

Modeling and Control of Complex Dynamic Systems: Applied Mathematical Aspects

Guest Editors: Zhiwei Gao, Dexing Kong, and Chuanhou Gao





Modeling and Control of Complex Dynamic Systems: Applied Mathematical Aspects

Journal of Applied Mathematics

Modeling and Control of Complex Dynamic Systems: Applied Mathematical Aspects

Guest Editors: Zhiwei Gao, Dexing Kong,
and Chuanhou Gao



Copyright © 2012 Hindawi Publishing Corporation. All rights reserved.

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

Editorial Board

Saeid Abbasbandy, Iran
Mina B. Abd-El-Malek, Egypt
Mohamed A. Abdou, Egypt
Subhas Abel, India
Mostafa Adimy, France
Carlos J. S. Alves, Portugal
Mohamad Alwash, USA
Igor Andrianov, Germany
Sabri Arik, Turkey
Francis T.K. Au, Hong Kong
Olivier Bahn, Canada
Roberto Barrio, Spain
Alfredo Bellen, Italy
J. Biazar, Iran
Hester Bijl, The Netherlands
J. R. Buchanan, USA
A. Cabada, Spain
Xiao C. Cai, USA
Jinde Cao, China
A. Carvalho, Brazil
Song Cen, China
Qianshun S. Chang, China
Shih-sen Chang, China
Tai-Ping Chang, Taiwan
Ke Chen, UK
Xinfu Chen, USA
Rushan Chen, China
Eric Cheng, Hong Kong
Francisco Chiclana, UK
Jen-Tzung Chien, Taiwan
Cheng-Sheng Chien, Taiwan
H. H. Choi, Republic of Korea
Tin-Tai Chow, China
Md S. H. Chowdhury, Malaysia
C. Conca, Chile
Vitor Costa, Portugal
Livija Cveticanin, Serbia
Andrea De Gaetano, Italy
Patrick De Leenheer, USA
Eric de Sturler, USA
Orazio Descalzi, Chile
Kai Diethelm, Germany

Vit Dolejsi, Czech Republic
Magdy A. Ezzat, Egypt
Meng Fan, China
Ya Ping Fang, China
A. Ferreira, Portugal
Michel Fliess, France
M. A. Fontelos, Spain
Luca Formaggia, Italy
Huijun Gao, China
B. Geurts, The Netherlands
Jamshid Ghaboussi, USA
Pablo González-Vera, Spain
Laurent Gosse, Italy
K. S. Govinder, South Africa
Jose L. Gracia, Spain
Yuantong Gu, Australia
Zhihong Guan, China
N. Guglielmi, Italy
F. Guimarães, Brazil
Vijay Gupta, India
Bo Han, China
Maoan Han, Slovenia
Pierre Hansen, Canada
Ferenc Hartung, Hungary
Tasawar Hayat, Pakistan
X. He, Hong Kong
Luis Javier Herrera, Spain
Ying Hu, France
Ning Hu, Japan
Zhilong L. Huang, China
Kazufumi Ito, USA
Takeshi Iwamoto, Japan
George Jaiani, Georgia
Zhongxiao Jia, China
Tarun Kant, India
Ido Kanter, Israel
A. H. Kara, South Africa
Hamid R. Karimi, Norway
Jae-Wook Kim, UK
J. H. Kim, Republic of Korea
K. Komori, Japan
Fanrong Kong, USA

V. Krysko, Russia
Jin L. Kuang, Singapore
M. Lachowicz, Poland
Hak-Keung Lam, UK
Tak-Wah Lam, HongKong
P. Leach, UK
Y. Li, China
Wan-Tong Li, China
Jin Liang, China
Ching-Jong Liao, Taiwan
Chong Lin, China
L. Ling, Hong Kong
Zhijun Liu, China
Chein-Shan Liu, Taiwan
S. Liu, China
Kang Liu, USA
Y. Liu, China
Fawang Liu, Australia
M. Z. Liu, China
Julia'n Lo'pez-Go'mez, Spain
Shiping Lu, China
Gert Lube, Germany
N. I. Mahmudov, Turkey
O. D. Makinde, South Africa
F. J. Marcellán, Spain
Guiomar Martín-Herrán, Spain
N. Mastronardi, Italy
M. McAleer, The Netherlands
S. Metens, France
M. Meylan, Australia
Alain Miranville, France
Jaime E. Munoz Rivera, Brazil
J. Murillo, Spain
R. Natalini, Italy
S. Natesan, India
Jiri Nedoma, Czech Republic
J. Niu, Hong Kong
Roger Ohayon, France
Javier Oliver, Spain
Donal O'Regan, Ireland
M. Ostoja-Starzewski, USA
Turgut Özis, Turkey



