda[(\mu_X^2 + \sigma_X^2) + p(\mu_Y^2 + \sigma_Y^2) + q(\mu_Z^2 + \sigma_Z^2) + \beta^2(1-\rho)]}. \end{aligned} \tag{11}$$

Proof. By Taylor's expansion, we have

$$\begin{aligned} M_X(-R) &= E\left[e^{-RX}\right] = E\left[1 - RX + \frac{R^2X^2}{2}\right] \\ &= 1 - R\mu_X + \frac{R^2}{2}\left(\mu_X^2 + \sigma_X^2\right) \\ M_Y(R) &= E\left[e^{RY}\right] = E\left[1 + RY + \frac{R^2Y^2}{2}\right] \\ &= 1 + R\mu_Y + \frac{R^2}{2}\left(\mu_Y^2 + \sigma_Y^2\right). \end{aligned} \tag{12}$$

Then,

$$\begin{aligned}0 &= g(R)\\ &= -RA\alpha + \frac{\lambda\left(M_X(-R)-1\right)}{1-\rho M_X(-R)} + \frac{\lambda p\left(M_Y(R)-1\right)}{1-\rho M_Y(R)}\\ &\quad + \frac{\lambda q\left(M_Z(R)-1\right)}{1-\rho M_Z(R)} + \frac{1}{2}\beta^2R^2\\ &\geq -RA\alpha + \frac{\lambda\left[-R\mu_X + \left(R^2/2\right)\left(\mu_X^2+\sigma_X^2\right)\right]}{1-\rho}\\ &\quad + \frac{\lambda p\left[R\mu_Y + \left(R^2/2\right)\left(\mu_Y^2+\sigma_Y^2\right)\right]}{1-\rho}\\ &\quad + \frac{\lambda q\left[R\mu_Z + \left(R^2/2\right)\left(\mu_Z^2+\sigma_Z^2\right)\right]}{1-\rho} + \frac{1}{2}\beta^2R^2.\end{aligned} \tag{13}$$

Dividing both sides of the above inequality by R, we obtain

$$R \leq \frac{2\left[(1-\rho)A\alpha + \lambda\mu_X - \lambda p\mu_Y - \lambda q\mu_Z\right]}{\lambda\left[\left(\mu_X^2+\sigma_X^2\right) + p\left(\mu_Y^2+\sigma_Y^2\right) + q\left(\mu_Z^2+\sigma_Z^2\right) + \beta^2(1-\rho)\right]}. \tag{14}$$

Similarly,

$$\begin{aligned}0 &= g(R)\\ &\leq -RA\alpha + \left[\lambda\left(M_X(-R)-1\right)\right] + \frac{\lambda p\left(M_Y(R)-1\right)}{1-M_Y(R)}\\ &\quad + \frac{\lambda q\left(M_Z(R)-1\right)}{1-M_Z(R)} + \frac{1}{2}\beta^2R^2\\ &\leq -RA\alpha + \lambda\left[-R\mu_X + \frac{R^2}{2}\left(\mu_X^2+\sigma_X^2\right)\right] - \lambda p - \lambda q\\ &\quad + \frac{1}{2}\beta^2R^2\\ &\leq -RA\alpha + \lambda\left[-R\mu_X + \frac{R^2}{2}\left(\mu_X^2+\sigma_X^2\right)\right] + \frac{1}{2}\beta^2R^2.\end{aligned} \tag{15}$$

Dividing both sides of the above inequality by R, we have

$$R \geq \frac{2\left[A\alpha + \lambda\mu_X\right]}{\lambda\left(\mu_X^2+\sigma_X^2\right) + \beta^2}. \tag{16}$$

For the profits process $\{S(t),\ t \geq 0\}$, let $F_t^S = \sigma(S(v),\ v \leq t)$ be a filtration. Let $T = \inf\{t : t \geq 0,\ U(t) < 0\}$ and $\psi(u) = \Pr\{T < \infty \mid U(0) = u\}$ be ruin time and ruin probability. □

Lemma 7. *T is F_t^S-stopping time.*

Theorem 8. *$\{H_u(t), F_t,\ t \geq 0\}$ is a martingale, where $H_u(t) = \exp[-rU(t) - tg(r)]$.*

Proof. In fact, $\{U(t),\ t \geq 0\}$ has stationary and independent increments, and from [2] we know that $H_u(t)$ is a martingale if and only if $E(\exp[-RU(t)]) = \exp[-Ru]$. By Lemmas 3 and 4, there is a function $g(r)$ such that $E(\exp[-RS(t)]) = \exp[tg(R)] = 1$; then,

$$\begin{aligned}E\left(\exp\left[-RU(t)\right]\right) &= E\left[\exp\left[-Ru\right]\right]E\left(\exp\left[-RS(t)\right]\right)\\ &= \exp\left[-Ru\right],\end{aligned} \tag{17}$$

so $\{H_u(t), F_t,\ t \geq 0\}$ is a martingale. □

Theorem 9. *For any real number r, the ruin probability $\psi(u)$ satisfies*

$$\psi(u) \leq \exp(-ru)\,E\left[\sup_{t\geq 0}\exp\left(tg(r)\right)\right]. \tag{18}$$

Proof. For a fixed time t_0, $t_0 \wedge T$ is a bounded stopping time. Using the theorem of martingale and stopping time, we have

$$\exp(-ru) = E\left[H_u(0)\right] = E\left[H_u\left(t_0 \wedge T\right)\right]. \tag{19}$$

By the full expectations formula, we have

$$\begin{aligned}\exp(-ru) &= E\left[H_u(T) \mid T \leq t_0\right]\Pr\left(T \leq t_0\right)\\ &\quad + E\left[H_u(t_0) \mid T > t_0\right]\Pr\left(T > t_0\right)\\ &\geq E\left[H_u(T) \mid T \leq t_0\right]\Pr\left(T \leq t_0\right),\end{aligned} \tag{20}$$

which implies

$$\begin{aligned}\Pr\left(T \leq t_0\right) &= \frac{\exp(-ru)}{E\left[H_u(T) \mid T \leq t_0\right]}\\ &\leq \frac{\exp(-ru)}{\inf_{0\leq t\leq t_0}\left[\exp\left(-tg(r)\right)\right]}\\ &= \exp(-ru)\cdot\sup_{0\leq t\leq t_0}\left[\exp\left(tg(r)\right)\right].\end{aligned} \tag{21}$$

By expectation on both sides of the above inequality and letting $t_0 \to \infty$, we can get the desired results. □

Theorem 10. *The ruin probability of surplus process $\{U(t);\ t \geq 0\}$ satisfies*

$$\psi(u) = \frac{\exp(-Ru)}{E\left[\exp\left(-RU(T)\right) \mid T < \infty\right]}. \tag{22}$$

Proof. For a fixed time t_0, $t_0 \wedge T$ is a bounded stopping time. Using the theorem of martingale and stopping time, we have

$$\begin{aligned}\exp(-ru) &= E\left[H_u\left(t_0 \wedge T\right)\right]\\ &= E\left[H_u(T) \mid T \leq t_0\right]\Pr\left(T \leq t_0\right)\\ &\quad + E\left[H_u(t_0) \mid T > t_0\right]\Pr\left(T > t_0\right).\end{aligned} \tag{23}$$

Let $r = R$, we have

$$\begin{aligned}&\exp(-Ru)\\ &\quad = E\left[\exp\left(-RU(T)\right) \mid T \leq t_0\right]\Pr\left(T \leq t_0\right)\\ &\qquad + E\left[\exp\left(-RU(T)\right) \mid T > t_0\right]\Pr\left(T > t_0\right),\end{aligned} \tag{24}$$

which implies

$$\begin{aligned} 0 &\le E\left[\exp\left(-RU\left(T\right)\right) \mid T > t_0\right] \Pr\left(T > t_0\right) \\ &\le E\left[\exp\left(-RU\left(t_0\right)\right) I\left(U\left(t_0\right) \ge 0\right)\right]. \end{aligned} \tag{25}$$

Since $0 \le \exp(-RU(t_0))I(U(t_0) \ge 0) \le 1$, by the law of large numbers, when $t_0 \to \infty$, $U(t_0) \to \infty$ (a.s.). By dominated convergence theorem, we have

$$\lim_{t_0 \to \infty} E\left[\exp\left(-RU\left(T\right)\right) \mid T > t_0\right] \Pr\left(T > t_0\right) = 0, \quad \text{(a.s.)}. \tag{26}$$

Then, when $t_0 \to \infty$ in (26), we can obtain (22). □

Corollary 11. *For the surplus process $\{U(t);\ t \ge 0\}$, the ruin probability $\psi(u)$ satisfies Lundberg inequality:*

$$\psi\left(u\right) \le \exp\left(-Ru\right), \quad u \ge 0, \tag{27}$$

where R is adjustment coefficient.

In order to get this inequality as good as possible, we shall choose r as large as possible under the restriction $\sup_{t\ge 0} \exp(tg(r)) < \infty$. Combined with Figure 1, we have $R = \sup\{r \mid g(r) \le 0\}$.

Theorem 12. *The ruin probability of insurance company before time t satisfies*

$$\psi_t\left(u\right) \le \begin{cases} e^{-R_y u}, & t < \dfrac{u}{g'\left(R\right)} \\ e^{-Ru}, & t \ge \dfrac{u}{g'\left(R\right)}, \end{cases} \tag{28}$$

where $R_y = f(r_y)$ and r_y is the solution of $g'(r) = u/t$.

Proof. By Lemma 4, we have

$$\begin{aligned} \psi_t\left(u\right) &\le e^{-ru} \cdot \sup_{0\le t\le\infty} \left[e^{tg(r)}\right] = e^{-ru} \cdot \max\left\{1, e^{tg(r)}\right\} \\ &\le e^{-u\min\{r, r-(t/u)g(r)\}} = e^{-u\min\{r, f(r)\}}, \end{aligned} \tag{29}$$

where $f(r) = r - yg(r)$, $y = t/u$. Obviously, the supremum of $\psi_t(u)$ is $e^{-u\sup_{r\ge 0}\{\min\{r, f(r)\}\}}$.

Since $f(r)$ is the convex function and $f(R) = R$, when $r > R$, we have $f(r) < r$; when $0 < r < R$, we have $f(r) > r$. Let r_y be the solution of $f'(r) = 0$ and $R_y = f(r_y)$; then, R_y is the maximum value of $f(r)$.

And because

$$\begin{aligned} r_y &\le R, \quad t > \frac{u}{g'\left(R\right)} \\ r_y &> R, \quad t \le \frac{u}{g'\left(R\right)}, \end{aligned} \tag{30}$$

then, we have

$$\sup_{r\ge 0}\left\{\min\left[r, f\left(r\right)\right]\right\} = \begin{cases} R_y, & t < \dfrac{u}{g'\left(R\right)} \\ R, & t \ge \dfrac{u}{g'\left(R\right)}. \end{cases} \tag{31}$$

Thus, Theorem 12 is obtained. □

Theorem 13. *Let $\tau = \inf\{t \ge 0,\ U(t) = x > u\}$ be the time when the surplus reaches a given level firstly; then, the Laplace transform of τ is as follows:*

$$E\left[e^{-s\tau}\right] = e^{rx}, \tag{32}$$

where s and r satisfy $s = g(r)$.

Proof. For the surplus process $\{U(t);\ t \ge 0\}$, using the theorem of martingale and stopping time, we see that τ is a stopping rime of F_t^S. Let $H(t) = e^{-rU(\tau)-s\tau}$. By Theorem 8, the surplus process $\{H(t);\ t \ge 0\}$ is a martingale. Hence, we have $E[H(\tau)] = E[H(0)]$; that is, $E[e^{-rU(\tau)-s\tau}] = 1$. Since $U(\tau) = x$, then we get $E[e^{-rx-s\tau}] = 1$; that is, $E[e^{-s\tau}] = e^{rx}$. □

Theorem 14. *The expectation and variance of τ satisfy*

$$\begin{aligned} E\left[\tau\right] &= \frac{x}{\eta} \\ \operatorname{Var}\left[\tau\right] &= \frac{x\omega}{\eta^3}, \end{aligned} \tag{33}$$

where

$$\begin{aligned} \eta &= A\alpha + \frac{\lambda\mu_X}{1-\rho} - \frac{\lambda p\mu_Y}{1-\rho} - \frac{\lambda q\mu_Z}{1-\rho}, \\ \omega &= \frac{\lambda}{\left(1-\rho\right)^2}\left\{\left(1-\rho\right)\left(\sigma_X^2 + \mu_X^2\right) + 2\rho\mu_X^2\right\} \\ &\quad + \frac{\lambda p}{\left(1-\rho\right)^2}\left\{\left(1-\rho\right)\left(\sigma_Y^2 + \mu_Y^2\right) + 2\rho\mu_Y^2\right\} \\ &\quad + \frac{\lambda q}{\left(1-\rho\right)^2}\left\{\left(1-\rho\right)\left(\sigma_Z^2 + \mu_Z^2\right) + 2\rho\mu_Z^2\right\} + \beta^2. \end{aligned} \tag{34}$$

Proof. By Theorem 13, we have $E[e^{-s\tau}] = e^{rx}$; let $\varphi(s) = \ln E[e^{-s\tau}]$; that is, $\varphi(s) = rx$; then, we have the following:

$$\begin{aligned} \varphi'\left(s\right) &= \frac{d\varphi\left(s\right)}{dr} \cdot \frac{1}{ds/dr} = \frac{x}{s'\left(r\right)} = \frac{x}{g'\left(r\right)} \\ \varphi''\left(s\right) &= \frac{d\varphi'\left(s\right)}{ds} = \frac{d\varphi'\left(s\right)}{dr} \cdot \frac{1}{s'\left(r\right)} = -\frac{xg''\left(r\right)}{\left[g'\left(r\right)\right]^3}. \end{aligned} \tag{35}$$

By Lemma 5, we have

$$\begin{aligned} E\left[\tau\right] &= -\left.\frac{d\varphi\left(s\right)}{ds}\right|_{s=r=0} = -\frac{x}{g'\left(0\right)} \\ &= -\frac{x}{-A\alpha - \lambda\mu_X/\left(1-\rho\right) + \lambda p\mu_Y/\left(1-\rho\right) + \lambda q\mu_Z/\left(1-\rho\right)} \\ &= \frac{x}{\eta} \\ \operatorname{Var}\left[\tau\right] &= \left.\varphi''\left(s\right)\right|_{s=r=0} = \frac{x\omega}{\eta^3}. \end{aligned} \tag{36}$$

□

Table 1: The upper bound of the ruin probability.

$1/b$	R	$\exp(-Ru)$										
		$u = 1000$	$u = 1100$	$u = 1200$	$u = 1300$	$u = 1400$	$u = 1500$	$u = 1600$	$u = 1700$	$u = 1800$	$u = 2000$	$u = 10000$
1/1000	0.0010	0.3677	0.3326	0.3010	0.2723	0.2464	0.2229	0.2017	0.1825	0.1651	0.1325	0.00004
1/1100	$9.0960e-004$	0.4027	0.3677	0.3357	0.3065	0.2799	0.2555	0.2333	0.2130	0.1945	0.1622	0.00011
1/1200	$8.3376e-004$	0.4344	0.3997	0.3677	0.3383	0.3112	0.2863	0.2634	0.2423	0.2230	0.1887	0.00024
1/1300	$7.6959e-004$	0.4632	0.4289	0.3971	0.3677	0.3405	0.3152	0.2919	0.2703	0.2503	0.2146	0.00045
1/1400	$7.1460e-004$	0.4894	0.4556	0.4242	0.3950	0.3677	0.3424	0.3187	0.2968	0.2763	0.2395	0.00079
1/1500	$6.6694e-004$	0.5133	0.4802	0.4492	0.4202	0.3931	0.3667	0.3440	0.3218	0.3010	0.2635	0.0013
1/1600	$6.2524e-004$	0.5351	0.5027	0.4722	0.4436	0.4167	0.3915	0.3677	0.3455	0.3245	0.2864	0.0019
1/1700	$5.8845e-004$	0.5552	0.5235	0.4935	0.4653	0.4387	0.4137	0.3900	0.3667	0.3467	0.3082	0.0028
1/1800	$5.5574e-004$	0.5736	0.5426	0.5133	0.4856	0.4953	0.4345	0.4110	0.3888	0.3678	0.3291	0.0039
1/1900	$5.2649e-004$	0.5907	0.5604	0.5316	0.5044	0.4785	0.4540	0.4307	0.4086	0.3876	0.3489	0.0052
1/2000	$5.0015e-004$	0.6064	0.5769	0.5487	0.5219	0.4965	0.4723	0.4492	04273	0.4065	0.3678	0.0067

4. Numerical Simulation and Analysis

4.1. Simulation of the Upper Bound of Ruin Probability. Since the adjustment coefficient R can be used to measure the risk, by formula $\psi(u) \leq \exp(-Ru)$, we know that the higher adjustment coefficient R results in the less ruin probability. The following will be used to simulate the size of ruin probability by adjustment coefficient R.

Suppose $\lambda = 40$, $\rho = 0.1$, $A = 400$, $\alpha = 0.1$, $\beta = 1$, $p = 0.005$, and $q = 0.0005$. Random variables X_k, Y_k, and Z_k obey the exponential distribution with mean $\mu_X = b$, $\mu_Y = 20$, and $\mu_Z = 10$, respectively. Taking different values about b and u, by MATLAB and Theorem 6, we can get the upper bound of the ruin probability. See Table 1.

Table 1 shows that ruin probability of insurance company varies tremendously in size with different value of u; the higher initial surplus u results in the less ruin probability. Magnitude of initial capital increase is far below the level of reduced number of ruin probability of insurance company. For example, in the first seven behaviors of Table 1, the initial capital u is only increased by 10 times, and the ruin probability decreases from 0.4894 to 0.00079. This is another example in the course of business; the availability of sufficient initial capital is crucial to the insurance company.

In addition, determining the value of distribution parameter $1/b$ of the premium X has a great impact on the ruin probability of insurance company. For the exponential distribution, the value of $1/b$ is smaller; the smaller the amount of premium charged by insurance company, the greater the probability of ruin. This suggests that a reasonable determination of the premium on the normal operation of insurance companies is very important, which leads to higher requirement for determination of the premium in the design of insurance products in the insurance company.

4.2. Simulation of the First Arrival Time. Suppose $\lambda = 40$, $\rho = 0.1$, $A = 5000$, $\alpha = 0.1$, $x = 2000$, $\beta = 1$, $p = 0.005$, and $q = 0.0005$ (in order to clear the trend of curve, in Figure 4, let $q = 0.0125$; in Figure 8, let $q = 0.005$). Random variables X_k, Y_k, and Z_k obey the exponential distribution with mean $\mu_X = 100$, $\mu_Y = 80$, and $\mu_Z = 60$, respectively. We obtained the trend chart of average time and the variance of time.

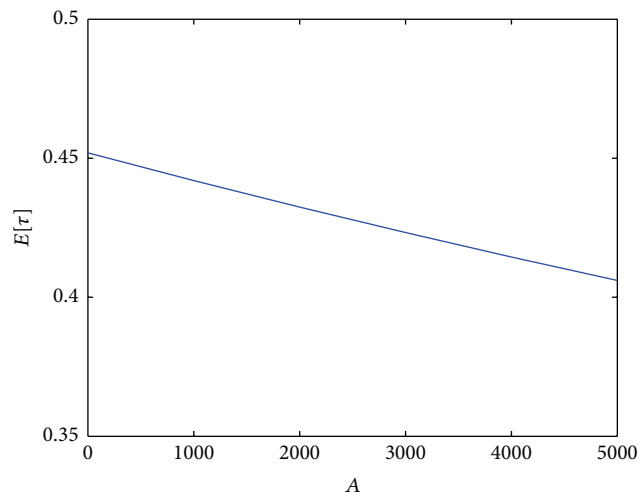


Figure 2

(1) The Trend Figure of Average Time. From Figures 2–5, we can clearly know that the first arrival time to reach a given level is a decreasing function of investment capital A and investment interest rate α and is an increasing function of claim rate λp and surrender rate λq.

In comparison, the change in the average time is more sensitive to the change of investment capital and investment rates and is not sensitive to the change of claim rate and surrender rate.

(2) The Trend Figure of Variance of the First Arrival Time. Figures 6–9 point out that the first arrival time to reach a given level is a decreasing function of investment capital A and investment interest rate α and is an increasing function of claim rate λp and surrender rate λq. In comparison, the change in variance is more sensitive to the change of investment capital and investment rates and is not sensitive to the change of claim rate and surrender rate.

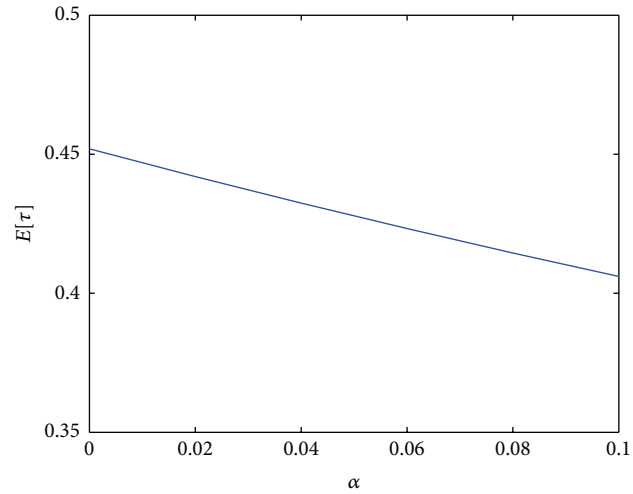


Figure 3

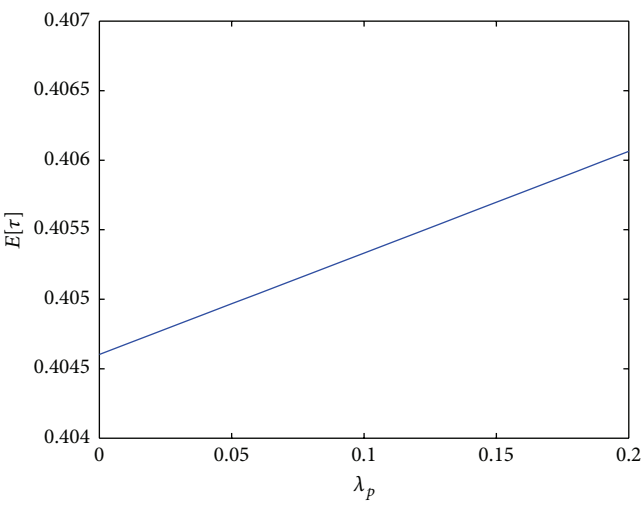


Figure 4

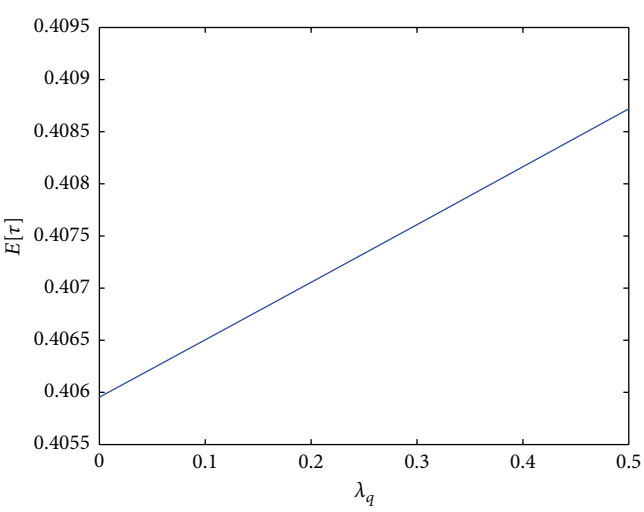


Figure 5

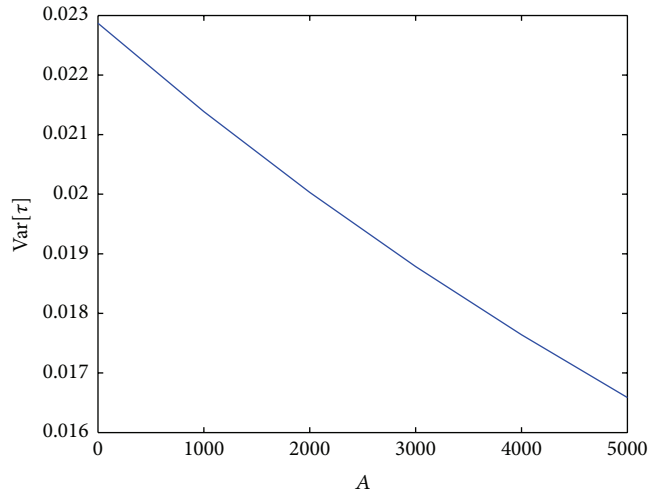


Figure 6

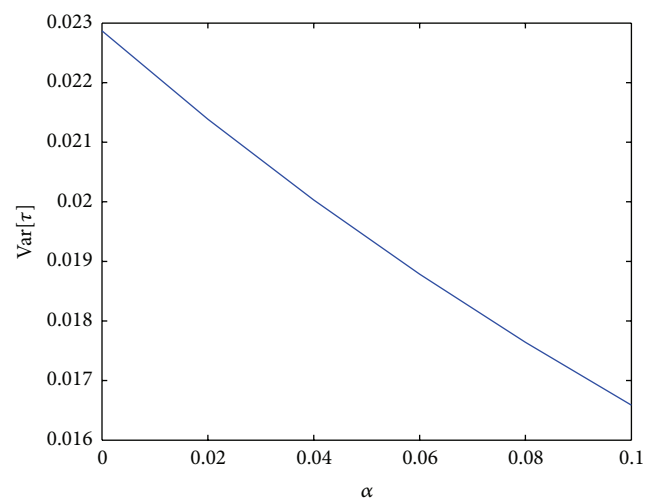


Figure 7

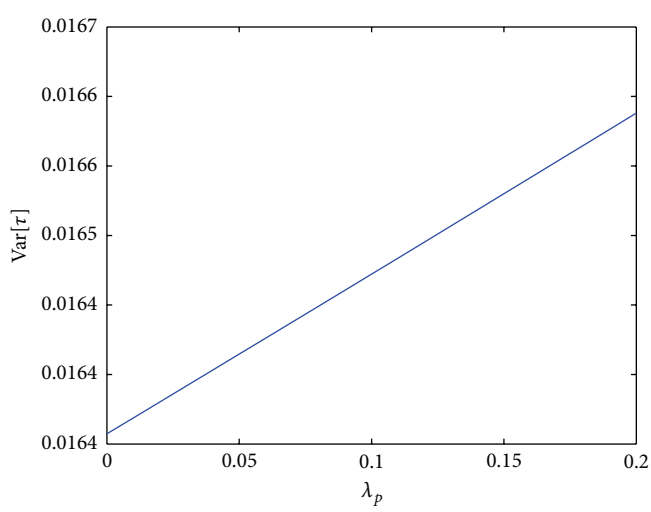


Figure 8

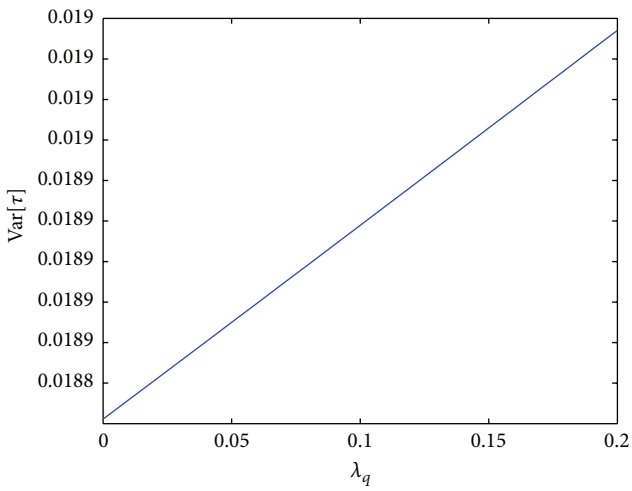


Figure 9

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by National Social Science Foundation of China (no. 15BJY007), National Natural Science Foundation of China (no. 11301303 and no. 11231005), Humanities and Social Sciences Project of the Ministry of Education of China (no. 14YJA630088, no. 14YJC790054, no. 13YJA790035, no. 13YJC630150, and no. 13YJC630162), Natural Science Foundation of Shandong Province (no. ZR2012AQ013 and no. ZR2010GL013), A Project of Shandong Province Higher Educational Science and Technology Program (J15LI03, J15LI53), the 2014 Youth Talent Support Program of Shandong University of Finance and Economics, and the Startup Foundation of Doctor Scientific Research of Shandong Jiaotong University.

References

[1] H. U. Gerber, *An Introduction to Mathematical Risk Theory*, vol. 8 of *S. S. Heubner Foundation Monograph Series*, 1979.

[2] H. U. Gerber and E. S. W. Shiu, "On the time value of ruin," *North American Actuarial Journal*, vol. 2, no. 1, pp. 48–78, 1998.

[3] J. Liu, J. C. Xu, and Y. J. Hu, "On the expected discounted penalty function in a Markov-dependent risk model with constant dividend barrier," *Acta Mathematica Scientia B*, vol. 30, no. 5, pp. 1481–1491, 2010.

[4] P. Shi, X. P. Feng, and A. Ivantsova, "Dependent frequency—severity modeling of insurance claims," *Insurance: Mathematics & Economics*, vol. 64, pp. 417–428, 2015.

[5] T. Jiang, Y. Wang, Y. Chen, and H. Xu, "Uniform asymptotic estimate for finite-time ruin probabilities of a time-dependent bidimensional renewal model," *Insurance: Mathematics and Economics*, vol. 64, pp. 45–53, 2015.

[6] Z. M. Zhang and H. Yang, "Gerber-Shiu analysis in a perturbed risk model with dependence between claim sizes and interclaim times," *Journal of Computational and Applied Mathematics*, vol. 235, no. 5, pp. 1189–1204, 2011.

[7] Y. F. Shi, P. Liu, and C. S. Zhang, "On the compound Poisson risk model with dependence and a threshold dividend strategy," *Statistics & Probability Letters*, vol. 83, no. 9, pp. 1998–2006, 2013.

[8] W. Zou, J.-W. Gao, and J.-H. Xie, "On the expected discounted penalty function and optimal dividend strategy for a risk model with random incomes and interclaim-dependent claim sizes," *Journal of Computational and Applied Mathematics*, vol. 255, pp. 270–281, 2014.

[9] Z. C. Mao and J. E. Liu, "The distribution about numbers of claims on homogeneous policyholders under NCD system and stop loss insurance," *Chinese Journal of Management Science*, vol. 13, no. 5, pp. 1–5, 2005.

[10] W. Yu, "Some results on absolute ruin in the perturbed insurance risk model with investment and debit interests," *Economic Modelling*, vol. 31, no. 1, pp. 625–634, 2013.

[11] G. H. Guan and Z. X. Liang, "Optimal reinsurance and investment strategies for insurer under interest rate and inflation risks," *Insurance: Mathematics and Economics*, vol. 55, pp. 105–115, 2014.

[12] W. Yu, "Randomized dividends in a discrete insurance risk model with stochastic premium income," *Mathematical Problems in Engineering*, vol. 2013, Article ID 579534, 9 pages, 2013.

[13] Y. J. Huang and W. G. Yu, "Studies on a double poisson-geometric insurance risk model with interference," *Discrete Dynamics in Nature and Society*, vol. 2013, Article ID 128796, 8 pages, 2013.

[14] Z. C. Mao and J. E. Liu, "The expression of ruin probability under claim number with compound Poisson-Geometric process," *Chinese Journal of Management Science*, vol. 15, no. 5, pp. 23–28, 2007.

[15] X. Lin and N. Li, "Ruin probability, optimal investment and reinsurance strategy for an insurer with compound poisson-geometric risk process," *Mathematica Applicata*, vol. 24, no. 1, pp. 174–180, 2011.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 591232, 9 pages
http://dx.doi.org/10.1155/2015/591232

Research Article

Coordinated Stability Control of Wind-Thermal Hybrid AC/DC Power System

Zhiqing Yao,[1] Zhenghang Hao,[2] Zhuo Chen,[2] and Zhiguo Yan[3]

[1] *School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China*
[2] *School of Electrical Engineering, Guizhou University, Guiyang 550025, China*
[3] *School of Electrical Engineering and Automation and Key Laboratory of Pulp and Paper Science & Technology of Ministry of Education of China, Qilu University of Technology, Jinan 250353, China*

Correspondence should be addressed to Zhiguo Yan; yanzg500@sina.com

Received 12 August 2015; Accepted 16 September 2015

Academic Editor: Xinguang Zhang

The wind-thermal hybrid power transmission will someday be the main form of transmitting wind power in China but such transmission mode is poor in system stability. In this paper, a coordinated stability control strategy is proposed to improve the system stability. Firstly, the mathematical model of doubly fed wind farms and DC power transmission system is established. The rapid power controllability of large-scale wind farms is discussed based on DFIG model and wide-field optical fiber delay feature. Secondly, low frequency oscillation and power-angle stability are analyzed and discussed under the hybrid transmission mode of a conventional power plant with wind farms. A coordinated control strategy for the wind-thermal hybrid AC/DC power system is proposed and an experimental prototype is made. Finally, real time simulation modeling is set up through Real Time Digital Simulator (RTDS), including wind power system and synchronous generator system and DC power transmission system. The experimental prototype is connected with RTDS for joint debugging. Joint debugging result shows that, under the coordinated control strategy, the experimental prototype is conductive to enhance the grid damping and effectively prevents the grid from occurring low frequency oscillation. It can also increase the transient power-angle stability of a power system.

1. Introduction

As one of the most efficient new energy sources that have the potential of large-scale development, wind power generation has developed speedily in China. Due to the limited ability of electricity consumption, wind power in northwest China, northeast China, and north China should be transmitted to the load center by long-distance transmission line [1, 2]. Given that wind power is not constant and it is not economical to transmit wind power alone, there arises the necessity to "bond" large-scale wind farms with thermal power plant so as to realize transregional transmission. However, hybrid transmission system easily triggers low frequency oscillation or angle instability [3, 4].

Damping features and controllable strategy of flywheel energy storage device [5], flexible power conditioner [6], and static series compensator [7] were studied to improve damping of the power system in previous research. But the cost of large-scale power electronic equipment is so high that it limits the application. In comparison, doubly fed wind generator can realize decoupling control of active and reactive power [8]. Excitation converter in the wind turbine system can be adjusted to control the active power in the transient process. As a result, for large-scale wind farms, active power in the whole wind farms can be adjusted randomly and quickly through the communication network, making the wind farms controllable. So long as the active power is able to be adjusted, it is possible to enhance the stability of the power system. Controllable power in the wind farms can not only help increase the damping and prevent low frequency oscillation, but also enhance transient angle stability of the power system, which is meaningful for ensuring grid safety.

In this paper, firstly, structural features of hybrid power system of wind farms and thermal power plant are analyzed and problems about stability of the power transmission system are pointed out. Secondly, the mathematical model of

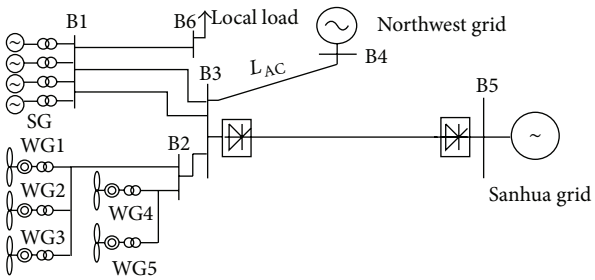


Figure 1: AC/DC power system with wind farms incorporated.

doubly fed wind farms and DC power transmission system is established. And it is proved that active power in the large-scale wind farms is controllable. Thirdly, for the purpose of increasing damping of the system and enhancing angel stability, a coordinated control strategy for wind farms and DC power transmission system is proposed. Finally, the experimental prototype is made and the control effect of the experimental prototype is also introduced in detail.

2. AC/DC Power Transmission System Incorporating Wind Farms

Chinese large-scale wind farms are usually located at remote areas. Due to small load capacity, power generated by the large-scale wind farms cannot be consumed. So long-distance transmission is an inevitable solution. But as the wind power is not constant, long-distance transmission is costly if wind power is the only thing to transmit. And irregular fluctuation of the wind power would make the grid unstable. So, at the present time, a hybrid transmission mode of wind farms and thermal power plant (also named as wind-thermal hybrid power system) [9] is the main form of transmission, as is shown in Figure 1. To be more specific, wind farms are connected with the thermal power plant nearby and the power is transmitted to other areas by extra high voltage (EHV) line.

Wind-thermal hybrid power transmission mode is for long-distance transmission. But as the utilization hours of wind power are lesser than the conventional thermal power plant, the transmission channel capacity of wind power is designed lower than the maximum power. When large-scale wind power is generated, the transmission line will be heavily loaded, and the thermal power plant also functions in a heavy-load state, resulting in low frequency oscillation. In addition, vector control of the turbine may pose influence on damping [10]. It is also worth noticing that the thermal power unit bonded with the wind farms is responsible for curbing wind power fluctuation. So the adjustment for the unit is frequent, which would damage the power system stabilizer (PSS) [11] and hamper damping characteristic. Therefore, effective measures on damping should be taken to ensure safe operation of the wind-thermal power transmission system.

In addition, the wind-thermal hybrid power system does not work in a conventional way. The dynamic behaviors of wind farms may weaken the angle stability of the thermal power plant, even the whole power system. Thus, the wind farms and the AC/DC power transmission system should be subjected to coordinated control in order to erase negative effect caused by the wind farms.

3. Mathematical Model for Wind-Thermal Hybrid Power Transmission

3.1. Wind Turbine System. The mechanical part of the wind turbine system includes wind turbine, transmission shaft, and gearbox. Wind turbine is used to capture wind energy through the turbines and transform it to the mechanical torque on the wheel hub. The shaft and gearbox is used to pass the driven force of the wind turbine to the generator and increase the revolving speed. The gear ratio can reach 100. To simplify calculation, the mechanical part is regarded as a concentrated mass expressed by first-order inertial element [8]:

$$\frac{\mathrm{d}P_m}{\mathrm{d}t} = \frac{1}{T_d}\left(P_T - P_m\right), \tag{1}$$

where P_m and P_T refer to mechanical power and electromagnetic power, respectively, on the rotor of the generator. T_d refers to inertia time constant.

3.2. Mathematical Model for DFIG. DFIG is actually the rotor asynchronous motor. There are symmetrical three-phase windings on the stator and the rotor. The modeling process is similar to that of asynchronous motor and synchronous generator, in which the primitive equation is confirmed in the three-phase static coordinate system and then coordinates are transformed. Unlike the modeling of synchronous generator, dq coordinates of DFIG can be oriented in different modes, such as stator flux mode, rotor flux mode, and stator voltage mode. And orientation of dq coordinates of synchronous generator is only to take physical location of the rotor. In this paper, the stator vector voltage of DFIG is taken as axis q. In the dq coordinates system, the stator flux (ψ_{ds}, ψ_{qs}) and rotor current (i_{dr}, i_{qr}) are taken as the state variables. The state equation is expressed as follows [11]:

$$\begin{aligned}
p\psi_{ds} &= -\frac{r_s}{l_s}\psi_{ds} + l'' r_s i_{dr} + \omega_1 \psi_{qs},\\
p\psi_{qs} &= -\frac{r_s}{l_s}\psi_{qs} + l'' r_s i_{qr} - \omega_1 \psi_{ds} + u_{qs},\\
pl' i_{dr} &= -r_r i_{dr} + u_{dr} + \omega_s l' i_{qr} + \omega_s l'' \psi_{qs} - l'' p\psi_{ds},\\
pl' i_{qr} &= -r_r i_{qr} + u_{qr} - \omega_s l' i_{dr} - \omega_s l'' \psi_{ds} - l'' p\psi_{qs},
\end{aligned} \tag{2}$$

where $l' = (l_r - l_m l_m / l_s)$ and $l'' = l_m / l_s$, l_s, l_r, and l_m being stator self-inductance, rotor self-inductance, and mutual inductance, respectively; r_s and r_r are stator and rotor resistance, respectively; ω_1 and ω_s are synchronous speed and slip, respectively; u_{dr} and u_{qr} are vertical and horizontal vector of excitation, respectively; u_{qs} is stator voltage; and p is differential operator.

3.3. UHVDC Power Transmission System Model. UHVDC power transmission refers to DC transmission based on thyristor inverter. It consists of the inverter, DC line, and

the auxiliary equipment [12, 13]. Quasi steady state model is used to simulate the UHVDC primary system. DC commutation is described by algebraic equation. DC line and smoothing reactor are described in the T-equivalent-circuit model [14–16]. Substitute the algebraic equation for the differential equation and get the mathematical model for UHVDC power transmission system expressed by [17, 18]

$$\begin{aligned} I_{dr} &= \frac{1}{L_{dr\Sigma}}\left(-R_d I_{dr} - U_c + \frac{3\sqrt{2}}{\pi}U_{dr}\cos\alpha - \frac{3}{\pi}X_r I_{dr}\right), \\ I_{di} &= \frac{1}{L_{di\Sigma}}\left(-R_d I_{di} - U_c + \frac{3\sqrt{2}}{\pi}U_{di}\cos\beta - \frac{3}{\pi}X_i I_{di}\right), \\ U_c &= \frac{1}{C}\left(I_{dr} - I_{di}\right), \end{aligned} \tag{3}$$

where I_{dr}, I_{di}, and U_c are state variable. I_{dr} and I_{di} are DC current of rectifier and inverter; U_c refers to the voltage in the middle of DC line; R_d is direct current resistance; C is earth capacitance equivalent to DC line; U_{dr} and U_{di} are DC voltage of rectifier and inverter; $L_{dr\Sigma}$ and $L_{di\Sigma}$ are equivalent inductance of rectifier and inverter; X_r and X_i are commutation reactance of rectifier and inverter; and α and β are trigger delay angle of rectifier and trigger angle of inverter.

4. Conditions for Quick Adjustment of the Wind Farms

4.1. Quick Adjustment of the Active Power of Single Turbine in Transient Process. Before the wind farms realize the effectively quick adjustment, it is made sure that every turbine is highly controllable. From (2), the rotor current can be controlled by rotor voltage. But i_{dr} and i_{qr} are cross-coupled, so feedforward compensation scheme is usually adopted to realize decoupling control. The feedforward compensation is calculated from rotor current and internal flux variables:

$$\begin{aligned} e_{dr} &= k\left(\omega_s l' i_{qr} + \omega_s l'' \psi_{qs}\right), \\ e_{qr} &= -k\left(\omega_s l' i_{dr} + \omega_s l'' \psi_{ds}\right). \end{aligned} \tag{4}$$

Add e_{dr} and e_{qr} to u_{dr} and u_{qr}, and get the new control variable. For the state equation of rotor current in (2), substituted feedforward compensation item, the relationship between rotor current and control command (u^*_{dr}, u^*_{qr}) is

$$\begin{aligned} pL' i_{dr} &= -r_r i_{dr} + u^*_{dr}, \\ pL' i_{qr} &= -r_r i_{qr} + u^*_{qr}. \end{aligned} \tag{5}$$

Equation (5) shows that the response of active current and reactive current to the control command is the first-order inertial link. Time constant of inertial element is τ. Typical parameters of DFIG are substituted into τ and we can get it is about 10 ms. This means that the single turbine can respond at the level of ms under external control command.