Claudio Padra, Argentina
R. M. Palhares, Brazil
F. Pellicano, Italy
J. Manuel Peña, Spain
Ricardo Perera, Spain
M. Peszynska, USA
J. F. Peters, Canada
M. A. Petersen, South Africa
M. Petkovic, Serbia
Vu Ngoc Phat, Vietnam
Andrew Pickering, Spain
H. Pomares, Spain
M. Porfiri, USA
M. Primicerio, Italy
M. Rafei, The Netherlands
B. V. Rathish Kumar, India
Roberto Renò, Italy
Jacek Rokicki, Poland
Dirk Roose, Belgium
Carla Roque, Portugal
D. Roy, India
Samir H. Saker, Egypt
M. A. Savi, Brazil
W. Schmidt, Germany
Eckart Schnack, Germany

Mehmet Sezer, Turkey
Naseer Shahzad, Saudi Arabia
Fateme Shakeri, Iran
Hui-Shen Shen, China
J. Hua Shen, China
F. Simões, Portugal
T. E. Simos, Greece
A.-M. Soliman, Egypt
Qiankun Song, China
Xinyu Song, China
Yuri N. Sotskov, Belarus
P. Spreij, The Netherlands
N. Strömberg, Sweden
Ray KL Su, Hong Kong
Jitao Sun, China
Wenyu Sun, China
XianHua Tang, China
M. H. Terra, Brazil
A. Timokha, Norway
M. Torrisi, Italy
Jung-Fa Tsai, Taiwan
Ch Tsitouras, Greece
K. Vajravelu, USA
A. Valencia, Chile
E. Van Vleck, USA

Ezio Venturino, Italy
J. Vigo-Aguiar, Spain
M. Vrahatis, Greece
Mingxin Wang, China
Baolin Wang, China
Q. Wang, China
Junjie Wei, China
Li Weili, China
Martin Weiser, Germany
Frank Werner, Germany
Shanhe Wu, China
Dongmei Xiao, China
Yuesheng Xu, USA
Suh-Yuh Yang, Taiwan
J. Yuan, Brazil
Alejandro Zarzo, Spain
G. Zhai, Japan
J. Zhan, China
Z. Zhang, China
J. Zhang, Australia
C. Zhao, Australia
Shan Zhao, USA
Renat Zhdanov, USA
H. Zhu, China
J. Hoenderkamp, The Netherlands

Contents

Modeling and Control of Complex Dynamic Systems: Applied Mathematical Aspects,
Zhiwei Gao, Dexing Kong, and Chuanhou Gao
Volume 2012, Article ID 869792, 5 pages

Interval Continuous Plant Identification from Value Sets, R. Hernández, J. A. García,
and C. Mañoso
Volume 2012, Article ID 840603, 32 pages

Improving the Solution of Least Squares Support Vector Machines with Application to a Blast Furnace System, Ling Jian, Shuqian Shen, and Yunquan Song
Volume 2012, Article ID 949654, 12 pages

Adaptive Fault Detection for Complex Dynamic Processes Based on JIT Updated Data Set,
Jinna Li, Yuan Li, Haibin Yu, Yanhong Xie, and Cheng Zhang
Volume 2012, Article ID 809243, 17 pages

Computation of the Added Masses of an Unconventional Airship, Naoufel Azouz,
Said Chaabani, Jean Lerbet, and Azgal Abichou
Volume 2012, Article ID 714627, 19 pages

Modules Identification in Gene Positive Networks of Hepatocellular Carcinoma Using Pearson Agglomerative Method and Pearson Cohesion Coupling Modularity, Jinyu Hu and Zhiwei Gao
Volume 2012, Article ID 248658, 21 pages

Multi-State Dependent Impulsive Control for Pest Management, Huidong Cheng, Fang Wang,
and Tongqian Zhang
Volume 2012, Article ID 381503, 25 pages

Nonlinearities in Drug Release Process from Polymeric Microparticles: Long-Time-Scale Behaviour, Elena Simona Bacaita, Costica Bejinariu, Borsos Zoltan, Catalina Peptu,
Gabriela Andrei, Marcel Popa, Daniela Magop, and Maricel Agop
Volume 2012, Article ID 653720, 26 pages

Application of the Poor Man's Navier-Stokes Equations to Real-Time Control of Fluid Flow,
James B. Polly and J. M. McDonough
Volume 2012, Article ID 653720, 18 pages

Data Fusion Based Hybrid Approach for the Estimation of Urban Arterial Travel Time,
S. P. Anusha, R. A. Anand, and L. Vanajakshi
Volume 2012, Article ID 587913, 17 pages