4.2. Quick Adjustment of the Active Power of Wind Farms in Transient Process. In the previous research, the wind farms were usually equaled to a single wind turbine [8, 11]. Obviously, they are different. Section 4.1 has already proved that the single turbine can respond at the level of ms, but whether this holds true to the wind farms still needs to be proved.

A large-scale wind farm has hundreds of wind turbines. Controlling them depends on wide-field communication technology. The control system of the wind farms has a master-slave structure. There are two methods of communication: (1) one-to-multiple answering transmission and (2) one-to-multiple global broadcast. For method (1), as all turbines (200 sets) are slave turbines, it means 200 messages are sent in a controlling cycle. For method (2), as every turbine receives the same message, only 1 message is sent in a controlling cycle.

Message transmission presents the following features: suppose the length of the message is 200 bits and the serial communication baud rate is 1 Mbps. It is calculated that it costs 0.2 ms to send the message. If fiber communication is adopted for long-distance transmission, 1.5 ms should be used in photovoltaic conversion. So the total time which is the delay time of the fiber communication for long-distance transmission is 1.7 ms. Therefore, time delay in a controlling cycle in method (1) is 200 ms and that in method (2) is 1.7 ms. Obviously, method (2) is suggested as 1.7 ms time delay will not pose significant influence on the closed-loop control system.

According to the analysis in Sections 4.1 and 4.2, based on global broadcast fiber communication technology, the wind farms can be an active power source which is able to be adjusted quickly.

5. Coordinated Control Strategy of Wind Farms and DC Power Transmission System

5.1. Basic Ideas and Purposes of Coordinated Control. Even though it may weaken the damping of conventional power plant and angle stability when the wind farms are connected with the thermal power plant, the ability of power adjustment of the wind farms would increase the damping of synchronous generator and enhance angle stability. Basic ideas of coordinated control of wind farms and DC power transmission system are mainly described as follows: the revolving speed or the frequency of synchronous generator of the conventional power plant is fed back to the controller; then the controlled quantity that can activate small-scale dynamic active power in the wind farms is produced through gain calculation and phase correction; and the dynamic active power of the wind farms propels the synchronous generator to produce electromagnetic torque with damping characteristic. So the damping can be increased and oscillation can be restrained. The key to supply the synchronous generators with the positive damping is that the wind farms must be controllable and can be controlled quickly.

5.2. Technical Framework of Coordinated Control. Based on the controllability of the wind farms and DC transmission system, the damping control principle of the thermal power

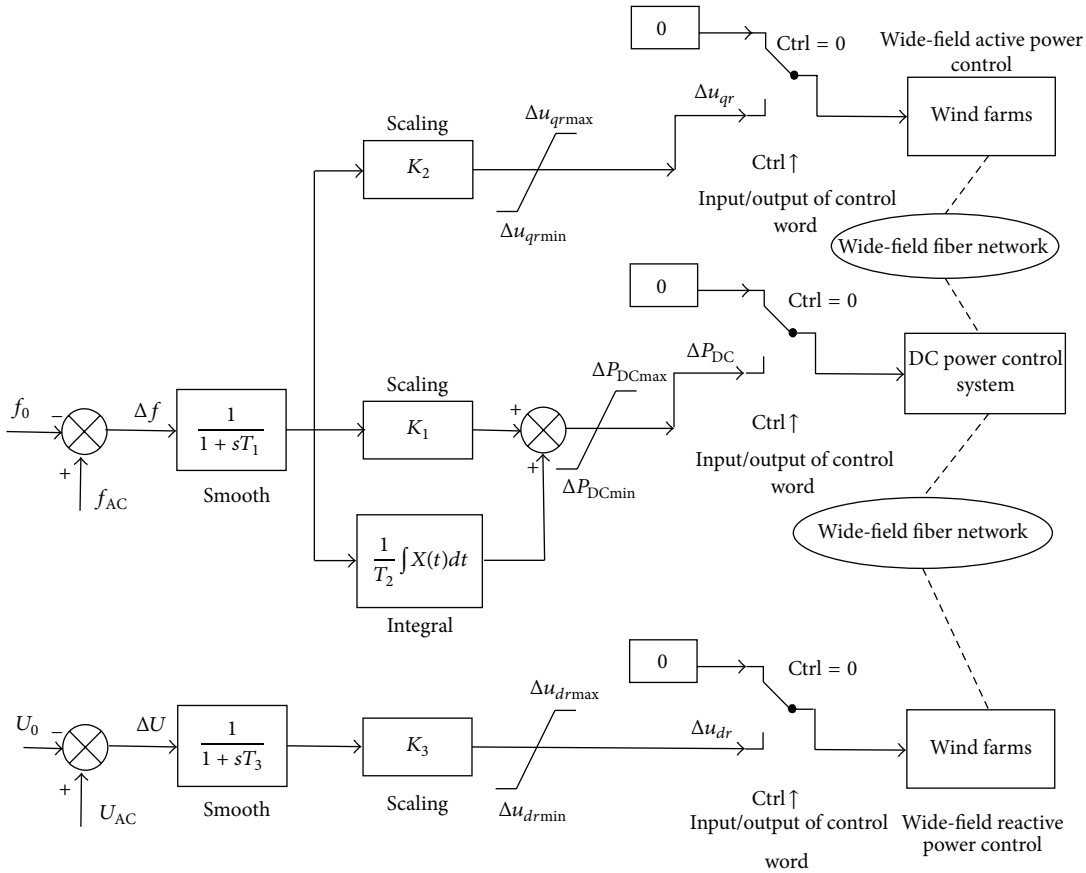


Figure 2: Coordinated control strategy framework.

plant is described in Figure 2. Firstly, the frequency of common DC-bus of the wind farms (Δf_{pcc}) or speed or angle of the synchronous generator is collected. Secondly, the collected signal passes smooth block and scaling block and integral link. Then the controlled signal of the ith wind turbine (Δu_{qri}) is confirmed according to its working state and the allocation algorithm. This controlled signal is sent to the active circuit of each wind turbine through wide-field fiber communication network (excitation voltage reference point at axis q of dq decoupling control of the excitation converter) so that the wind turbines can adjust the active power synchronically. As a result, the active power in the transient process can increase the damping of the synchronous generator and prevent low frequency oscillation.

Parameter design of damping controller is expressed as follows: the value of T_1 is set up under the condition that the low frequency signals are able to pass through; and the angle for compensation is figured out by calculating the dynamic frequency before adjustment and the values of T_2 and T_3 are calculated; based on these values, the value of K is confirmed according to the expected dynamic frequency.

When the power transmission system is disturbed, the most important thing is to extract fault characteristic quantity and analyze the type and the place of the fault, in other words, to judge whether it occurs in the DC system or the AC system. If the fault occurs in the DC system, DC block results in the great reduction of power. At this moment, the active output of the wind farm should be lowered within controllable time to protect the synchronous generator from instability. And the reactive output is captured to prevent abrupt rise of the voltage and the instability of the wind turbine.

When the fault occurs in the AC system, the power reduces substantially. The output power of the DC system should be increased within set time. But the sudden supply of DC power may increase the demand of reactive power and decrease AC bus voltage at the converter station. Thus, the reactive output of the wind farms should be adjusted quickly to prevent voltage fall. At the same time, the reactive power demand in the transient process is calculated according to the output power of DC system and the reactive power is sent from the wind farms.

6. Modeling and Prototype Test Based on RTDS

An experimental prototype is designed according to the coordinated control strategy framework mentioned in Section 5.2. To test the coordinated control strategy and verify the effectiveness of the experimental prototype, "hardware in-the-loop simulation" is conducted. The experimental prototype is the real object and the wind-thermal power transmission system is the visual object based on Real Time Digital System (RTDS).

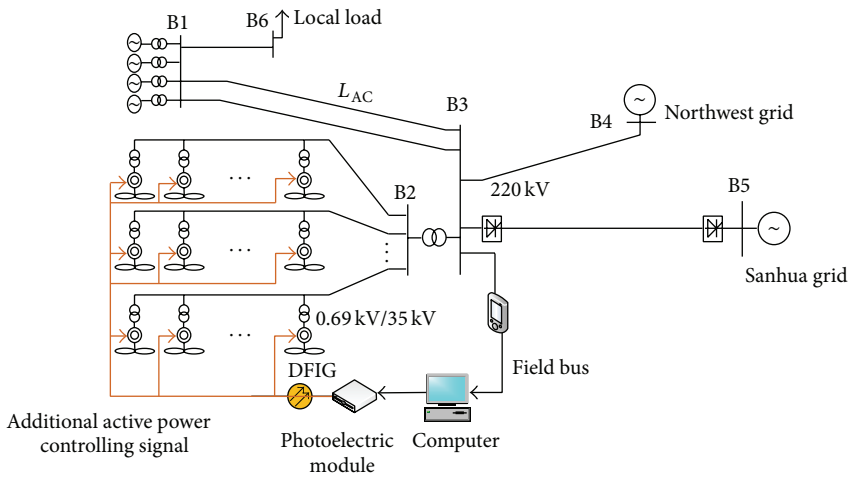


Figure 3: Simulation case.

6.1. Simulation of the Wind-Thermal Power Transmission System. The wind-thermal power transmission system is simulated and tested, as shown in Figure 3. Compared to the conventional power plant, each wind turbine has small capacity and the wind farms have a large number of wind power units [19]. It is impossible to simulate every turbine set [14]. Therefore, "equivalent similitude ratio" method is adopted in the simulation. In other words, the large-scale wind turbine system is replaced by a relatively small DFIG in which there are many wind turbine sets closely related to each other. Parameters are scaled down to a proper proportion. Thus, a large-scale wind farm is divided into sections and each section is simulated by DFIG. As a result, electromagnetism and transient process can be better reflected and the process is simplified to make the simulation close to the real situation.

The proposed wind-thermal power transmission system model based on RTDS consists of six wind turbines and one synchronous generator set. Wind-thermal capacity is in ratio of 1 : 1.5. According to the principle of "equivalent per-unit value of parameter," the capacity of the synchronous generator is scaled down to the level of MW. The rated capacity of a DFIG is 2.2 MVA and the rated frequency is 60 Hz.

In Figure 3, parameters of the DFIG are as follows: stator winding resistance is 0.00462 p.u., stator leakage inductance is 0.102 p.u., rotor winding resistance is 0.00736, rotor leakage inductance is 0.11 p.u., and stator and rotor mutual inductance is 2.62 p.u.; parameters of synchronous generator are (refer to literature [16] for name and physical definition)

$$\begin{aligned} x_d &= 0.51 \text{ p.u.}, \\ x_d' &= 0.042 \text{ p.u.}, \\ x_d'' &= 0.032 \text{ p.u.}, \\ x_q &= 0.375 \text{ p.u.}, \\ x_q'' &= 0.011 \text{ p.u.}, \\ T_d' &= 0.33 \text{ s}, \\ T_d'' &= 0.03 \text{ s}, \\ T_q' &= 0.03 \text{ s}, \\ H &= 6.98 \text{ s}; \end{aligned} \tag{6}$$

and parameters of the additional damping controller are

$$\begin{aligned} T_1 &= 5.32 \text{ s}, \\ T_2 &= 0.06 \text{ s}, \\ T_3 &= 0.38 \text{ s}, \\ K &= 12.9. \end{aligned} \tag{7}$$

6.2. RTDS Hardware Requirement and Calculation Assignment. RTDS hardware has the following requirement: 10 processors (GPC-PB5) of 2 RACKs and 1 12-channel analog input card (GTAO) are used in the whole model. RTDS is the real time simulation equipment, and the processors must be properly allocated when modeling. To enhance simulation accuracy, small-step (<2 us) RTDS/RSCAD system is used as the carrier of DFIG system model. Each model includes DFIG and PWM frequency converter and transformer. Small-step is set up in the VSC module in the small-step model base. Each VSC module has a corresponding processor (GPC-PB5). So among 10 processors, 6 of them correspond to 6 DFIGs' models, respectively, and the remaining 4 are for controlling calculation and the synchronous generator simulation and the grid simulation.

6.3. Experimental Prototype and Interface of RTDS. The wind farms, the thermal power system, and the AC/DC power transmission system shown in Figure 3 are simulated by RTDS. As the analog state variables, the frequency of the common bus of wind farms, the speed of synchronous generator, and the power-angle are the output from GTAO of RTDS, and they are put into the data collection module of the experimental prototype by the signal cable. To simulate actual

Figure 4: Joint debugging of the experimental prototype and RTDS.

fiber channel, the photoelectric conversion module and the 3 km single-mode fiber are set up in the experimental device. The additional damping control signals are produced after the signals collected from RTDS are processed through the data management module and the algorithm producing module. Then, the additional damping control signals are connected to GTAI of RTDS through profibus and photoelectric module to control the stability of the system.

In Figure 3, the first-order part of DC power transmission system is simulated by RTDS and the controller of DC power transmission system is a special controlling system developed on DPS3000 platform. It is connected with RTDS through signal cable. After the connection, RTDS and the experimental device construct a closed controlling system in which the experimental device is the controller and RTDS is being controlled, as is shown in Figure 4.

7. Joint Debugging of Experimental Device and RTDS

7.1. Experiment Analysis of Damping Characteristic. The experiment is described as follows: rectify excitation parameters and active power of the synchronous generator to produce weak damping; set up three-phase transient circuit fault at the common DC-bus to activate low frequency oscillation; and record the speed of the synchronous generator, active power of the wind turbine, stator current, and rotor current.

(1) Record the speed of the synchronous generator and observe the additional damping control. The speed is shown in Figure 5(a) under the condition that the experimental prototype is not put into operation. Compare Figures 5(a) and 5(b) and it is clear that the amplitude of the curve under additional damping control is smaller and smaller, which presents good damping characteristic. This indicates that, under additional damping control, the damping characteristic of the system gets improved.

(2) Record the active power of DFIG and analyze the active power regulation ability of the wind turbine in the transient process. Figure 6(a) shows the real time active power of DFIG without damping control strategy when short circuit occurs; Figure 5(b) shows the active power of DFIG under damping control. Compare two figures and it is seen that when low frequency oscillation occurs in the system, DFIG rectifies its active power according to additional controlling signals sent by the experimental device to activate additional damping control. When the experimental device is not connected, DFIG outputs constant active power only according to the given value and does not provide any damping for the synchronous generator.

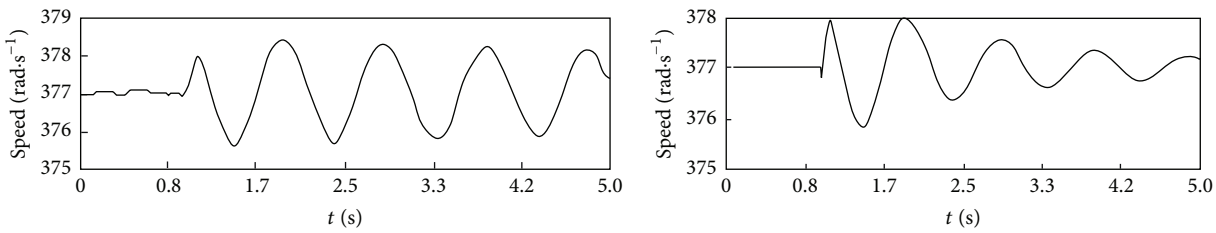
(3) Record the stator current and rotor current of the DFIG and observe the variation of current under additional damping control. Figures 6(a) and 6(b) show rotor current of DFIG when the experimental prototype is not put into operation and when it is. Figures 7(a) and 7(b) show the stator current of the DFIG when the experimental prototype is not put into operation and when it is. Compare these four situations and it is found that the rotor current does not show significant change when there is experimental prototype and when there is not. Although the stator current increases substantially when the experimental prototype is available, the current changes within safe range because it is not directly connected with other power electric devices. It indicates that additional damping control would not bring negative effect to the wind turbine and the system.

From the results it is seen that the active power of each wind turbine in the wind farms can be concentrated to be adjusted based on the integrated control platform of the wind farms and the technology of wide-field communication. When the system is working under weak damping, the action of additional damping control is excited to increase the damping of the system and prevents low frequency oscillation. At the same time, there is no negative influence on the wind turbine and the system, which proves that the method is available. In addition, wind-thermal ratio is an important factor influencing the damping effect. If the ratio is too small, the damping effect is limited. Compulsory damping may lead to overload of the rotor.

7.2. Experimental Analysis of Angle Stability

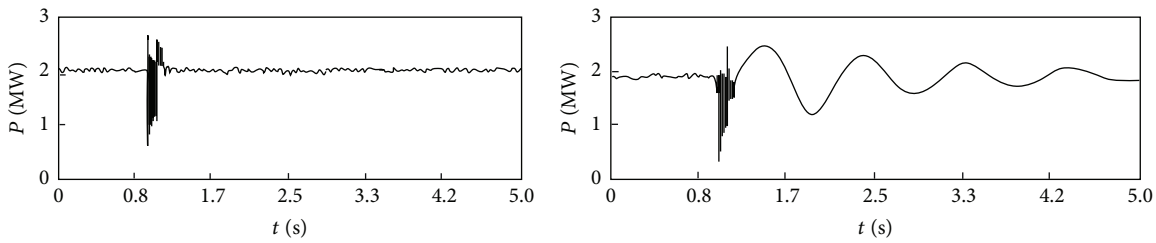
7.2.1. Fault Analysis of the Wind Farms. After the large-scale wind farms are connected to the power system, the overall stability of the power system declines greatly. An experiment is carried out to find out reasons. The wind farms in Figure 3 are replaced by a thermal power plant (SG1) of the same level. SG and SG1 constitute a large-scale thermal power plant. Fault simulation is compared between the single thermal power plant and the wind-thermal hybrid power system to see how the wind farms affect the stability of the power system. Assume that short circuit fault occurs on AC (B3-B4) for 0.3 s. Simulation results under two power modes are shown in Figures 8(a) and 8(b).

From Figure 8(a), it is seen that, after short circuit fault occurs, two thermal power plants have experienced an abrupt decline of power. The active power of SG and SG1 is reduced to 0.11 p.u., respectively. It suggests that the power decline is shared between two thermal power plants. From Figure 8(b), it is found that, after short circuit fault happens, the power of SG is decreased to 0.22 p.u. but that of the wind farms remains unchanged. It suggests that the power decline occurs only in the thermal power plant rather than both. This is not conductive to the thermal power plant and may result in its power-angle instability. Thus, a coordinated control strategy is proposed to address the power imbalance during the fault.



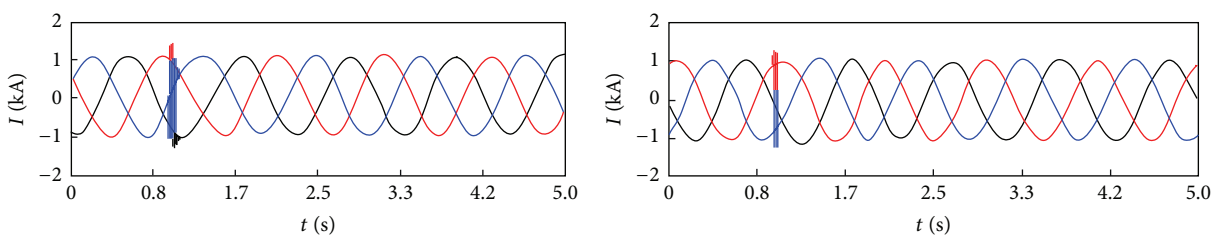
(a) Experimental prototype not put into operation (b) Experimental prototype put into operation

Figure 5: The speed of the synchronous generator.



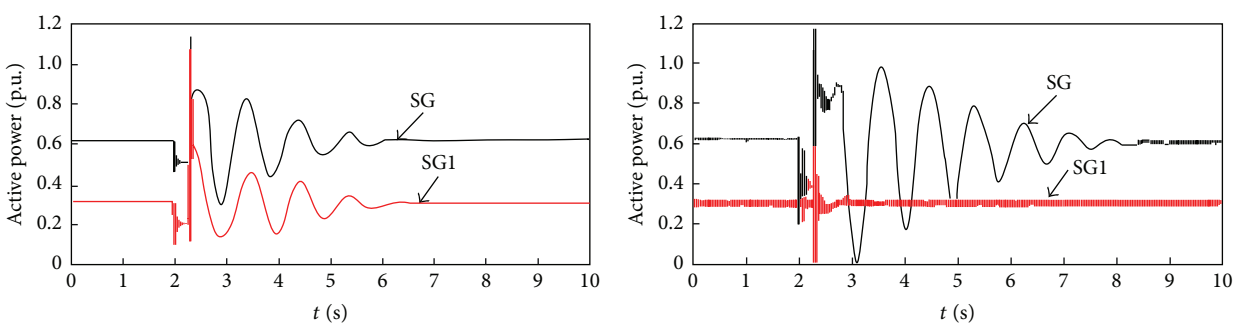
(a) Experimental prototype not put into operation (b) Experimental prototype put into operation

Figure 6: Rotor current of the doubly fed generator.



(a) Experimental prototype not put into operation (b) Experimental prototype put into operation

Figure 7: Stator current of DFIG.

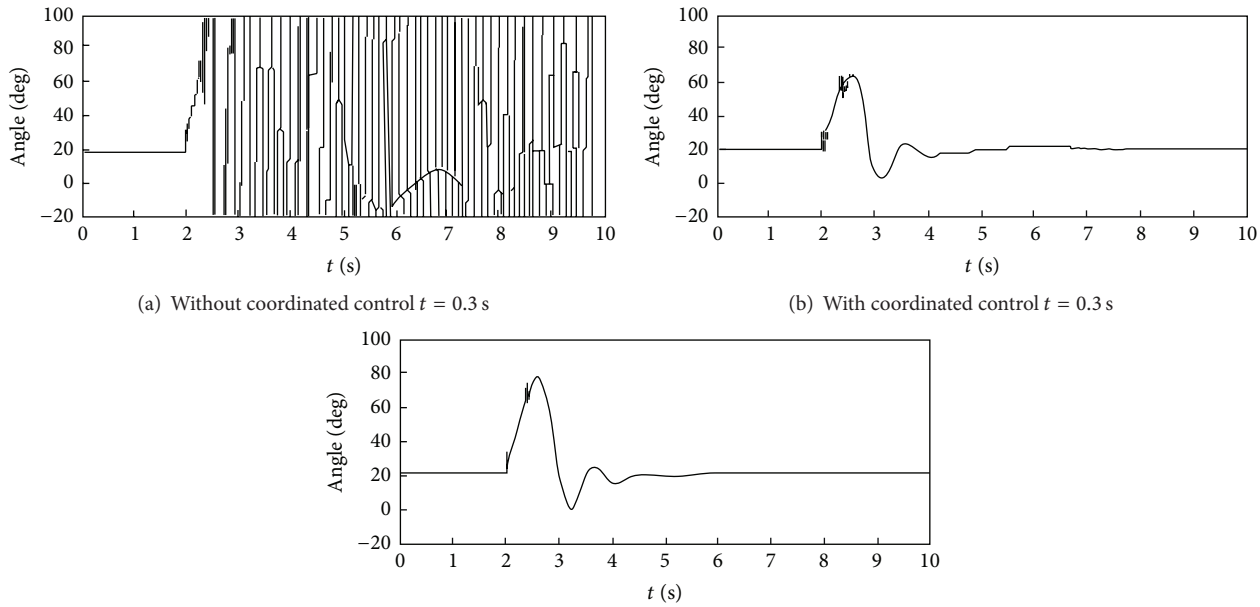


(a) Thermal power plant power (b) Wind-thermal hybrid power

Figure 8: Power comparisons under fault.

7.2.2. Angle Stability under Coordinated Control. In order to test the effectiveness of coordinated control strategy, a short circuit default simulation is carried out based on the experiment in Figure 4. Fault occurs on the AC circuit. Results are shown in Figure 9. Figure 9(a) shows the variation of power-angle of SG due to 0.3 s fault without the coordinated strategy. It shows that the stability of the synchronous generator is violated. Under the same fault condition, Figure 9(a) shows the variation of power-angle of SG due to 0.3 s fault with the coordinated strategy. The maximum oscillation of power-angle is 69°, which indicates that the synchronous generator is within its stability.

The duration of fault can reflect how bad the fault is. Thus, the duration is increased to 0.41 s. And the variation of power-angle is shown in Figure 9(c). It shows that the maximum oscillation of angle is 79°, which indicates that



(a) Without coordinated control $t = 0.3$ s

(b) With coordinated control $t = 0.3$ s

(c) With coordinated control $t = 0.41$ s

Figure 9: Analysis of the influence of coordinated control strategy.

the synchronous generator maintains its stability. Consequently, the coordinated control strategy proposed in this paper is proved to enhance the transient stability of the wind farms thermal hybrid power system.

8. Conclusion

Conclusions can be drawn as follows:

(1) The active power and the reactive power of DFIG have quick responses. Under the support of wide-field optical fiber network, doubly fed wind farms can be adjusted quickly.

(2) Quick adjustment of the active power of doubly fed wind farms is accomplished, which can restrain the low frequency oscillation of the thermal power unit and can supply the positive damping to the unit.

(3) Through coordinated control for wind farms and DC power transmission system, the transient stability of the hybrid power system can be enhanced.

(4) The experimental prototype and RTDS are connected for joint debugging. Results show that the experimental prototype can pose significant damping effect on thermal power plant and lower the risk of low frequency oscillation of the grid. The experimental prototype can also enhance the transient stability of hybrid AC/DC power transmission system.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors gratefully acknowledge the support of Science and Technology Project of Henan Province (112101210700), the Introduce Talents Research Fund of Guizhou University (2014-07), the National Natural Science Fund of China (51467003, 61403221), Open Foundation of Key Laboratory of Pulp and Paper Science and Technology of Ministry of Education of China under Grant nos. KF201419 and 08031347, and the Social Development Research Project of Guizhou Province (SY[2011]3081).

References

[1] S. Wang, D. Yu, and J. Yu, "A coordinated dispatching strategy for wind power rapid ramp events in power systems with high wind power penetration," *International Journal of Electrical Power & Energy Systems*, vol. 64, pp. 986–995, 2015.

[2] H. Li, S. Guo, L. Cui, J. Yan, J. Liu, and B. Wang, "Review of renewable energy industry in Beijing: development status, obstacles and proposals," *Renewable and Sustainable Energy Reviews*, vol. 43, pp. 711–725, 2015.

[3] A. A. Eldesouky, "Security constrained generation scheduling for grids incorporating wind, photovoltaic and thermal power," *Electric Power Systems Research*, vol. 116, pp. 284–292, 2014.

[4] Y. Zhang, F. Yao, H. H. C. Iu, T. Fernando, and H. Trinh, "Wind-thermal systems operation optimization considering emission problem," *International Journal of Electrical Power and Energy Systems*, vol. 65, no. 2, pp. 238–245, 2015.

[5] Y. Yubisui, S. Kobayashi, R. Amano, and T. Sugiura, "Effects of nonlinearity of magnetic force on passing through a critical speed of a rotor with a superconducting bearing," *IEEE Transactions on Applied Superconductivity*, vol. 23, no. 3, pp. 338–340, 2013.

[6] B. Basu, A. Staino, and M. Basu, "Role of flexible alternating current transmission systems devices in mitigating grid fault-induced vibration of wind turbines," *Wind Energy*, vol. 17, no. 7, pp. 1017–1033, 2014.

[7] A. Moharana, R. K. Varma, and R. Seethapathy, "SSR alleviation by STATCOM in induction-generator-based wind farm connected to series compensated line," *IEEE Transactions on Sustainable Energy*, vol. 5, no. 3, pp. 947–957, 2014.

[8] S.-Y. Yang, Y.-K. Wu, H.-J. Lin, and W.-J. Lee, "Integrated mechanical and electrical DFIG wind turbine model development," *IEEE Transactions on Industry Applications*, vol. 50, no. 3, pp. 2090–2102, 2014.

[9] H. Hinz, B. Zuber, and J. Kilz, "Development of a hybrid power generation system," in *Proceedings of the International Exhibition and Conference for Power Electronics, Intelligent Motion, Renewable Energy and Energy Management (PCIM Europe '14)*, pp. 651–658, IEEE, Nuremberg, Germany, May 2014.

[10] R. Zhu, Z. Chen, X. Wu, and F. Deng, "Virtual damping flux-based LVRT control for DFIG-based wind turbine," *IEEE Transactions on Energy Conversion*, vol. 23, no. 1, pp. 186–192, 2015.

[11] T. Surinkaew and I. Ngamroo, "Coordinated robust control of DFIG wind turbine and pss for stabilization of power oscillations considering system uncertainties," *IEEE Transactions on Sustainable Energy*, vol. 5, no. 3, pp. 823–833, 2014.

[12] S. Nouri, E. Babaei, and S. H. Hosseini, "A new AC/DC converter for the interconnections between wind farms and HVDC transmission lines," *Journal of Power Electronics*, vol. 14, no. 3, pp. 592–597, 2014.

[13] Y. Li, Z. Zhang, Y. Yang, Y. Li, H. Chen, and Z. Xu, "Coordinated control of wind farm and VSC-HVDC system using capacitor energy and kinetic energy to improve inertia level of power systems," *International Journal of Electrical Power & Energy Systems*, vol. 59, pp. 79–92, 2014.

[14] A. Rabiee, A. Soroudi, and A. Keane, "Information gap decision theory based OPF with HVDC connected wind farms," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 3396–3406, 2014.

[15] I. Erlich, C. Feltes, and F. Shewarega, "Enhanced voltage drop control by VSC-HVDC systems for improving wind farm fault ridethrough capability," *IEEE Transactions on Power Delivery*, vol. 29, no. 1, pp. 378–385, 2014.

[16] A. Rabiee and A. Soroudi, "Stochastic multiperiod OPF model of power systems with HVDC-connected intermittent wind power generation," *IEEE Transactions on Power Delivery*, vol. 29, no. 1, pp. 336–344, 2014.

[17] K. E. Okedu, S. M. Muyeen, R. Takahashi, and J. Tamura, "Effectiveness of current-controlled voltage source converter excited doubly fed induction generator for wind farm stabilization," *Electric Power Components and Systems*, vol. 40, no. 5, pp. 556–574, 2012.

[18] A. A. Elserougi, A. S. Abdel-Khalik, A. M. Massoud, and S. Ahmed, "A new protection scheme for HVDC converters against DC-side faults with current suppression capability," *IEEE Transactions on Power Delivery*, vol. 29, no. 4, pp. 1569–1577, 2014.

[19] J.-H. Lee, T.-H. Kim, G.-H. Kim, S. Heo, M. Park, and I.-K. Yu, "RTDS-based modeling of a 100 MW class wind farm applied to an integrated power control system," in *Proceedings of the IEEE Vehicle Power and Propulsion Conference (VPPC '12)*, pp. 1441–1443, Seoul, Republic of Korea, October 2012.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 836283, 14 pages
http://dx.doi.org/10.1155/2015/836283

Research Article

On Some Boundedness and Convergence Properties of a Class of Switching Maps in Probabilistic Metric Spaces with Applications to Switched Dynamic Systems

M. De la Sen[1] and A. Ibeas[2]

[1] *Institute of Research and Development of Processes, University of the Basque Country, Campus of Leioa (Biscay), P.O. Box 644, Bilbao, Barrio Sarriena, 48940 Leioa, Spain*
[2] *Department of Telecommunications and Systems Engineering, Universitat Autònoma de Barcelona (UAB), Bellaterra, Cerdanyola del Vallès, 08193 Barcelona, Spain*

Correspondence should be addressed to A. Ibeas; asier.ibeas@uab.cat

Received 10 June 2015; Accepted 9 September 2015

Academic Editor: Wenguang Yu

This paper investigates some boundedness and convergence properties of sequences which are generated iteratively through switched mappings defined on probabilistic metric spaces as well as conditions of existence and uniqueness of fixed points. Such switching mappings are built from a set of primary self-mappings selected through switching laws. The switching laws govern the switching process in between primary self-mappings when constructing the switching map. The primary self-mappings are not necessarily contractive but if at least one of them is contractive then there always exist switching maps which exhibit convergence properties and have a unique fixed point. If at least one of the self-mappings is nonexpansive or an appropriate combination given by the switching law is nonexpansive, then sequences are bounded although not convergent, in general. Some illustrative examples are also given.

1. Introduction

The background literature on fixed point theory and applications and associated convergence properties in metric spaces, Banach spaces, probabilistic metric spaces, Menger spaces, and some fuzzy-type versions is very abundant. See, for instance, [1–19] and the references therein. In particular, the theory focused on probabilistic metric spaces, including their specialization to Menger spaces, is also abundant. See, for instance, [1–4, 15, 16, 20] and the references therein. There are also studies in the graph framework for fixed point theory and problems of stability. See, for instance, [21, 22] and the references therein. On the other hand, fixed point theory has a wide range of applications, for instance, in the study of convergence of iterative schemes [17], in particular, of Mann and Jungck types or their many variants [18, 19], and in that of stability of dynamic systems and that of differential and difference equations. A particular class of real world applications refer to the stability of the so-called switched dynamic systems where a switching law assigns active parameterization for the dynamic system through time (or through an iterative discrete process) [23–27].

This paper investigates some boundedness and convergence properties of sequences which are generated through a class of switched mappings defined on probabilistic metric spaces, as well as conditions of existence and uniqueness of fixed points. The above switching mappings are defined via the selection as active of a set of primary self-mappings with the activation process governed by a "so-called" switching law. In this way, such switching laws govern the switching process in between primary self-mappings when constructing the switching map. The primary self-mappings are not necessarily contractive but if at least one of them is contractive then there always exist switching maps which exhibit convergence properties and have a unique fixed point. On the other hand, if at least one of the primary self-mappings is nonexpansive or an appropriate combination given by the switching law is nonexpansive, then sequences

are bounded although not convergent, in general. Some illustrative examples are also discussed. Section 2 introduces $C_{\rho k}$ and C_k classes of primary self-mappings in probabilistic metric spaces as well as associated upper- and lower-bounding constraints of the probability density of the built sequences. The above class allows the characterization of strict contractions as well as nonexpansive or expansive self-mappings in the probabilistic metric spaces. In parallel, some needed definitions are revisited while some preliminary results of convergence of sequences, Cauchy sequences, and boundedness of sequences in probabilistic metric spaces and in Menger spaces are obtained. Section 3 gives formalism in probabilistic metric spaces related to the switched maps defined via the activation of the primary self-mappings through switching laws. The obtained results for switched maps rely on boundedness and convergence of sequences in a probabilistic context.

2. On $C_{\rho k}$ and C_k Classes of Self-Mappings in Probabilistic Metric Spaces

Let us define a probabilistic distance $\mathbf{F} : X \times X \to \Delta_{\mathbf{F}}$, where X is a nonempty abstract set represented by $F_{x,y}$ for each $(x, y) \in X \times X$, where $\Delta_{\mathbf{F}}$ is a set of distribution functions. A distribution function $F \in \Delta_{\mathbf{F}}$ is a mapping $F : \mathbf{R} \to \mathbf{R}_{0+}$ which is nondecreasing and left-continuous with $\inf_{t\in\mathbf{R}} F(t) = 0$ and $\sup_{t\in\mathbf{R}} F(t) = 1$.

The ordered pair $(X, \mathbf{F})$ is a probabilistic metric (PM) space if for any $x, y, z \in X$ and all $t, s \in \mathbf{R}_+$ the following conditions hold [1]:

(1) $F_{x,y}(t) = H(t) \Longleftrightarrow x = y$,

where $H \in \Delta_{\mathbf{F}}$ is defined by $H(t) = \begin{cases} 0, & \text{if } t \le 0, \\ 1, & \text{if } t > 0; \end{cases}$

(2) $F_{x,y}(t) = F_{y,x}(t)$; (1)

(3) if $F_{x,y}(t) = 1$,

$F_{y,z}(s) = 1$

then $F_{x,z}(t + s) = 1$.

The triplet $(X, \mathbf{F}, \Delta)$ is a Menger space where $(X, \mathbf{F})$ is a PM-space and Δ is a triangular norm which satisfies the inequality $F_{x,z}(t + s) \ge \Delta(F_{x,y}(t), F_{y,z}(s))$, $\forall x, y, z \in X$, $\forall t, s \in \mathbf{R}_+$.

Note that $F_{x,y}(0) = F_{x,y}(t) = 0$ for $t \le 0$ and $F_{x,y}(t) = F_{x,y}(0^+) = 1$ for $t > 0$ if $x = y$ since $H \in \Delta_{\mathbf{F}}$ is nondecreasing and left-continuous. Note also that every metric space (X, d) can be realized as a PM-space by taking $\mathbf{F} : X \times X \to \Delta_{\mathbf{F}}$ being defined by $F_{x,y}(t) = H(t - d(x, y))$ for all $x, y \in X$ [1–4]. In the following, D_+ is the space of all mappings $F : \mathbf{R} \to [0, 1]$ which are left-continuous and nondecreasing with $F(0) = 0$ and $\ell^- F(+\infty) = 1$. The space D_+ is partially ordered by the usual pointwise ordering of functions; namely, $F \le G$ if and only if $F(t) \le G(t)$, $\forall t \in \mathbf{R}$, and its maximal element is the distribution $H(t)$ [4].

Definition 1. Let $(X, \mathbf{F})$ be a PM-space. A mapping $T : X \to X$ is said to be of C_k-class for some function $k : X \times X \to \mathbf{R}_+$ if

$$F_{Tx,Ty}(t) \ge F_{x,y}\left(k^{-1}(x, y)\, t\right); \quad \forall x, y \in X, \ \forall t \in \mathbf{R}_+. \quad (2)$$

Definition 2. Let $(X, \mathbf{F})$ be PM-space. A mapping $T : X \to X$ is said to be of $C_{\rho k}$-class for some functions $\rho, k : X \times X \to \mathbf{R}_+$ if

$$F_{x,y}\left(\rho^{-1}(x, y)\, t\right) \ge F_{Tx,Ty}(t) \ge F_{x,y}\left(k^{-1}(x, y)\, t\right); \quad \forall x, y \in X, \ \forall t \in \mathbf{R}_+, \quad (3)$$

where the functions $\rho, k : X \times X \to \mathbf{R}_+$ satisfy $\rho(x, y) \le k(x, y)$, $\forall x, y \in X$.

Note that if $T : X \to X$ is of $C_{\rho k}$-class, then it is also of C_k-class. Note also that $T : X \to X$ is nonexpansive if it is of C_k-class with $\sup_{x,y\in X} k(x, y) \le 1$ and, in particular, a probabilistic strict contraction if it is of C_k-class with $\sup_{x,y\in X} k(x, y) < 1$. Also, if $T : X \to X$ is of $C_{\rho k}$-class with $\sup_{x,y\in X} k(x, y) \le 1$ ($\sup_{x,y\in X} k(x, y) < 1$), then it is nonexpansive (probabilistic strictly contractive) [1–4]. If $T : X \to X$ is of $C_{\rho k}$-class with $1 < \inf_{x,y\in X} \rho(x, y) \le \inf_{x,y\in X} k(x, y)$, then it is expansive [1–4]. If there is some $\rho : X \times X \to \mathbf{R}_+$ with $\inf_{x,y\in X} \rho(x, y) > 1$ such that

$$F_{x,y}\left(\rho^{-1}(x, y)\, t\right) \ge F_{Tx,Ty}(t); \quad \forall x, y \in X, \ \forall t \in \mathbf{R}_+, \quad (4)$$

then $T : X \to X$ is expansive (even if $T : X \to X$ is not of $C_{\rho k}$-class for some $k : X \times X \to \mathbf{R}_+$ subject to $1 < \inf_{x,y\in X} \rho(x, y) \le \inf_{x,y\in X} k(x, y)$).

The following technical result follows.

Lemma 3. *The following properties hold:*

(i) *Let $(X, \mathbf{F})$ be a PM-space and let $T : X \to X$ be a mapping of $C_{\rho k}$-class. Consider the sequences $\{x_n\} \subseteq X$ and $\{y_n\} \subseteq X$ built by $x_{n+1} = Tx_n$, $y_{n+1} = Ty_n$, $\forall n \in \mathbf{Z}_{0+}$ with $x_0 = x$, $y_0 = y$ for some given $x, y \in X$. Then,*

$$F_{x,y}\left(\prod_{i=0}^{n-1}\left[\rho_i^{-1}(x, y)\right] t\right) \ge F_{T^n x, T^n y}(t) \ge F_{x,y}\left(\prod_{i=0}^{n-1}\left[k_i^{-1}(x, y)\right] t\right), \quad (5)$$

where $k_n(x, y) = k(T^n x, T^n y)$, $\rho_n(x, y) = \rho(T^n x, T^n y)$, $\forall n \in \mathbf{Z}_{0+}$.

(ii) *If $T : X \to X$ is of C_k-class, then $F_{T^n x, T^n y}(t) \ge F_{x,y}(\prod_{i=0}^{n-1}[k_i^{-1}(x, y)]t)$, $\forall n \in \mathbf{Z}_{0+}$.*

(iii) *If $T : X \to X$ is a mapping of either C_k-class or $C_{\rho k}$-class with $\lim_{n\to\infty}\prod_{i=0}^{n}[k_i(x, y)] = 0$ for the given $x, y \in X$, then $\lim_{n\to\infty} F_{T^n x, T^n y}(t) = 1$, $\forall t \in \mathbf{R}_+$.*

(iv) *If $T : X \to X$ is a mapping of $C_{\rho k}$-class with $\beta_n = \prod_{i=0}^{n-1}[\rho_i^{-1}(x,y)]$, $\alpha_n = \prod_{i=0}^{n-1}[k_i^{-1}(x,y)]$, $\forall n \in \mathbf{Z}_{0+}$, and*

$$\alpha = \alpha(x,y) = \liminf_{n\to\infty} \alpha_n, \qquad \beta = \beta(x,y) = \limsup_{n\to\infty} \beta_n \tag{6}$$

are in $cl\mathbf{R}_{0+} = \mathbf{R}_{0+} \cup \{+\infty\}$ (i.e., $cl\mathbf{R}_{0+}$ is the closure of $\mathbf{R}_{0+}$, i.e., the extended nonnegative real semiline) for the given $x, y \in X$, then

$$F_{x,y}(\beta_n t) \geq F_{T^n x, T^n y}(t) \geq F_{x,y}(\alpha_n t); \quad \forall n \in \mathbf{Z}_{0+}\ \forall t \in \mathbf{R}_+, \tag{7a}$$

$$F_{x,y}(\beta t) \geq \limsup_{n\to\infty} F_{T^n x, T^n y}(t) \geq \liminf_{n\to\infty} F_{T^n x, T^n y}(t) \geq F_{x,y}(\alpha t); \quad \forall t \in \mathbf{R}_+. \tag{7b}$$

If $T : X \to X$ is a mapping of C_k-class, then $\liminf_{n\to\infty} F_{T^n x, T^n y}(t) \geq F_{x,y}(\alpha t)$, $\forall t \in \mathbf{R}_+$.