The Chaotic Prediction for Aero-Engine Performance Parameters Based on Nonlinear PLS Regression, Chunxiao Zhang and Junjie Yue
Volume 2012, Article ID 615618, 14 pages

Adaptive Control of a Two-Item Inventory Model with Unknown Demand Rate Coefficients,
Ahmad M. Alshamrani
Volume 2012, Article ID 810635, 16 pages

Optimal Control of a Spatio-Temporal Model for Malaria: Synergy Treatment and Prevention, Malicki Zorom, Pascal Zongo, Bruno Barbier, and Blaise Somé
Volume 2012, Article ID 854723, 20 pages

Dynamic Analysis of a Predator-Prey (Pest) Model with Disease in Prey and Involving an Impulsive Control Strategy, Min Zhao, Yanzhen Wang, and Lansun Chen
Volume 2012, Article ID 969425, 18 pages

Pareto Design of Decoupled Sliding-Mode Controllers for Nonlinear Systems Based on a Multiobjective Genetic Algorithm, M. J. Mahmoodabadi, A. Bagheri, N. Nariman-zadeh, A. Jamali, and R. Abedzadeh Maafi
Volume 2012, Article ID 639014, 22 pages

Synchronization of Coupled Networks with Mixed Delays by Intermittent Control, Kun Yuan, Jinde Cao, and Shumin Fei
Volume 2012, Article ID 927609, 13 pages

Dynamic Analysis of a Hybrid Squeeze Film Damper Mounted Rub-Impact Rotor-Stator System, Cai-Wan Chang-Jian
Volume 2012, Article ID 279827, 21 pages

Nonsmooth Recursive Identification of Sandwich Systems with Backlash-Like Hysteresis, Ruili Dong, Yonghong Tan, Hui Chen, and Yangqiu Xie
Volume 2012, Article ID 457603, 16 pages

The Hybrid Adaptive Control of T-S Fuzzy System Based on Niche, Tong Zhao, Guo-ping Lu, Yun-li Hao, and Yi-min Li
Volume 2012, Article ID 158720, 15 pages

Modeling and Analysis of Epidemic Diffusion within Small-World Network, Ming Liu and Yihong Xiao
Volume 2012, Article ID 841531, 14 pages

Stability and Limit Oscillations of a Control Event-Based Sampling Criterion, M. De la Sen, J. C. Soto, and A. Ibeas
Volume 2012, Article ID 684292, 25 pages

The Analysis and Control for Singular Ecological-Economic Model with Harvesting and Migration, Qingling Zhang, Hong Niu, Lichun Zhao, and Fenglan Bai
Volume 2012, Article ID 973869, 17 pages

A Matrix Method for Determining Eigenvalues and Stability of Singular Neutral Delay-Differential Systems, Jian Ma, Baodong Zheng, and Chunrui Zhang
Volume 2012, Article ID 749847, 11 pages

Type-K Exponential Ordering with Application to Delayed Hopfield-Type Neural Networks, Bin-Guo Wang
Volume 2012, Article ID 580482, 9 pages

Contents

Mean Square Almost Periodic Solutions for Impulsive Stochastic Differential Equations with Delays, Ruojun Zhang, Nan Ding, and Linshan Wang
Volume 2012, Article ID 414320, 14 pages

Bank Liquidity and the Global Financial Crisis, Frednard Gideon, Mark A. Petersen, Janine Mukuddem-Petersen, and Bernadine De Waal
Volume 2012, Article ID 743656, 27 pages

Optimal Control for a Class of Chaotic Systems, Jianxiong Zhang and Wansheng Tang
Volume 2012, Article ID 859542, 20 pages

Bifurcation Analysis for a Kind of Nonlinear Finance System with Delayed Feedback and Its Application to Control of Chaos, Rongyan Zhang
Volume 2012, Article ID 316390, 18 pages

Modeling Optimal Scheduling for Pumping System to Minimize Operation Cost and Enhance Operation Reliability, Yin Luo, Shouqi Yuan, Yue Tang, Jianping Yuan, and Jinfeng Zhang
Volume 2012, Article ID 370502, 19 pages

Editorial

Modeling and Control of Complex Dynamic Systems: Applied Mathematical Aspects

Zhiwei Gao,¹ Dexing Kong,² and Chuanhou Gao²

¹ *Division of Electric Engineering and Physics, Faculty of Engineering and Environment,
Northumbria University, Newcastle upon Tyne NE1 8ST, UK*

² *Department of Mathematics, College of Science, Zhejiang University, Hang Zhou 310027, China*