Proof. It follows recursively from (3) with $x_{n+1} = Tx_n$, $y_{n+1} = Ty_n$, $\forall n \in \mathbf{Z}_{0+}$, with $x_0 = x$, $y_0 = y$ for the given $x, y \in X$ that

$$\begin{aligned}
&F_{x,y}\left(\rho_0^{-1}(x,y)\,t\right) \geq F_{Tx,Ty}(t) \geq F_{x,y}\left(k_0^{-1}(x,y)\,t\right),\\
&F_{x,y}\left(\rho_0^{-1}(x,y)\,\rho_1^{-1}(x,y)\,t\right) \geq F_{Tx,Ty}\left(\rho_1^{-1}(x,y)\,t\right)\\
&\quad \geq F_{T^2x,T^2y}(t) \geq F_{Tx,Ty}\left(k_1^{-1}(x,y)\,t\right)\\
&\quad \geq F_{x,y}\left(k_0^{-1}(x,y)\,k_1^{-1}(x,y)\,t\right),\\
&\qquad \vdots\\
&F_{x,y}\left(\prod_{i=0}^{n-1}\left[\rho_i^{-1}(x,y)\right]t\right) \geq F_{T^n x, T^n y}(t)\\
&\quad \geq F_{x,y}\left(\prod_{i=0}^{n-1}\left[k_i^{-1}(x,y)\right]t\right);\\
&\qquad \forall x, y \in X,\ \forall t \in \mathbf{R}_+,\ \forall n \in \mathbf{Z}_{0+}.
\end{aligned} \tag{8}$$

Property (i) has been proved and the proof of Property (ii) follows directly by just using the lower-bounding part of the recursion. Property (iii) follows since $\lim_{n\to\infty}\prod_{i=0}^{n-1}[k_i^{-1}(x,y)] = +\infty$; then, $\lim_{n\to\infty}(\prod_{i=0}^{n-1}[k_i^{-1}(x,y)])t = +\infty$, $\forall t \in \mathbf{R}_+$, and the conditions that $F_{x,y}(t)$ is nondecreasing in the argument t and $\sup_{t\in\mathbf{R}_+} F_{x,y}(t) = \limsup_{t\to\infty} F_{x,y}(t) = 1$ lead from (8) to the existence of the limit $\lim_{n\to\infty} F_{T^n x, T^n y}(t) = 1$, $\forall t \in \mathbf{R}_+$. Property (iv) is proved closely to Property (iii) by first getting (7a) and (7b) directly from the definitions of β_n, α_n, β, and α, $\forall n \in \mathbf{Z}_{0+}$. □

The subsequent example illustrates that Lemma 3 is useful for the characterization of probabilities which can be less than one (i.e., the probabilistic certainty) through lower-bounds and upper-bounds in probabilistic metric spaces.

Example 4. Let us consider the metric space (X, d) with $F : X \times X \to \Delta_F$ being defined by $F_{x,y}(t) = H(t - d(x,y))$ for all $x, y \in X$ for the distribution function $H_{ab} \in \Delta_F$ defined by:

$$H_{ab}(t) = \begin{cases} a(t), & \text{if } t \leq 0,\\ b(t), & \text{if } t > 0 \end{cases} \tag{9}$$

for some left-continuous nondecreasing functions $a, b : \mathbf{R}_{0+} \to [0,1]$ with

$$\begin{aligned}
&b(x,y,t) \geq a(x,y,t) = a(x,y,-t); \quad \forall t \in \mathbf{R},\\
&\lim_{t\to-\infty} a(x,y,t) = 0,\\
&\lim_{t\to-\infty} b(x,y,t) = 1.
\end{aligned} \tag{10}$$

Assume also that $a : \mathbf{R}_{0+} \to [0,1]$ is everywhere lower-semicontinuous and $b : \mathbf{R}_{0+} \to [0,1]$ is everywhere upper-semicontinuous. Then,

$$H_{ab}(\alpha t - d(x,y)) = \begin{cases} a(\alpha t - d(x,y)), & \text{if } t \leq \dfrac{d(x,y)}{\alpha},\\ b(\alpha t - d(x,y)), & \text{if } t > \dfrac{d(x,y)}{\alpha}, \end{cases} \quad \forall x, y \in X \tag{11}$$

with $a(x,y,0^-) = a(x,y,0^+) = b(x,y,0^-) = 0$, $\lim_{t\to-\infty} a(x,y,t) = 0$, and $\lim_{t\to+\infty} b(x,y,t) = 1$, $\forall x, y \in X$. Assume following Lemma 3(iv) that $\beta = \beta(x,y) = \limsup_{n\to\infty}\beta_n$ and $\alpha = \alpha(x,y) = \liminf_{n\to\infty}\alpha_n$ with $\beta_n = \beta_n(x,y) = \prod_{i=0}^{n-1}[\rho_i^{-1}(x,y)]$, $\alpha_n = \alpha_n(x,y) = \prod_{i=0}^{n-1}[k_i^{-1}(x,y)]$, $\forall n \in \mathbf{Z}_{0+}$. Note that α, β, α_n, and β_n are allowed to be dependent on x, y. Then, if $T : X \to X$ is a mapping of $C_{\rho k}$-class so that (7a) and (7b) of Lemma 3 hold, one gets for any given $x, y \in X$

$$\begin{aligned}
b(\beta_n t - d(x,y)) &= F_{x,y}(\beta_n t)\\
&= H_{ab}(\beta_n t - d(x,y))\\
&\geq \limsup_{n\to\infty} F_{T^n x, T^n y}(t)\\
&= \limsup_{n\to\infty} H_{ab}(t - d(T^n x, T^n y))\\
&\geq \liminf_{n\to\infty} F_{T^n x, T^n y}(t)\\
&= \liminf_{n\to\infty} H_{ab}(t - d(T^n x, T^n y))\\
&\geq F_{x,y}(\alpha_n t) = b(\alpha_n t - d(x,y))\\
&= H_{ab}(\alpha_n t - d(x,y));\\
&\qquad \forall n \in \mathbf{Z}_{0+},
\end{aligned} \tag{12a}$$

$$\begin{aligned} b(\beta t - d(x,y)) &= F_{x,y}(\beta t) = H_{ab}(\beta t - d(x,y)) \\ &\geq \limsup_{n\to\infty} F_{T^n x, T^n y}(t) \\ &= \limsup_{n\to\infty} H_{ab}(t - d(T^n x, T^n y)) \\ &\geq \liminf_{n\to\infty} F_{T^n x, T^n y}(t) \\ &= \liminf_{n\to\infty} H_{ab}(t - d(T^n x, T^n y)) \\ &\geq F_{x,y}(\alpha t) = b(\alpha t - d(x,y)) \\ &= H_{ab}(\alpha t - d(x,y)) \end{aligned} \tag{12b}$$

from (7a) and (7b). Thus, one gets for any given $x, y \in X$ the following:

(a) If $t_n > d(x,y)/\alpha_n$ for some given $n \in \mathbf{Z}_{0+}$, then, since $t_n > d(x,y)/\beta_n$ as well, one gets

$$\begin{aligned} b(\beta_n t_n - d(x,y)) &\geq F_{T^n x, T^n y}(t_n) \\ &\geq b(\alpha_n t_n - d(x,y)) \end{aligned} \tag{13a}$$

and if $t > d(x,y)/\alpha$ since $t > d(x,y)/\beta$, then

$$\begin{aligned} b(\beta t - d(x,y)) &\geq \limsup_{n\to\infty} F_{T^n x, T^n y}(t) \\ &\geq \liminf_{n\to\infty} F_{T^n x, T^n y}(t) \\ &\geq b(\alpha t - d(x,y)) \end{aligned} \tag{13b}$$

and $\exists \lim_{t\to+\infty} \lim_{n\to\infty} F_{T^n x, T^n y}(t) = 1$ if $\alpha > 0$ since the above superior and inferior limits equalize unity.

(b) If $d(x,y)/\beta_n < t \leq d(x,y)/\alpha_n$, then

$$\begin{aligned} H_{ab}(\alpha_n t - d(x,y)) &= a(\alpha_n t - d(x,y)) \\ &\leq \liminf_{n\to\infty} F_{T^n x, T^n y}(t) \\ &\leq \limsup_{n\to\infty} F_{T^n x, T^n y}(t) \\ &\leq H_{ab}(\beta_n t - d(x,y)) \\ &= b(\beta_n t - d(x,y)). \end{aligned} \tag{14}$$

If, furthermore, $\beta = \alpha > 0$, then $\exists \lim_{t\to+\infty} \lim_{n\to\infty} F_{T^n x, T^n y}(t) = 1$. If, in addition, $\alpha = +\infty$, then $\beta = +\infty$, $a(t) = 0$, and $b(t) = 1$, $\forall t \in \mathbf{R}_+$; then $\exists \lim_{n\to\infty} F_{T^n x, T^n y}(t) = 1$, $\forall t \in \mathbf{R}_+$, which is the basic convergence suitable property in probabilistic metric spaces for probabilistic strictly contractive mappings in the existing literature. Note that this case includes the case under Lemma 3(iii) when $\lim_{n\to\infty} \prod_{i=0}^{n} [k_i(x,y)] = 0$ leading to

$$\begin{aligned} \lim_{n\to\infty} F_{T^n x, T^n y}(t) &= \lim_{n\to\infty} H_{ab}(t - d(T^n x, T^n y)) \\ &\geq F_{x,y}(+\infty) = H(+\infty) = 1; \end{aligned} \tag{15}$$

that is, $\lim_{n\to\infty} F_{T^n x, T^n y}(t) = 1$, $\forall x, y \in X$, $\forall t \in \mathbf{R}_+$, or, in other words, for any distance $d(x,y)$ from a given $x \in X$ to a given $y \in X$, $\lim_{n\to\infty} d(T^n x, T^n y) = 0$.

(c) If $t \leq d(x,y)/\beta_n$, then $\liminf_{n\to\infty} F_{T^n x, T^n y}(t) \geq a(\beta_n t - d(x,y))$.

(d) Now, assume that $T : X \to X$ is a mapping of $C_{\rho k}$-class with $1 < \inf_{x,y\in X} \rho(x,y) \leq \inf_{x,y\in X} k(x,y)$ and then $\beta = \alpha = 0$; that is, the mapping is expansive. Then, if $t \to +\infty$ implying that $t > d(x,y)/\beta$ (and also $t > d(x,y)/\beta$ since $\alpha = \beta = 0$), one concludes from (13b) that

$$\begin{aligned} &\inf_{t > d(x,y)/\beta_n} F_{T^n x, T^n y}(t) = 1, \\ &\sup_{t \leq d(x,y)/\beta_n} F_{T^n x, T^n y}(t) < 1; \\ &\qquad\qquad \forall x, y (\neq x) \in X, \\ &\lim_{n\to+\infty} \inf_{t > d(x,y)/\beta_n} F_{T^n x, T^n y}(t) = 1, \\ &\lim_{n\to+\infty} \sup_{t \leq d(x,y)/\beta_n} F_{T^n x, T^n y}(t) < 1; \\ &\qquad\qquad \forall x, y (\neq x) \in X \end{aligned} \tag{16}$$

since $\{\alpha_n\} \to 0$ and $\{\beta_n\} \to 0$. The constraints (13a) still hold for each $n \in \mathbf{Z}_{0+}$ such that $t_n > d(x,y)/\alpha_n$ but the sequence $\{t_n\}$ diverges to $+\infty$ while $\{\alpha_n\} \to 0$ and $\{\beta_n\} \to 0$.

Example 5. Assume that $\alpha = \beta = \alpha_n = \beta_n = 1$, $\forall n \in \mathbf{Z}_{0+}$, independent of $x, y \in X$. Then, $T : X \to X$ is of C_{11}-class and nonexpansive but also probabilistic noncontractive. If $t > d(x,y)$, then one gets from (13a) and (13b)

$$\begin{aligned} &F_{T^n x, T^n y}(t) = F_{x,y}(t) = b(t - d(x,y)), \\ &\lim_{t\to+\infty} \lim_{n\to\infty} F_{T^n x, T^n y}(t) = \lim_{t\to+\infty} b(t - d(x,y)) \\ &\quad = \lim_{t\to+\infty} H_{ab}(t) = 1; \\ &\qquad\qquad \forall x, y \in X. \end{aligned} \tag{17}$$

Assume instead that $\alpha_{n_k} = \beta_{n_k} = 1$ for some sequence $\{n_k\} \subseteq \mathbf{Z}_{0+}$, $\forall k \in \mathbf{Z}_{0+}$. Then, if $t_{n_k} > d(x,y)$, one has for any $x, y \in X$ that

$$\begin{aligned} &F_{T^n x, T^n y}(t_{n_k}) = b(t_{n_k} - d(x,y)) \\ &\qquad\qquad \text{if } t_{n_k} > d(x,y); \\ &b(\alpha_n t - d(x,y)) \leq F_{T^n x, T^n y}(t) \leq b(\beta_n t - d(x,y)), \\ &\forall t \in (t_{n_k}, t_{n_{k+1}}],\ \forall n \in (n_k, n_{k+1}] \text{ if } t > \frac{d(x,y)}{\beta_n}, \end{aligned} \tag{18}$$

which simplifies as $b(\alpha_n t - d(x,y)) \le F_{T^n x, T^n y}(t) \le b(\beta_n t - d(x,y))$, $\forall n \in \mathbf{Z}_{0+}$, if $t > d(x,y)/\beta_n$:

$$\begin{aligned}\lim_{t\to+\infty}\lim_{k\to\infty} F_{T^{n_k}x,T^{n_k}y}(t) &= \lim_{k\to\infty} b\left(t_{n_k} - d(x,y)\right)\\ &= \lim_{t\to+\infty} b(t - d(x,y)) = \lim_{k\to\infty} H_{ab}\left(t_{n_k}\right) = 1.\end{aligned} \tag{19}$$

From Lemma 3(iii), one gets directly the subsequent result.

Proposition 6. *Let $(X,\mathbf{F})$ be a PM-space and let $T : X \to X$ be a mapping of either C_k-class or $C_{\rho k}$-class and there is a strictly increasing sequence of nonnegative integers $\{n_k\}$ fulfilling $\lim_{k\to\infty}\sup(n_{k+1} - n_k) < +\infty$ such that, for some given $x, y \in X$, $\prod_{i=n_k}^{n_k+n_{k+1}}[k_i(x,y)] < 1$, $\forall k \in \mathbf{Z}_{0+}$; then $\lim_{n\to\infty} F_{T^n x, T^n y}(t) = 1$, $\forall t \in \mathbf{R}_+$.*

Proof. It follows from Lemma 3 that if $\prod_{i=n_k}^{n_k+n_{k+1}}[k_i(x,y)] < 1$, $\forall k \in \mathbf{Z}_{0+}$, then $\lim_{k\to\infty}\prod_{i=0}^{\sum_{i=0}^{k} n_i}[k_i(x,y)] = 0$, $\lim_{k\to\infty}\prod_{i=0}^{\sum_{i=0}^{k} n_i+n}[k_i(x,y)] = 0$, $\lim_{k\to\infty}\prod_{i=0}^{\sum_{i=0}^{k} n_i+n}[k_i^{-1}(x,y)] = \infty$, $\forall n \in (n_k, n_{k+1}]$, $\forall k \in \mathbf{Z}_{0+}$, and then $\lim_{n\to\infty}\prod_{i=0}^{n}[k_i^{-1}(x,y)] = \infty$. □

Note that Proposition 6 includes as a particular case that of probabilistic strict contractions $T : X \to X$ which are then mappings of C_k-class with $0 < \alpha^{-1} = \sup_{x,y\in X} k(x,y) < 1$.

Definition 7 (see [2]). Let $(X,\mathbf{F})$ be a PM-space and A a nonempty subset of X. The probabilistic diameter of A is a mapping $D_A : \mathbf{R}_{0+} \to [0,1]$ defined by $D_A(z) = \sup_{t<z}\inf_{x,y\in A} F_{x,y}(t)$.

Definition 8 (see [2, 4]). Let $(X,\mathbf{F})$ be a PM-space and A a nonempty subset of X. The nonempty set A is said to be probabilistically bounded if $\sup_{z\in\mathbf{R}_{0+}} D_A(z) = 1$, that is, if the supremum of its probabilistic diameter $D_A \in D_+$.

We can define the set unboundedness as the concept opposite to Definition 8 as follows.

Definition 9 (see [2, 4]). Let $(X,\mathbf{F})$ be a PM-space and A a nonempty subset of X. The nonempty set A is said to be probabilistically unbounded if $\sup_{z\in\mathbf{R}_{0+}} D_A(z) < 1$, that is, if $D_A \notin D_+$.

The boundedness and unboundedness of sequences $\{x_n\} \subseteq X$ can be easily defined as supported by Definitions 8 and 9 as follows.

Definition 10 (see [2, 4]). Let $(X,\mathbf{F})$ be a PM-space. The sequence $\{x_n\} \subseteq X$ is probabilistically bounded if $\sup_{z\in\mathbf{R}_{0+}} \sup_{t<z}\inf_{n,m\in\mathbf{Z}_{0+}}\inf_{x_n,x_m\in X} F_{x_n,x_m}(t) = 1$.

Definition 11 (see [1]). Let $(X,\mathbf{F})$ be a PM-space. Then, the sequence $\{x_n\} \subseteq X$ is

(1) probabilistically convergent to a point $x \in X$, denoted by $\{x_n\} \to x$, if for every $\varepsilon \in \mathbf{R}_+$ and $\lambda \in (0,1)$ there exists some $N = N(\varepsilon,\lambda) \in \mathbf{Z}_{0+}$ such that

$$F_{x_n,x}(\varepsilon) > 1 - \lambda; \quad \forall n\left(\in \mathbf{Z}_{0+}\right) \ge N; \tag{20}$$

(2) Cauchy if for every $\varepsilon \in \mathbf{R}_+$ and $\lambda \in (0,1)$ there exists some $N = N(\varepsilon,\lambda) \in \mathbf{Z}_{0+}$ such that

$$F_{x_n,x_m}(\varepsilon) > 1 - \lambda; \quad \forall n,m\left(\in \mathbf{Z}_{0+}\right) \ge N. \tag{21}$$

A PM-space $(X,\mathbf{F})$ is complete if every Cauchy sequence is probabilistically convergent.

Proposition 12. *Let $(X,\mathbf{F})$ be a PM-space. Then,*

(1) *$\{x_n\}(\to x) \subseteq X$ for some $x \in X$ if and only if the following limit exists: $\lim_{\varepsilon\to 0^+}\lim_{n\to\infty} F_{x_n,x}(\varepsilon) = 1$;*

(2) *$\{x_n\} \subseteq X$ is a Cauchy sequence if and only if $\lim_{\varepsilon\to 0^+}\lim_{n\to\infty} F_{x_n,x_{n+m}}(\varepsilon) = 1$, $\forall m \in \mathbf{Z}_{0+}$.*

Proof. If $\{x_n\} \to x$, then there exists some $N = N(\varepsilon,\lambda) \in \mathbf{Z}_{0+}$ such that $F_{x_n,x}(\varepsilon) > 1 - \lambda$, $\forall n(\in \mathbf{Z}_{0+}) \ge N$, for every $\varepsilon \in \mathbf{R}_+$ and $\lambda \in (0,1)$. Thus, since $0 \le F_{x_n,x}(\varepsilon) \le 1$, $\forall n \in \mathbf{Z}_{0+}$, then, by taking $\lambda \to 0^+$, one gets $\lim_{n\to\infty} F_{x_n,x}(0^+) = \lim_{\varepsilon\to 0^+}\lim_{n\to\infty} F_{x_n,x}(\varepsilon) = 1$.

Conversely, if $\lim_{n\to\infty} F_{x_n,x}(0^+) = 1$, then for every $\varepsilon \in \mathbf{R}_+$ and $\lambda \in (0,1)$ there exists some $N = N(\varepsilon,\lambda) \in \mathbf{Z}_{0+}$ such that $F_{x_n,x}(\varepsilon) > 1-\lambda$, $\forall n(\in \mathbf{Z}_{0+}) \ge N$; thus $\{x_n\} \to x$. Assume that this is not true. Thus, there is some subsequence $\{x_{n_k}\} \subseteq \{x_n\}$ such that $F_{x_{n_k},x}(\varepsilon) \le 1-\lambda$ for some $\varepsilon \in \mathbf{R}_+$ and $\lambda \in (0,1)$ while $\lim_{n\to\infty} F_{x_n,x}(\varepsilon) = \lim_{n\to\infty} F_{x_n,x}(0^+) = 1$ for any $\varepsilon \in \mathbf{R}_+$ since $F_{x_n,x}(\varepsilon)$ is nondecreasing in the argument ε and one gets the following contradiction for some $\lambda \in (0,1)$:

$$\begin{aligned}1 - \lambda \ge \lim_{k\to\infty} F_{x_{n_k},x}(\varepsilon) &= \lim_{n\to\infty} F_{x_n,x}(\varepsilon)\\ &= \lim_{n\to\infty} F_{x_n,x}\left(0^+\right) = 1.\end{aligned} \tag{22}$$

Proposition 12(1) has been proved. The proof of Proposition 12(2) is very close and it is omitted. □

Proposition 13. *Let $(X,\mathbf{F})$ be a PM-space. Then, the sequence $\{x_n\} \subseteq X$ is probabilistically bounded if and only if $D_{aM\{x_n\}} = 1$, where $D_{aM\{x_n\}} = \sup_{z\in\mathbf{R}_{0+}} D_{a\{x_n\}}(z)$ with $D_{a\{x_n\}}(z) = \sup_{t<z}\inf_{x_n\in X} F_{a,x_n}(t)$ for some $a \in X$, that is, if and only if $D_{aM\{x_n\}} \in D_+$ for some $a \in X$.*

Proof. If X is bounded, then the result is direct for any sequence $S = \{x_n\} \subseteq X$. Assume that X is not bounded and proceed by contradiction by assuming that $S \subseteq X$ is probabilistically bounded and $D_{aMS} \notin D_+$ for some $a \in X$. On the other hand, since $S \subseteq X$ is probabilistically bounded, then, for all $x_k, x_m \in S$, there is $t_0 = t_0(x_k, x_m) \in \mathbf{R}_{0+}$ such that $F_{x_k,x_m}(t) = 1$, $\forall t(\in \mathbf{R}_{0+}) > t_0$. Since $D_{aMS} \notin D_+$, there is $x_k \in S$ such that $F_{a,x_k}(t) < 1$, $\forall t(\in \mathbf{R}_{0+}) > t_{01}$, and some

$t_{01} = t_{01}(x_k, a) \in \mathbf{R}_{0+}$. This implies from the contrapositive equivalent logic proposition to the third property of (1) of $(X, \mathbf{F})$ being a PM-space that either $F_{x_k,x_m}(t/2) < 1$, and then the sequence $S \subseteq X$ is not probabilistically bounded, $\forall t(\in \mathbf{R}_{0+}) > t_{01}$ (a contradiction), or $F_{a,x_m}(t/2) < 1$, $\forall t(\in \mathbf{R}_{0+}) > t_{01}$, for any given $x_m \in S$, and then $\limsup_{t\to+\infty} F_{a,x_m}(t/2) < 1$ for some fixed $a, x_m \in X$. Now, assume that, for all $a \in X$ such that $a \notin S$, $\limsup_{t\to+\infty} F_{a,x_m}(t/2) < 1$. Thus, the subset $A_a = \{a, x_m\}$ of X, $\forall a \in X$, is unbounded since its probabilistic diameter is less than one; that is, $D_{A_a} \notin D_+$, $\forall a \in X$, and then the sequence S is probabilistically unbounded, again a contradiction. It has been proved that if $S \subset X$ is bounded, then $D_{aMS} \in D_+$ for some $a \in X \cap \overline{S}$, where $\overline{S}$ is the complementary to S in X. It remains to prove that if $D_{aMS} \in D_+$ for some $a \in X \cap \overline{S}$, then $S \subset X$ is probabilistically bounded. Since $D_{aMS} \in D_+$, then $F_{a,x_k}(t) = 1$, $\forall t(\in \mathbf{R}_{0+}) > t_0$, for some $t_0 = t_0(a, x_k) \in \mathbf{R}_{0+}$ and all $x_k \in S$. It follows from the third property of (1) that $F_{x_n,x_m}(t) = 1$, $\forall t(\in \mathbf{R}_{0+}) > 2t_0$, $\forall x_n, x_m \in S$. Thus, $S \subset X$ is probabilistically bounded. □

3. Switched Maps Defined by $C_{\rho k}$ and C_k Classes of Primary Self-Mappings and a Class of Dynamic Systems

Switching processes are a very important tool in some applications of discrete-time and continuous-time dynamic systems. The basic idea is how to switch in-between alternative parameterizations of a system by using either "ad hoc" or even arbitrary switching laws while keeping or improving essential suitable properties like global or asymptotic stability or convergence to the equilibrium points. See [23–26] and some references therein. The formalism can also rely on the definitions of iteration-dependent maps in iterative schemes of Mann or Jungck type or its generalizations so as to get appropriate convergence properties [18, 19, 23]. Note that a switching process in an iterative scheme can be interpreted as the choice under a switching rule of certain primary self-maps from an available collection of them at certain iteration points; that is, the iterative scheme or the solution equation of a dynamic system is being governed by a switching rule [28]. Based on the above elementary idea, this section relies on switching maps built with a prefixed number of either $C_{\rho k}$-class or C_k-class, self-mappings on PM-spaces subject to switching rules which select the new selected mapping and the points at which such new switching occurs. For exposition simplicity, it is assumed that $C_{\rho k}$-class, or C_k-class, self-mappings are characterized by constants instead of functions in Definitions 1 and 2.

Let $(X, \mathbf{F})$ be PM-space and let $T_i : X \to X$ be a set of (primary) self-mappings of $C_{\rho k}$-class for some constants $\rho_i(\leq k_i)$, $k_i \in \mathbf{R}_+$ for $i \in \overline{q} = \{1, 2, \ldots, q\}$. A switching map $T = T_{\sigma_n}(x)$ from $\mathbf{Z}_{0+} \times X$ to X with respect to the switching law $\sigma : \mathbf{Z}_{0+} \times X \to \overline{q}$ generates a sequence

$$x_{n+1} = Tx_n = T_{\sigma_n}(x_n)\, x_n = T_i x_n, \quad \forall n \in \mathbf{Z}_{0+}, \tag{23}$$

for each given $x_0 \in X$ for some $i = i(n) \in \overline{q}$ and we informally can say that the ith primary self-mapping $T_i : X \to X$ is "active" at the nth value (or sample) of the sequence $\{x_n\}$ [28]. See also [23–27]. In other words, the switching map $T = T_{\sigma_n}$ on X is defined by one of the self-mappings $T_i : X \to X$ $(i \in \overline{q})$ for each $n \in \mathbf{Z}_{0+}$ and it has associated piecewise constant functions $\rho_{\sigma_n}, k_{\sigma_n} : \mathbf{Z}_{0+} \times X \times X \to \mathbf{R}_+$ such that $\rho = \rho_{\sigma_n}(x, y) \in \{\rho_1, \rho_2, \ldots, \rho_q\}$, $k = k_{\sigma_n}(x, y) \in \{k_1, k_2, \ldots, k_q\}$ for each $n \in \mathbf{Z}_{0+}$ and each $x, y \in X$. The set of switching samples of a sequence $\{x_n\} \subset X$ is a (proper or improper) subset $Z_S = Z_S(\{x_n\}, \sigma)$ of $\mathbf{Z}_{0+}$, so-called the switching set, defined by $Z_S = \{n \in \mathbf{Z}_+ : T_{\sigma_n} x_n \neq T_{\sigma_{n-1}} x_n\} = \{n_0, n_1, \ldots, n_k, \ldots\}$. Note that a switching set is a strictly ordered set for the standard strict ordering relation "<". Since $T_i : X \to X$, $\forall i \in \overline{q}$, are self-mappings of $C_{\rho k}$ with constants $\rho_i(\leq k_i)$, $k_i \in \mathbf{R}_+$, then

$$F_{T_i x,x}\left(\rho_i^{-1} t\right) \geq F_{T_i^2 x, T_i x}(t) \geq F_{T_i x,x}\left(k_i^{-1} t\right); \quad \forall t \in \mathbf{R}_+, \tag{24}$$

for any $x \in X$ so that one has recursively from (3) for a sequence $\{x_n\} \subset X$ generated by $x_{n+1} = T_{\sigma_n} x_n$, $\forall n \in \mathbf{Z}_{0+}$, for any given $x_0 \in X$

$$\begin{aligned} &F_{x_1,x_0}\left(\left(\prod_{i=0}^{n}\left[\rho_{\sigma_i}^{-1}\right]\right)t\right) \\ &\quad = F_{x_1,x_0}\left(\left(\prod_{i=1}^{q}\left[\rho_i^{-\sum_{j=0}^{k-1}\gamma_i(n_j,n_{j+1})-\gamma_i(n_k,n_k+\ell)}\right]\right)t\right) \\ &\quad \geq F_{x_{n_k+\ell+1},x_{n_k+\ell}}(t) \geq F_{x_1,x_0}\left(\left(\prod_{i=0}^{n_k+\ell}\left[k_{\sigma_i}^{-1}\right]\right)t\right) \\ &\quad = F_{x_1,x_0}\left(\left(\prod_{i=1}^{q}\left[k_i^{-\sum_{j=0}^{k-1}\gamma_i(n_j,n_{j+1})-\gamma_i(n_k,n_k+\ell)}\right]\right)t\right); \\ &\qquad\qquad \forall k \in \mathbf{Z}_{0+},\ \forall t \in \mathbf{R}_+, \end{aligned} \tag{25}$$

where $\ell(\in \mathbf{Z}_{0+}) \leq n_{j+1} - n_j$, $n_k \in Z_S$, $\forall k \in \mathbf{Z}_{0+}$, is

$$n_k = n_{k-1} + \sum_{i=1}^{q}\gamma_i(n_{k-1}, n_k) = \sum_{j=0}^{k-1}\sum_{i=1}^{q}\gamma_i(n_j, n_{j+1}) \tag{26}$$

and $\gamma_i(n_j, n_j + \ell) \in \mathbf{Z}_{0+}$ is the number of times that the ith self-mapping $T_i : X \to X$ for some $i \in \overline{q}$ is "active" in the interval $[n_j, n_j + \ell)$ for each $i \in \overline{q}$. If $T_i : X \to X$, $\forall i \in \overline{q}$, are self-mappings of C_k-class, then one has instead of (25)

$$\begin{aligned} &F_{x_{n_k+\ell+1},x_{n_k+\ell}}(t) \\ &\quad \geq F_{x_1,x_0}\left(\left(\prod_{i=1}^{q}\left[k_i^{-\sum_{j=0}^{k-1}\gamma_i(n_j,n_{j+1})-\gamma_i(n_k,n_k+\ell)}\right]\right)t\right); \\ &\qquad\qquad \forall k \in \mathbf{Z}_{0+},\ \forall t \in \mathbf{R}_+. \end{aligned} \tag{27}$$

Theorem 14. *Let $(X, \mathbf{F})$ be a PM-space and let $T_i : X \to X$ be self-mappings of C_k-class for some constants $k_i \in \mathbf{R}_+$ for $i \in \overline{q} = \{1, 2, \ldots, q\}$. Then, the following properties hold:*

(i) *Assume that there is (at least) a self-mapping $T_i : X \to X$ for some $i \in \overline{q}$ which is a probabilistic strict contraction. Then, there are infinitely many switching laws $\sigma : \mathbf{Z}_{0+} \times X \to \overline{q}$ such that their associate switching maps $T : \mathbf{Z}_{0+} \times X \to X$ are probabilistic strict contractions.*

(ii) *Under the conditions of the above proposition, there are infinitely many switching laws* $\sigma : \mathbf{Z}_{0+} \times X \to \overline{q}$ *such that their associate switching maps* $T : \mathbf{Z}_{0+} \times X \to X$ *are probabilistic strict contractions and, furthermore, they consist of infinitely many alternate active switching maps of the form* (i, j) *or* (j, i) *with* $j \in \overline{q} \setminus \{i\}$.

(iii) *If, in addition,* $(X, \mathbf{F})$ *is complete, then any sequence* $x_{n+1} = T_{\sigma_n} x_n$, $n \in \mathbf{Z}_{0+}$ *under any switching law* $\sigma : \mathbf{Z}_{0+} \times X \to \overline{q}$ *fulfilling either Property (i) or Property (ii) for any given initial point* $x_0 \in X$ *is Cauchy and probabilistically convergent.*

Proof. It follows from (27) that Property (i) is fulfilled for sequences $\{x_n\} \subset X$ generated as $x_{n+1} = T_{\sigma_n} x_n$ for any $x_0 \in X$ by any of the infinitely many switching maps $T : \mathbf{Z}_{0+} \times X \to X$ built under switching laws $\sigma : \mathbf{Z}_{0+} \times X \to \overline{q}$ which fulfil

$$\lim_{k \to \infty} \sum_{j=0}^{k-1} \gamma_i \left(n_j, n_{j+1}\right) = +\infty, \qquad \sum_{\ell(\neq i)=1}^{q} \sum_{k=0}^{\infty} \gamma_\ell \left(n_k, n_{k+1}\right) < +\infty \tag{28}$$

since there is a finite nonnegative integer $n^* = n^*(\sigma)$ depending on the subsequence $\{x_n : n < n^*\}$ which is a terminal switching point such that $\sigma_n = i$ for $i \geq n^*$ for any such a sequence $\{x_n\}$. Thus, one gets from (27) that

$$\lim_{n \to \infty} F_{x_{n+m}, x_n}(t) = F_{T_{\sigma_0} x_0, x_0}(+\infty) = 1, \quad \forall t \in \mathbf{R}_+, \ \forall m \in \mathbf{Z}_+, \tag{29}$$

since $F : \mathbf{R} \to \mathbf{R}_{0+}$ is nondecreasing and left-continuous with $\sup_{t \in \mathbf{R}} F(t) = 1$ and then the sequence $\{x_n\}$ built as $x_{n+1} = Tx_n = T_{\sigma_n} x_n$, $x_0 \in X$ is a Cauchy sequence and $T : \mathbf{Z}_{0+} \times X \to X$ is a probabilistic strict contraction. Property (i) has been proved.

Property (ii) follows with alternate (probabilistic strict contraction versus remaining self-mapping) switching laws $\sigma : \mathbf{Z}_{0+} \times X \to \overline{q}$ defined by a switching set Z_S fulfilling the fact that if, for any $k \in \mathbf{Z}_{0+}$, $n_{k+j} \in Z_S$ for $j = 0, 1, \ldots, \ell_k - 2$ has active (perhaps nonprobabilistic strict contractions) self-mappings $T_j : X \to X$ for $j(\neq i) \in \overline{q}$, then $n_{k+\ell_k} \in Z_S$ is defined such that $T_i : X \to X$ is active on $[n_{k+\ell_k - 1}, n_{k+\ell_k}]$; that is, $x_{n+1} = T_{\sigma_n} x_n = T_i x_n$ for $n \in \{n_{k+\ell_k-1}, n_{k+\ell_k}\}$ with $n_{k+\ell_k}$ being defined with ℓ_k large enough such that

$$k_i^{n_{k+\ell_k} - n_{k+\ell_k - 1}} < \prod_{j(\neq i)=1}^{q} \left[k_j^{\sum_{\ell=k}^{k+\ell_k-1} \gamma_j(n_\ell, n_{\ell+1})} \right] \tag{30}$$

which lead to

$$\lim_{k \to \infty} \sum_{j=0}^{k-1} \gamma_i \left(n_j, n_{j+1}\right) = \lim_{k \to \infty} \sum_{\ell(\neq i)=1}^{q} \sum_{k=0}^{\infty} \gamma_\ell \left(n_k, n_{k+1}\right) = +\infty \tag{31}$$

so that we get again (29) and a similar conclusion. Property (iii) is obvious from the fact that $(X, \mathbf{F})$ is complete and $\{x_n\}$ is Cauchy under switching laws fulfilling either Property (i) or Property (ii). □

The following result is a direct consequence of Theorem 14 since mappings of $C_{\rho k}$-class are also of C_k-class.

Corollary 15. *Let* $(X, \mathbf{F})$ *be a PM-space and let* $T_i : X \to X$ *be self-mappings of* $C_{\rho k}$*-class for some constants* $\rho_i (\leq k_i)$, $k_i \in \mathbf{R}_+$, *for* $i \in \overline{q} = \{1, 2, \ldots, q\}$. *Then, Theorem 14 still holds.*

Theorem 16. *Let* $(X, \mathbf{F}, \Delta)$ *be a complete Menger space with* $\Delta(a, b) = \min(a, b)$ *and let* $T_j : X \to X$ *be self-mappings of* C_k*-class for some constants* $k_j \in \mathbf{R}_+$ *for all* $j \in \overline{q} = \{1, 2, \ldots, q\}$ *with at least* $T_i : X \to X$ *being a probabilistic strict contraction for some* $i \in \overline{q}$. *Let* $\sigma : \mathbf{Z}_{0+} \times X \to \overline{q}$ *be a switching law and let* $\{x_n\} \subset X$ *be a sequence generated as* $x_{n+1} = T_{\sigma_n} x_n$, $n \in \mathbf{Z}_{0+}$, *for any given* $x_0 \in X$ *such that their associate switching map* $T : \mathbf{Z}_{0+} \times X \to X$ *is defined by* $Tx_n = T_{\sigma_n} x_n = T_i x_n$ *for* $j \in \overline{q}$ *and all* $n \geq n^*$ *and some finite* $n^* \in \mathbf{Z}_{0+}$. *Then,* $\{x_n\} \to z_i$ *which is the unique fixed point of the strict contraction* $T_i : X \to X$.

Proof. Since $T_i : X \to X$ is a strict probabilistic contraction, $\lim_{n \to \infty} F_{x_{n+m}, x_n}(t) = F_{T_i x_{n^*}, x_{n^*}}(+\infty) = 1$, $\forall t \in \mathbf{R}_+$, $\forall m \in \mathbf{Z}_+$, from (29) since $Tx_n = T_{\sigma_n} x_n = T_i x_n$ for $n \geq n^*$. Then $\{x_n\}$ is probabilistically convergent to $z_i \in X$ which is a fixed point of the probabilistic strict contraction $T_i : X \to X$ as proved by contradiction. Assume that this is false so that $z_i \neq T_i z_i$ and then since $\{x_n\}$ is Cauchy, $\{x_n\} \to z_i$, $F : \mathbf{R} \to \mathbf{R}_{0+}$ which is nondecreasing and left-continuous, and $T_i : X \to X$ is a probabilistic strict contraction, one gets, for any given $t \in \mathbf{R}_+$ and $\lambda \in (0, 1)$ and all n $(\in \mathbf{Z}_{0+}) \geq N$ and some $N = N(t, \lambda) \in \mathbf{Z}_{0+}$,

$$\begin{aligned} 1 - \lambda_0(t) &\geq F_{z_i, T_i z_i}(t) \\ &\geq \Delta\left(\Delta\left(F_{z_i, x_n}\left(\frac{t}{4}\right), F_{T_i x_n, x_n}\left(\frac{t}{4}\right)\right), F_{T_i x_n, T_i z_i}\left(\frac{t}{2}\right)\right) \\ &\geq \Delta\left(\Delta\left(F_{z_i, x_n}\left(\frac{t}{4}\right), F_{T_i x_n, x_n}\left(\frac{t}{4}\right)\right), F_{T_i x_n, T z_i}\left(\frac{t}{4}\right)\right) \\ &\geq \Delta\left(\Delta\left(F_{z_i, x_n}\left(\frac{t}{4}\right), F_{T_i x_n, x_n}\left(\frac{t}{4}\right)\right), F_{x_n, z_i}\left(k_i^{-1}\frac{t}{4}\right)\right) \\ &\geq \Delta\left(\Delta\left(F_{z_i, x_n}\left(\frac{t}{4}\right), F_{T_i x_n, x_n}\left(\frac{t}{4}\right)\right), F_{x_n, z_i}\left(\frac{t}{4}\right)\right) \\ &> 1 - \lambda \end{aligned} \tag{32}$$

for some $\lambda_0 = \lambda_0(t) > 0$, $\forall t \in \mathbf{R}_+$, which implies that $\lambda \in (\lambda_0, 1)$ but since $\lambda \in (0, 1)$ can be chosen arbitrarily, it suffices to take $\lambda \in (0, \lambda_0]$ to get a contradiction. Then, $z_i = Tz_i$ which is proved to be unique again by contradiction. Assume that this is not the case so that there exist $z_{i1} = T^n z_{i1}$ and $z_{i2} = T^n z_{i2} \neq z_{i1}$, $\forall n \in \mathbf{Z}_{0+}$, which are fixed points of the probabilistic strict contraction $T_i : X \to X$. Thus, one gets the contradiction

$$\begin{aligned} 1 &> F_{z_{i1}, z_{i2}}(0^+) = F_{z_{i1}, z_{i2}}(t) = F_{T^n z_{i1}, T^n z_{i2}}(t) \\ &\geq F_{z_{i1}, z_{i2}}(k_i^{-n} t) = F_{z_{i1}, z_{i2}}(+\infty) = 1, \quad \forall t \in \mathbf{R}_+, \ \forall n \in \mathbf{Z}_{0+}, \end{aligned} \tag{33}$$

so that $z_{i2} = z_{i1} = z_i = T_i z_i$. Since $Tx_n = T_{\sigma_n} x_n = T_i x_n$ for $n \geq n^*$ for some finite $n^* \in \mathbf{Z}_{0+}$, we can write $x_{n^*} = T_{\sigma_{n^*-1}} \cdots \cdot T_{\sigma_0} x_0$ and $x_n = T_i^{n-n^*} x_{n^*}$ for $n(\in \mathbf{Z}_{0+}) \geq n^*$ to get that $\{x_n\}$ generated by $x_{n+1} = T_{\sigma_n} x_n$, $\forall n \in \mathbf{Z}_{0+}$, for any given arbitrary $x_0 \in X$ is probabilistically convergent to $z_i = T_i z_i$. □

The above result is a direct consequence of Theorem 14 which is also valid if $T_i : X \to X$ is of $C_{\rho k}$-class. However, note that, under the alternate switching laws in Theorem 14(ii), the limit points of sequences generated through the switching maps $T : \mathbf{Z}_{0+} \times X \to X$ are, in general, dependent on the initial points of the sequences and on the switching law.

The following result generalizes Theorem 16 without assuming any special contractive condition on at least one of $T_i : X \to X, \forall i \in \overline{q}$, with the only condition on the operators being that all of them are either of $C_{\rho k}$-class or C_k-class.

Theorem 17. *Let $(X, \mathbf{F})$ be a PM-space, let $T_i : X \to X$ be self-mappings of $C_{\rho k}$-class for some constants $\rho_i (\leq k_i)$, $k_i \in \mathbf{R}_+$, $\forall i \in \overline{q}$, and let $T : \mathbf{Z}_{0+} \times X \to X$ be a switching mapping associated with a switching law $\sigma : \mathbf{Z}_{0+} \times X \to \overline{q}$ which generates a sequence $\{x_n\} \subset X$ as $x_{n+1} = T_{\sigma_n} x_n$, $\forall n \in \mathbf{Z}_{0+}$, for some given $x_0 \in X$. Let $Z_S \subseteq \mathbf{Z}_{0+}$ be the set of switching points, that is, for any given $n_k \in Z_S$, $\forall k \in \mathbf{Z}_+$, provided that $n_{k-1} \in Z_S$, if and only if $\sigma_{n_k} \neq \sigma_{n_k - 1} = \sigma_{n_{k-1}}$. Then, the following properties hold:*

(i)

$$F_{Tx_0,x_0}(\overline{\rho}_n t) \geq F_{x_{n+1},x_n}(t) \geq F_{Tx_0,x_0}(\overline{\gamma}_n t); \quad \forall t \in \mathbf{R}_+, \tag{34}$$

for all $n \in [n_k, n_k + \ell) \cap \mathbf{Z}_{0+}$, $\ell(\in \mathbf{Z}_{0+}) \leq n_{k+1} - n_k$, $n_k \in Z_S$, $\forall k \in \mathbf{Z}_{0+}$, where

$$\overline{\gamma}_n = \prod_{i=1}^{q} \left[k_i^{-\sum_{j=0}^{k-1} \gamma_i(n_j, n_{j+1}) - \gamma_i(n_k, n_k + \ell)} \right], \tag{35}$$

$$\overline{\rho}_n = \prod_{i=1}^{q} \left[\rho_i^{-\sum_{j=0}^{k-1} \gamma_i(n_j, n_{j+1}) - \gamma_i(n_k, n_k + \ell)} \right], \tag{36}$$

where $\gamma_i(n_j, n_j + \ell) \in \mathbf{Z}_{0+}$ is the number of times that the ith self-mapping $T_i : X \to X$ for each $i \in \overline{q}$ is "active" in the interval $[n_j, n_j + \ell)$ for some $i \in \overline{q}$.

(ii) *If $F : \mathbf{R} \to [0,1]$ is upper-semicontinuous at $\overline{\rho} t$ for some given $t \in \mathbf{R}_+$, where $\overline{\rho} = \limsup_{n\to\infty} \overline{\rho}_n$, then $\{x_n\}$ has the following property:*

$$\limsup_{n\to\infty} F_{x_{n+1},x_n}(t) \leq F_{Tx_0,x_0}(\overline{\rho} t). \tag{37}$$

If $F : \mathbf{R} \to [0,1]$ is lower-semicontinuous at $\overline{\rho} t$ for some given $t \in \mathbf{R}_+$, where $\underline{\gamma} = \liminf_{n\to\infty} \overline{\gamma}_n$, then $\{x_n\}$ has the following property:

$$\liminf_{n\to\infty} F_{x_{n+1},x_n}(t) \geq F_{Tx_0,x_0}(\underline{\gamma} t). \tag{38}$$

(iii) *If $T_i : X \to X$ are of C_k-class for some constants $k_i \in \mathbf{R}_+$, $\forall i \in \overline{q}$, then $\{x_n\}$ has the following property:*

$$F_{x_{n+1},x_n}(t) \geq F_{Tx_0,x_0}(\overline{\gamma}_n t); \quad \forall t \in \mathbf{R}_+. \tag{39}$$

And all $n \in [n_k, n_k + \ell) \cap \mathbf{Z}_{0+}$, $\ell(\in \mathbf{Z}_{0+}) \leq n_{k+1} - n_k$, $n_k \in Z_S$, $\forall k \in \mathbf{Z}_{0+}$. If $F : \mathbf{R} \to [0,1]$ is lower-semicontinuous at $\overline{\rho} t$ for a given $t \in \mathbf{R}_+$, where $\underline{\gamma} = \liminf_{n\to\infty} \overline{\gamma}_n$, then (38) holds.

Proof. Property (i) follows since (34), subject to (35)-(36), is obtained directly from (25)-(26). If $F : \mathbf{R} \to [0,1]$ is upper-semicontinuous at $\overline{\rho} t$, then

$$\begin{aligned} \limsup_{n\to\infty} F_{x_{n+1},x_n}(t) &\leq \limsup_{\overline{\rho}_n \to \overline{\rho}} F_{Tx_0,x_0}(\overline{\rho}_n t) \\ &\leq F_{Tx_0,x_0}\left(\left(\limsup_{\overline{\rho}_n \to \overline{\rho}} \overline{\rho}_n \right) t \right) \\ &= F_{Tx_0,x_0}(\overline{\rho} t). \end{aligned} \tag{40}$$

In the same way, if $F : \mathbf{R} \to [0,1]$ is lower-semicontinuous at $\underline{\gamma} t$, we get $\liminf_{n\to\infty} F_{x_{n+1},x_n}(t) \geq F_{Tx_0,x_0}(\underline{\gamma} t) = F_{Tx_0,x_0}((\liminf_{\overline{\gamma}_n \to \underline{\gamma}} \overline{\gamma}_n) t)$. This proves Property (ii). Property (iii) is a restriction of Properties (i)-(ii) for the case when $T_i : X \to X$ are of C_k-class for some constants $k_i \in \mathbf{R}_+$, $\forall i \in \overline{q}$. □

Note that although $F : \mathbf{R} \to [0,1]$ is assumed to be everywhere left-continuous in the probabilistic metric framework, this does not mean that it is everywhere lower- and/or upper-semicontinuous. Therefore, some extra related conditions are imposed in Theorem 17(ii)-(iii) allowing obtaining limit upper- and lower-bounds of $F_{x_{n+1},x_n}(t)$ as $n \to \infty$ via the limit superior and the limit inferior.

The following two results related to bounded and unbounded sequences follow from Theorem 17.

Corollary 18. *Let $(X, \mathbf{F})$ be a PM-space, let $T_i : X \to X$ be self-mappings of C_k-class for some constants $k_i \in \mathbf{R}_+$, $\forall i \in \overline{q}$, and let $T : \mathbf{Z}_{0+} \times X \to X$ be a switching mapping associated with a switching law $\sigma : \mathbf{Z}_{0+} \times X \to \overline{q}$ which generates a sequence $\{x_n\} \subset X$ as $x_{n+1} = T_{\sigma_n} x_n$, $\forall n \in \mathbf{Z}_{0+}$, $\forall n \in \mathbf{Z}_{0+}$, for some given $x_0 \in X$ with switching points in the switching set Z_S. Assume also that $F : \mathbf{R} \to [0,1]$ is everywhere lower-semicontinuous with $\underline{\gamma} = \liminf_{n\to\infty} \overline{\gamma}_n > 0$ for such a sequence $\{x_n\}$. Then, the following properties hold:*

(i) *$\{x_n\}$ is probabilistically bounded.*

(ii) *If $F_{x_0,Tx_0}(t) = H(t - d(x, Tx))$, $\forall x_0 \in X$, $\forall t \in \mathbf{R}$, then $\exists \lim_{n\to\infty} H(t - d(T^{n+1}x_0, T^n x_0)) = 1$ if $t \in (d(x_0, Tx_0)/\underline{\gamma}, +\infty)$.*

Proof. From (39), one concludes (38); that is, $\liminf_{n\to\infty} F_{x_{n+1},x_n}(t) \geq F_{Tx_0,x_0}(\underline{\gamma} t)$, $\forall t \in \mathbf{R}_+$. Since $\underline{\gamma} > 0$, then $\sup_{t\in\mathbf{R}_{0+}} F_{Tx_0,x_0}(\underline{\gamma} t) = 1$ and $\{x_n\}$ is probabilistically bounded since a nonempty set $A = A(\{x_{n_k}\}) \subset X$ which contains all the points of some subsequence $\{x_{n_k}\} \subseteq \{x_n\}$ has the property that $D_A \in D_+$. Property (i) has been proved.

On the other hand, it follows from (25) and Property (i) that if $F_{x_0,Tx_0}(t) = H(t - d(x,Tx))$, $\forall x_0 \in X$, $\forall t \in \mathbf{R}$, then

$$\begin{aligned} &\lim_{n\to\infty} H\left(t - d\left(T^{n+1}x_0, T^n x_0\right)\right) \\ &\quad \geq H\left(\underline{\gamma} t - d\left(x_0, Tx_0\right)\right) = 1; \\ &\qquad\qquad t \in \left(\frac{d\left(x_0, Tx_0\right)}{\underline{\gamma}}, +\infty\right) \end{aligned} \tag{41}$$

and Property (ii) is proved. □

Since mappings of C_k-class are also of $C_{\rho k}$-class, then Corollary 18 also holds if some of $T_i : X \to X$ are self-mappings of $C_{\rho k}$-class for some constants $\rho_i (\leq k_i)$, $k_i \in \mathbf{R}_+$, $\forall i \in \overline{q}$.

Corollary 19. *Let $(X, \mathbf{F})$ be a PM-space, let $T_i : X \to X$ be self-mappings of C_k-class for some constants $k_i \in \mathbf{R}_+$, $\forall i \in \overline{q}$, and let $T : \mathbf{Z}_{0+} \times X \to X$ be a switching mapping associated with a switching law $\sigma : \mathbf{Z}_{0+} \times X \to \overline{q}$ which generates a sequence $\{x_n\} \subset X$ as $x_{n+1} = T_{\sigma_n} x_n$, $\forall n \in \mathbf{Z}_{0+}$, $\forall n \in \mathbf{Z}_{0+}$, for some given $x_0 \in X$ with switching points in the switching set Z_S. Assume also that $F : \mathbf{R} \to [0,1]$ is everywhere upper-semicontinuous with $\overline{\rho} = \limsup_{n\to\infty} \overline{\rho}_n = 0$ for such a sequence $\{x_n\}$. Then, the following properties hold:*

(i) *$\{x_n\}$ is probabilistically unbounded.*

(ii) *If $F_{x_0,Tx_0}(t) = H(t - d(x,Tx))$, $\forall x_0 \in X$, $\forall t \in \mathbf{R}$, then $\exists \lim_{t\to+\infty} \lim_{n\to\infty} H(t - d(T^{n+1}x_0, T^n x_0)) = 1$.*

Proof. From (39), one concludes (38); that is, $\limsup_{n\to\infty} F_{x_{n+1},x_n}(t) \leq F_{Tx_0,x_0}(\overline{\rho} t) = F_{Tx_0,x_0}(0) = 0$, $\forall t \in \mathbf{R}_+$. Consider a nonempty set $A = A(\{x_{n_k}\}) \subset X$ which contains all the points of some subsequence $\{x_{n_k}\} \subseteq \{x_n\}$. It is obvious that $D_A \notin D_+$. Then $\{x_n\}$ is probabilistically unbounded since it has a probabilistically unbounded subsequence. Property (i) has been proved. On the other hand, it follows from (25) and Property (i) that if $F_{x_0,Tx_0}(t) = H(t - d(x,Tx))$, $\forall x_0 \in X$, $\forall t \in \mathbf{R}$, then

$$\begin{aligned} &\lim_{t\to+\infty} \lim_{n\to\infty} H\left(t - d\left(T^{n+1}x_0, T^n x_0\right)\right) \\ &\quad \leq \lim_{t\to+\infty} H\left(\overline{\rho} t - d\left(x_0, Tx_0\right)\right) = H\left(-d\left(x_0, Tx_0\right)\right) \\ &\quad = H(0) = 0 \end{aligned} \tag{42}$$

and Property (ii) is proved. □

Remarks. (1) Note that Corollary 18 is fulfilled, in particular, by switched sequences with switching set of finite cardinal with a terminal point of switching to some nonexpansive $C_{\rho k}$ (or C_k) self-mapping $T_i : X \to X$; that is, $k_i \in (0,1]$ for some $i \in \overline{q}$. Note also that Corollary 18 is not fulfilled for terminal switching to an expansive self-mapping.

(2) Note that Corollary 19 for $C_{\rho k}$ self-mappings can also be applied to self-mappings $T_i : X \to X$, $i \in \overline{q}$, which are only subject to the upper-bounding rule of the probability density function (say self-mappings "of C_ρ-class"); that is,

$$F_{T_i x_0, x_0}\left(\rho_i^{-1} t\right) \geq F_{T_i x_0, T_i^2 x_0}(t); \quad \forall t \in \mathbf{R}_+, \tag{43}$$

while leading to a similar conclusion about probabilistic unboundedness. Note that the result applies, in particular, to switched sequences with switching set of finite cardinal with a terminal point of switching to some expansive $C_{\rho k}$ (or C_ρ) self-mapping $T_i : X \to X$ for some $i \in \overline{q}$; that is, $\rho_i > 1$ for some $i \in \overline{q}$.

(3) Note that Corollary 18 is also fulfilled by switched sequences with switching set of infinite cardinal if $\underline{\gamma} > 0$. Corollary 18 excludes switching laws $\sigma : \mathbf{Z}_{0+} \times X \to \overline{q}$ being built under some expansive self-mapping $T_i : X \to X$ for some $i \in \overline{q}$ being used so infinitely often such that $\underline{\gamma} = 0$ with either finite terminal switching point or not. This includes, in particular, the case when the switched map is built with a finite terminal switching point to an expansive self-mapping. Then, the probabilistic unboundedness result of Corollary 19 applies.

Note that an expansive mapping can give a fixed point which unbounded sequences do not converge to as the next example visualizes.

Example 20. Assume that $(X, \mathbf{F})$ is a PM-space with $0 \in X$ and consider a self-map $T : X \to X$ such that $T0 = 0$. Thus, $z = 0$ is trivially a fixed point of $T : X \to X$. Assume the following cases:

(a) $F_{0,x}(t) \geq F_{0,Tx}(\rho t)$, $\forall t \in \mathbf{R}_+$, and all $x (\neq 0) \in X$ for some real constant $\rho > 1$ so that $T : X \to X$ is expansive. Then, $F_{0,T^n x}(t) \leq F_{0,x}(\rho^{-n} t)$, $\forall t \in \mathbf{R}_+$, $\forall n \in \mathbf{Z}_{0+}$, so that $\lim_{n\to\infty} F_{0,T^n x}(t) = 0$, $\forall t \in \mathbf{R}_+$, $\forall n \in \mathbf{Z}_{0+}$, and the sequence $\{T^n x\}$ is unbounded and does not converge in probability to the fixed point $z = 0$. There is an analogy with the expansive deterministic counterpart examples. For instance, consider a scalar difference equation $x_{n+1} = \rho x_n$, $\forall n \in \mathbf{Z}_{0+}$, with $|\rho| > 1$ and $x_0 \neq 0$. Then, $z = 0$ is a fixed point of the self-mapping on $\mathbf{R}$ defining the sequence trajectory solution which is clearly expansive for the metric space $(\mathbf{R}, d)$ with d being any metric, for instance, the Euclidean norm. However, $|x_n| \to +\infty$ as $n \to \infty$ so that the unique fixed point is an unstable equilibrium point.

(b) $F_{0,Tx}(t) \geq F_{0,x}(k^{-1} t)$, $\forall t \in \mathbf{R}_+$, and all $x (\neq 0) \in X$ for some real constant $k \in (0,1]$ so that $T : X \to X$ is nonexpansive. Then, if furthermore $k \in (0,1)$, then $T : X \to X$ is strictly contractive and $F_{0,T^n x}(t) \geq F_{0,x}(k^{-n} t)$, $\forall n \in \mathbf{Z}_{0+}$, so that $\lim_{n\to\infty} F_{0,T^n x}(t) = 1$ and $\{T^n x\} \to z = 0$, which is a fixed point, with probability one, is also Cauchy and probabilistically bounded. The deterministic counterpart can be the example $x_{n+1} = k x_n$ with $x_0 \neq 0$, $\forall n \in \mathbf{Z}_{0+}$, with $|k| < 1$ in a metric space $(\mathbf{R}, d)$, where $z = 0$ is the unique fixed point (and a globally

asymptotically stable equilibrium point) of the self-mapping defining the sequence trajectory solution for any initial condition.

(c) In the above case, assuming $k = 1$, then, for any $x \in X$, $F_{T^{n+1}x,T^n x}(t) \geq F_{x,Tx}(t)$, $\forall t \in \mathbf{R}_+$, $\forall n \in \mathbf{Z}_{0+}$. If $Tx = x$, then $F_{T^{n+1}x,T^n x}(t) = 1$, $T^n x = x$, $\forall n \in \mathbf{Z}_{0+}$; the mapping $T : X \to X$ defining the solution is not expansive but not contractive and any point is a fixed point. A deterministic counterpart can be visualized with the example $x_{n+1} = x_n$, $\forall n \in \mathbf{Z}_{0+}$, for any $x_0 \in X$ which has infinitely many fixed points which are also (nonasymptotically) stable equilibrium points.

(d) The above discussion can be directly extended to the case of switching maps built under switching laws with finite terminal switching point to an expansive primary self-mapping or for appropriate switching laws with no terminal switching point in the presence of at least one expansive primary self-mapping.

A worked numerical example follows.

Example 21. This example aims at numerically illustrating the main results stated and proved in Section 3 through Theorem 17 (as a generalization of Theorem 16). For this purpose, consider

$$F_{x,y}(t) = H(t - d(x,y)) = \begin{cases} 0, & t \leq d(x,y), \\ 1, & t > d(x,y), \end{cases} \tag{44}$$

where $d(x,y)$ is the distance induced by the 2-norm, $d(x,y) = \|x - y\|_2$. Consider also the switched self-mapping described by the discrete dynamical system $x_{n+1} = T_{\sigma_n} x_n$ with $T_{\sigma_n} = A_{\sigma_n} + F_{\sigma_n}$, $A_{\sigma_n} \in \{A_1, A_2, A_3\}$, and $F_{\sigma_n} \in \{F_1, F_2, F_3\}$. This way, A matrices can be regarded as dynamics matrices while F matrices can be understood as perturbation ones. The dynamics matrices are selected in such a way that in conjunction with the perturbation ones the switched map exhibits different characters (contractive, nonexpansive, and expansive) for each $i = 1, 2, 3$ so that the effect of switching can be positively noticed. Thus, the dynamics matrices are given by

$$A_1 = \begin{pmatrix} 0.1 & 0.3 & 0.2 \\ 0.1 & 0.2 & 0.2 \\ 0.1 & 0.2 & 0.1 \end{pmatrix},$$
$$A_2 = \begin{pmatrix} 0.2 & 0.3 & 0.4 \\ 0.425 & 0.2 & 0.3 \\ 0.1 & 0.3 & 0.1 \end{pmatrix}, \tag{45}$$
$$A_3 = \begin{pmatrix} 1.2 & 0.3 & 1.4 \\ 0.5 & 0.2 & 1.3 \\ 0.1 & 0.3 & 1.1 \end{pmatrix},$$

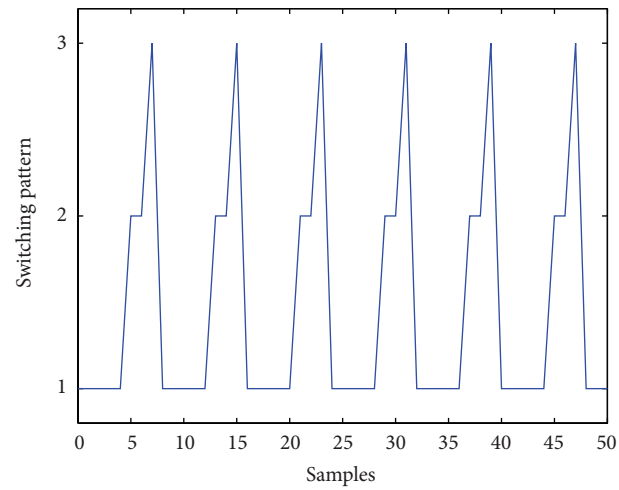


Figure 1: Periodic switching map for the first experiment.

whose eigenvalues are, respectively, given by

$$\begin{aligned} &\operatorname{spec}(A_1) = \{0.741, -0.17 + 0.153j, -0.17 - 0.153j\}, \\ &\operatorname{spec}(A_2) \\ &\quad = \{0.7722, -0.1361 + 0.1191j, -0.1361 - 0.1191j\}, \\ &\operatorname{spec}(A_3) = \{1.8691, 0.7381, -0.1073\}. \end{aligned} \tag{46}$$

The constants characterizing each one of these matrices, interpreted as operators, are calculated from (2) $F_{Tx,Ty}(t) \geq F_{x,y}(k^{-1}t)$, which in this particular case takes the form

$$\begin{aligned} H(t - d(A_i x, A_i y)) &\geq H\left(k^{-1}t - d(x,y)\right) \\ &= H\left(\frac{t - kd(x,y)}{k}\right) \end{aligned} \tag{47}$$

for each one of the matrices, $i = 1, 2, 3$. The latter condition is satisfied if $d(A_i x, A_i y) \leq k_i d(x,y)$ which results in the considered metric in $\|A_i x - A_i y\|_2 \leq k_i \|x - y\|_2$. Therefore, for the first matrix, we have

$$\begin{aligned} \|A_1 x - A_1 y\|_2 &\leq \|A_1\|_2 \|x - y\|_2 = 0.534 \|x - y\|_2 \\ &< 0.54 \|x - y\|_2 \end{aligned} \tag{48}$$

so that $k_1 = 0.54$. The 2-norms of the remaining matrices are $\|A_2\|_2 = 0.8$ and $\|A_3\|_2 = 2.528$. As it can be seen, A_1 and A_2 are contractive operators while A_3 is expansive. However, the perturbation matrices F will shape the behavior of the operator $T_{\sigma_n} = A_{\sigma_n} + F_{\sigma_n}$ in a different way. To this end, fix

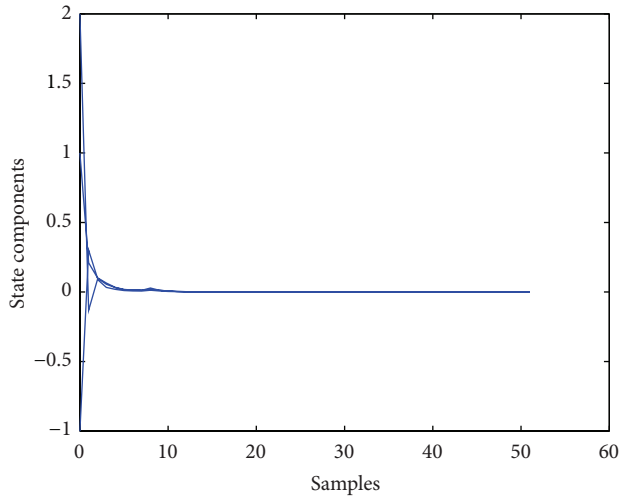


FIGURE 2: Evolution of the state variables for the first experiment.

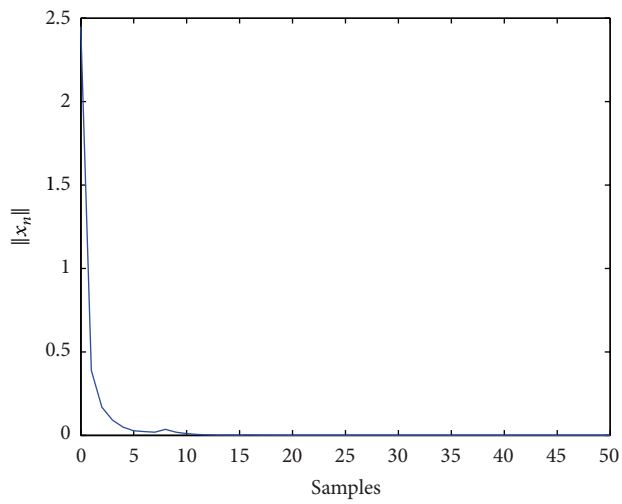


FIGURE 3: Evolution of the norm of the state for the first experiment.

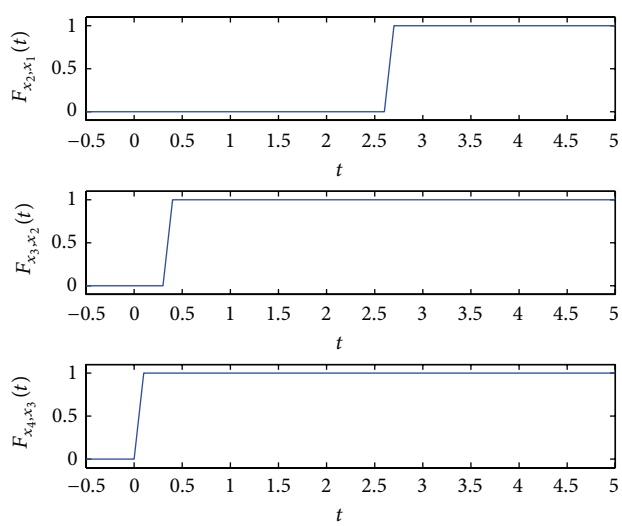


FIGURE 4: Evolution of $F_{x_{n+1},x_n}(t)$ as n increases for the first experiment.

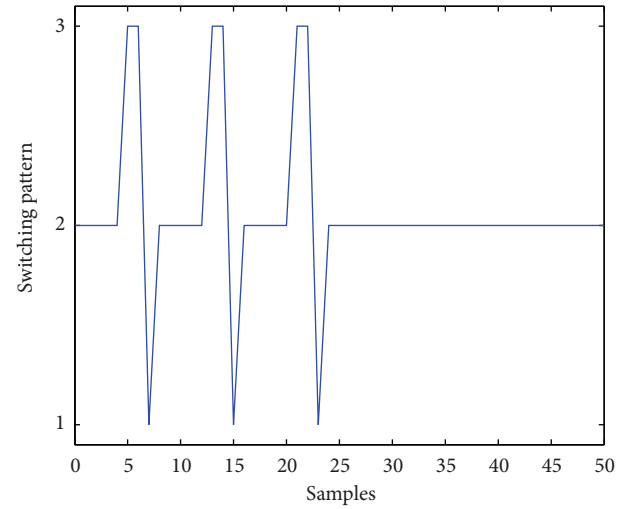


FIGURE 5: Switching map for the second experiment.

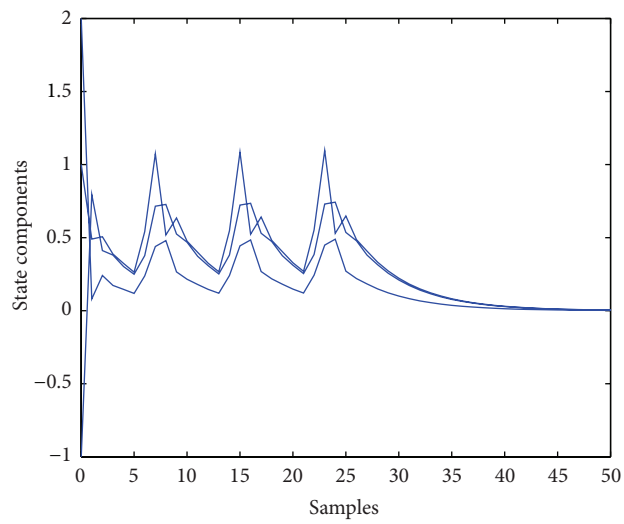


FIGURE 6: Evolution of the state variables in the second experiment.

$\varepsilon = 0.2$ and set the perturbation matrices in such way that $\|F_i\|_2 = \sqrt{\lambda_{\max}(F^T F)} \leq \varepsilon$ for $i = 1, 2, 3$. Thus, let them be

$$F_1 = \begin{pmatrix} 0.01 & 0.02 & -0.01 \\ 0.02 & -0.08 & -0.01 \\ -0.01 & 0.18 & -0.01 \end{pmatrix},$$

$$F_2 = \begin{pmatrix} -0.01 & 0.18 & -0.01 \\ 0.01 & 0.01 & 0.02 \\ 0.02 & -0.08 & -0.01 \end{pmatrix}, \tag{49}$$

$$F_3 = \begin{pmatrix} 0.02 & -0.08 & -0.01 \\ -0.01 & 0.18 & -0.01 \\ 0.01 & 0.02 & -0.01 \end{pmatrix}.$$

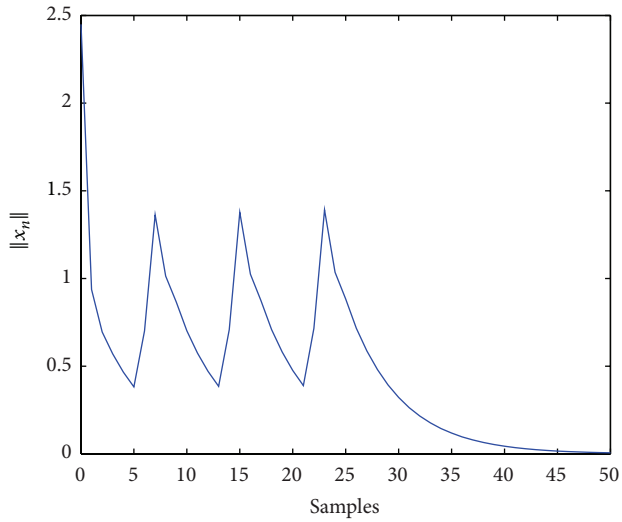


Figure 7: Evolution of the norm of the state in the second experiment.

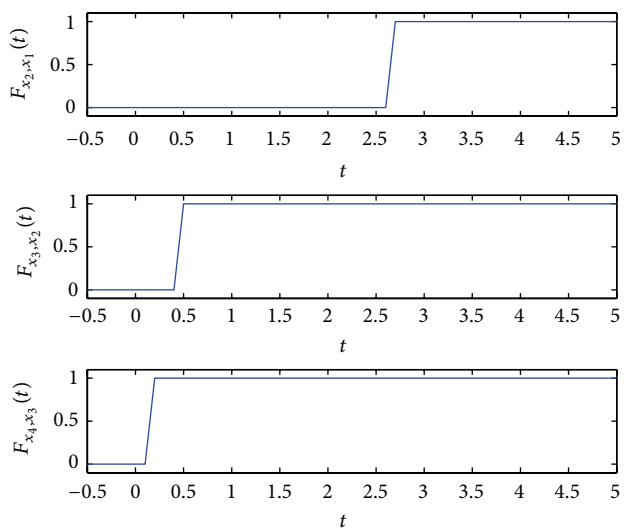


Figure 8: Evolution of $F_{x_{n+1},x_n}(t)$ as n increases for the second experiment.

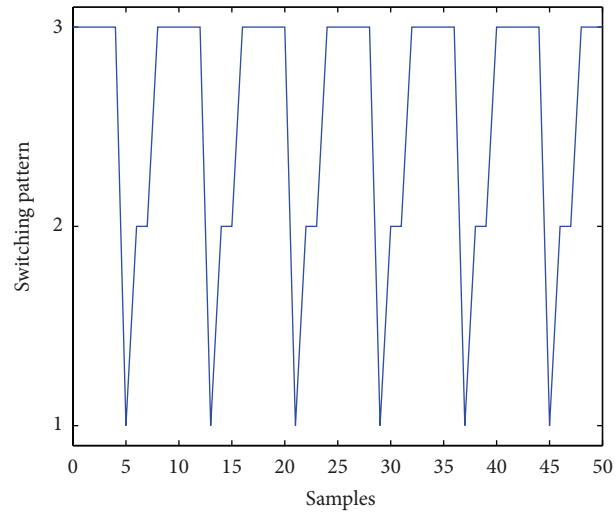


Figure 9: Switching patterns for the third experiment.

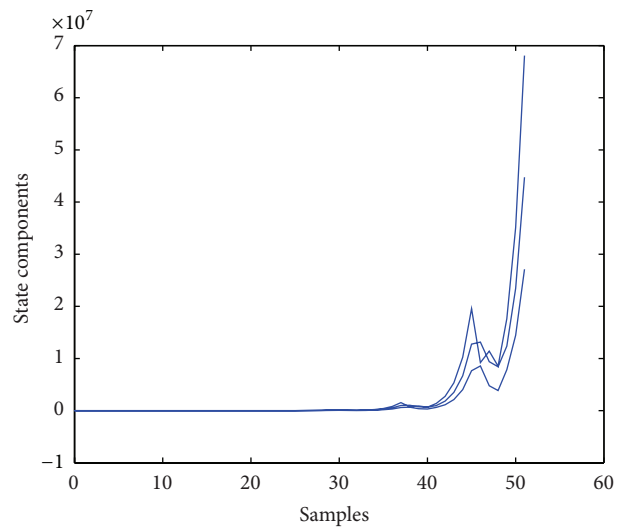


Figure 10: Evolution of the state variables for the third experiment.

The perturbation matrices are generated by swapping a file for each subsystem. Therefore, $\|F_1\|_2 = \|F_2\|_2 = \|F_3\|_2 = 0.1987$ near the upper bounding value of $\varepsilon = 0.2$. Under these circumstances, the worst-case operator $A_1 + \varepsilon I$ is still contractive, $A_2 + \varepsilon I$ is nearly nonexpansive and noncontractive, and $A_3 - \varepsilon I$ is always expansive providing different dynamics to each of the subsystems. With this setup, we will perform three simulation experiments to illustrate the diversity of dynamical behaviours that can be achieved by modifying the switching law. The initial condition in all experiments is $x_0 = [1 \ -1 \ 2]^T$.

Experiment 1. The switching law never stops and the first subsystem prevails over the other two. The switching law is selected to be periodic, with a period of 8 samples according to the pattern displayed in Figure 1. This means that in one period we have $5 = \gamma_1 > \gamma_2 + \gamma_3 = 2 + 1$; that is, the time interval within which the first subsystem is active is larger than the sum of the intervals corresponding to the other subsystems. Thus, Theorem 17 ensures that the sequence of iterates is probabilistically bounded and converges to the fixed point $x = 0$ as Figure 2 displays for the state components and Figure 3 for the norm of the state. Moreover, Figure 4 shows how $F_{x_{n+1},x_n(t)}$ converges to the Heaviside function $H(t)$ as the iteration variable n increases.

Experiment 2. The switching map converges in finite time to the second subsystem, which is nearly to be nonexpansive and noncontractive. Thus, the switching map is depicted in Figure 5. Figures 6 and 7 show, respectively, the evolution of the state variables and their norm as n increases. As Theorem 17 states, the sequence of iterates is bounded and converges to the fixed point because, due to numerical round-off errors, the operator T_2 is slightly contractive. However,

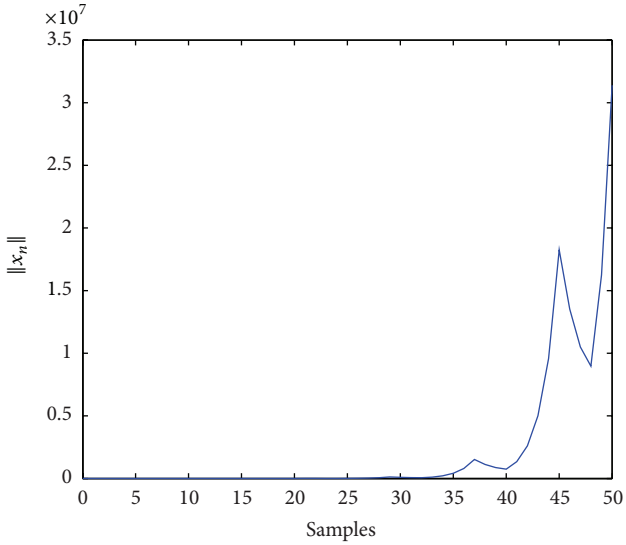


Figure 11: Norm of the state for the third experiment.

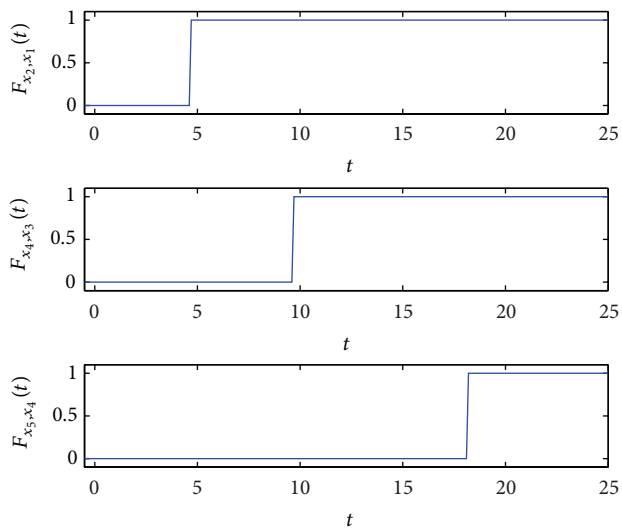


Figure 12: Evolution of $F_{x_{n+1},x_n}(t)$ as n increases for the third experiment.

it converges at a lower rate than in Experiment 1 since the operator T_2 is very nearly to be noncontractive and nonexpansive. In this case, $F_{x_{n+1},x_n}(t)$ also converges to the Heaviside function $H(t)$ as the iteration variable n increases, as it can be noticed in Figure 8.

Experiment 3. The switching law never stops and the third subsystem prevails over the other two. A periodic switching signal is considered again with $\gamma_3 > \gamma_1 + \gamma_2$ in a period. Figure 9 gives the switching law. In addition, Figures 10 and 11, respectively, display the evolution of the state variables and their norm as n increases. As we could have expected from Theorem 17, they are not bounded and diverge asymptotically since the expansive operator dominates the period of the switching pattern. Contrarily to the previous experiments, $F_{x_{n+1},x_n}(t)$ converges now to the identically null function, since, as n increases, the Heaviside function $H(t)$ is displaced to the right, as it is represented in Figure 12.

Conflict of Interests

The authors declare that they do not have competing interests.

Authors' Contribution

The authors declare that all the authors contributed equally to all of the parts of this paper.

Acknowledgments

The authors are very grateful to the Spanish Government for Grants DPI2012-30651 and DPI2013-47825-C3-1-R and also to the Basque Government and UPV/EHU for Grants IT378-10, SAIOTEK SPE13UN039, and UFI 2011/07.

References

[1] S. Kumar and S. Chauhan, "Fixed points of expansion mappings in Menger spaces using weak compatibility," *The Mathematics Student*, vol. 82, no. 2–4, pp. 157–164, 2013.

[2] R. J. Egbert, "Products and quotients of probabilistic metric spaces," *Pacific Journal of Mathematics*, vol. 24, pp. 437–455, 1968.

[3] V. M. Sehgal and A. T. Bharucha-Reid, "Fixed points of contraction mappings on probabilistic metric spaces," *Mathematical Systems Theory*, vol. 6, pp. 97–102, 1972.

[4] D. Miheţ, "Altering distances in probabilistic Menger spaces," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 71, no. 7-8, pp. 2734–2738, 2009.

[5] C.-M. Chen, "Fixed point theorems of generalized cyclic orbital Meir-Keeler contractions," *Fixed Point Theory and Applications*, vol. 2013, article 91, 2013.

[6] M. A. Al-Thagafi and N. Shahzad, "Convergence and existence results for best proximity points," *Nonlinear Analysis, Theory, Methods & Applications*, vol. 70, no. 10, pp. 3665–3671, 2009.

[7] S. Rezapour, M. Derafshpour, and N. Shahzad, "Best proximity points of cyclic φ-contractions on reflexive Banach spaces," *Fixed Point Theory and Applications*, vol. 2010, Article ID 946178, 7 pages, 2010.

[8] C. Di Bari, T. Suzuki, and C. Vetro, "Best proximity points for cyclic Meir-Keeler contractions," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 69, no. 11, pp. 3790–3794, 2008.

[9] S. Rezapour, M. Derafshpour, and N. Shahzad, "On the existence of best proximity points of cyclic contractions," *Advances in Dynamical Systems and Applications*, vol. 6, no. 1, pp. 33–40, 2011.

[10] N. Hussain, M. A. Kutbi, and P. Salimi, "Best proximity point results for modified $\alpha - \psi$ proximal rational contractions," *Abstract and Applied Analysis*, vol. 2013, Article ID 927457, 14 pages, 2013.

[11] M. Gabeleh, "Best proximity point theorems via proximal non-self mappings," *Journal of Optimization Theory and Applications*, vol. 164, no. 2, pp. 565–576, 2015.

[12] M. De la Sen, R. P. Agarwal, and A. Ibeas, "Results on proximal and generalized weak proximal contractions including the case

of iteration-dependent range sets," *Fixed Point Theory and Applications*, vol. 2014, article 169, 2014.

[13] S. Manro and Sumitra, "Common new fixed point theorem in modified intuitionistic fuzzy metric spaces using implicit relation," *Applied Mathematics*, vol. 4, no. 9, pp. 27–31, 2013.

[14] A. F. Roldán-López-de-Hierro, M. de la Sen, J. Martínez-Moreno, and C. Roldán-López-de-Hierro, "An approach version of fuzzy metric spaces including an ad hoc fixed point theorem," *Fixed Point Theory and Applications*, vol. 2015, article 33, 2015.

[15] M. De la Sen and E. Karapınar, "Some results on best proximity points of cyclic contractions in probabilistic metric spaces," *Journal of Function Spaces*, vol. 2015, Article ID 470574, 11 pages, 2015.

[16] M. De la Sen, "On fixed and best proximity points of cyclic C-contractions in probabilistic complete metric and Banach spaces," *Bulletin of the Malaysian Mathematical Sciences Society*, 2015.

[17] S. A. Sahab, M. S. Khan, and S. Sessa, "A result in best approximation theory," *Journal of Approximation Theory*, vol. 55, no. 3, pp. 349–351, 1988.

[18] M. De la Sen and E. Karapinar, "On a cyclic Jungck modified TS-iterative procedure with application examples," *Applied Mathematics and Computation*, vol. 233, pp. 383–397, 2014.

[19] A. R. Khan, V. Kumar, and N. Hussain, "Analytical and numerical treatment of Jungck-type iterative schemes," *Applied Mathematics and Computation*, vol. 231, pp. 521–535, 2014.

[20] J.-Z. Xiao, X.-H. Zhu, and X.-Y. Liu, "An alternative characterization of probabilistic Menger spaces with H-type triangular norms," *Fuzzy Sets and Systems*, vol. 227, pp. 107–114, 2013.

[21] T. Dinevari and M. Frigon, "Fixed point results for multivalued contractions on a metric space with a graph," *Journal of Mathematical Analysis and Applications*, vol. 405, no. 2, pp. 507–517, 2013.

[22] M. de la Sen and A. Ibeas, "On the stability properties of linear dynamic time-varying unforced systems involving switches between parameterizations from topologic considerations via graph theory," *Discrete Applied Mathematics*, vol. 155, no. 1, pp. 7–25, 2007.

[23] M. D. Sen, S. Alonso-Quesada, and A. Ibeas, "On the asymptotic hyperstability of switched systems under integral-type feedback regulation Popovian constraints," *IMA Journal of Mathematical Control and Information*, vol. 32, no. 2, pp. 359–386, 2015.

[24] M. De la Sen and A. Ibeas, "Properties of convergence of a class of iterative processes generated by sequences of self-mappings with applications to switched dynamic systems," *Journal of Inequalities and Applications*, vol. 2014, article 498, 2014.

[25] M. de la Sen and A. Ibeas, "On the global asymptotic stability of switched linear time-varying systems with constant point delays," *Discrete Dynamics in Nature and Society*, vol. 2008, Article ID 231710, 31 pages, 2008.

[26] C. Duan and F. Wu, "New results on switched linear systems with actuator saturation," *International Journal of Systems Science*, 2014.

[27] M. Rajchakit and G. Rajchakit, "Exponential stability of stochastic hybrid systems with nondifferentiable and interval time-varying delay," *International Journal of Pure and Applied Mathematics*, vol. 95, no. 1, pp. 79–88, 2014.

[28] M. De la Sen, "Linking contractive self-mappings and cyclic Meir-Keeler contractions with Kannan self-mappings," *Fixed Point Theory and Applications*, vol. 2010, Article ID 572057, 2010.