Correspondence should be addressed to Zhiwei Gao, zhiwei.gao@northumbria.ac.uk

Received 25 November 2012; Accepted 25 November 2012

Copyright © 2012 Zhiwei Gao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

The concept of complex dynamic systems arises in many varieties, including the areas of energy generation, storage and distribution, ecosystems, gene regulation and health delivery, safety and security systems, telecommunications, transportation networks, and the rapidly emerging research topics seeking to understand and analyse. Such systems are often concurrent and distributed, because they have to react to various kinds of events, signals, and conditions. They may be characterized by a system with uncertainties, time delays, stochastic perturbations, hybrid dynamics, distributed dynamics, chaotic dynamics, and a large number of algebraic loops. This special issue provides a platform for researchers to report their recent results on various mathematical methods and techniques for modelling and control of complex dynamic systems and identifying critical issues and challenges for future investigation in this field. This special issue amazingly attracted one-hundred-and-eighteen submissions, and twenty-seven of them are selected through a rigorous review procedure.

The selected papers contribute mathematical modelling, parameter identification, monitoring and diagnosis, optimization, and control for a variant of complex systems such as chaotic systems (4 papers), impulsive and singular systems (4 papers), nonlinear systems (12 papers), delay systems (3 papers), stochastic systems (3 papers), and gene complex network (1 paper). From the viewpoint of modelling, the related papers mainly investigate mathematical modelling for complex systems by using physical laws or deal with model identification by using the data based training and estimation. Along with the dynamic

equation models, graph-based model is of interest for modelling complex gene regulation network. From the viewpoint of control, the adopted techniques mainly include adaptive control, impulsive control, sliding mode control, fuzzy control, nonlinear optimization, and optimal control. From the view of application, the involved systems include mechanical dynamic systems, biological systems, aerospace systems, fluid dynamic systems, energy systems, and ecological-economic systems.

We are going to introduce the selected papers in the categories of chaotic systems, impulsive and singular systems, nonlinear systems/processes, time-delay systems, stochastic systems, and gene complex network.

2. Modelling and Control for Chaotic Systems

For modelling and control of chaotic systems, the paper entitled *"The chaotic prediction for aero-engine performance parameters based on nonlinear PLS regression"* is contributed by C. Zhang and J. Yue. On the basis of the actual air-borne data of the aircraft communications addressing and reporting system, a nonlinear partial least square regression method is proposed to predict chaotic function of the engine exhaust temperature time series of the aero-engine, which will benefit the condition monitoring and diagnosis for aero engines.

The paper entitled *"Dynamic analysis of a hybrid squeeze film damper mounted rub-impact rotor-stator system,"* authored by C. C.-Jian, presents a detailed analysis on dynamic behaviour of a rotor-stator system. It reveals that its dynamic behaviours may suddenly evolve into chaos from periodic motions without any transition. The presented results provide useful insights into the design and development of a rotor-bearing complex system.

The paper entitled *"Optimal control for a class of chaotic systems,"* authored by J. Zhang and W. Tang, presents an optimal state-feedback control strategy for a class of chaotic systems. The system considered is transformed into a set of uncertain piecewise linear systems, and an optimal robust controller is designed by solving an optimisation problem under the constraints of nonlinear inequalities.

In the paper entitled *"Bifurcation analysis for a kind of nonlinear finance system with delayed feedback and its application to control of chaos,"* R. Zhang presents a quantitative distribution analysis on the roots of the associate characteristic equation for a kind of nonlinear finance systems with time-delay feedback. The conditions of ensuring the existence of Hopf bifurcation are addressed and the explicit formulas are derived for determining the stability and direction of the bifurcating periodic solutions.

3. Modelling and Control of Impulsive and Singular Systems

For modelling and control of impulsive and singular systems, the paper entitled *"The analysis and control for singular ecological-economic model with harvesting and migration"* is contributed by Q. Zhang et al. This paper presents a singular ecological-economic model for the populations with harvesting and migration. The local stability and dynamic behaviour of the model are investigated. By transforming the singular model into a single-input single-output model, variable structure control is applied to eliminate the complex behaviours of the considered problem.

The paper entitled *"Multi-state dependent impulsive control for pest management,"* authored by H. Cheng et al., presents a combined control strategy for pest management. The combined control strategy adopts different control methods according to different thresholds.

The existence and convergence of the first-order periodic solution of such systems are further analysed.

The paper entitled *“Dynamic analysis of a predator-prey (Pest) model with disease in prey and involving an impulsive control strategy,”* is contributed by M. Zhao et al., which presents a dynamic analysis of a predator-prey (pest) model with disease in prey. The model involves an impulsive control strategy to release infected prey at fixed times. Theoretical analysis on the local asymptotical stability and global attractivity for the semitrivial periodic solution is also presented.

In paper entitled *“A matrix method for determining eigenvalues and stability of singular neutral delay-differential systems,”* J. Ma et al. present an analysis on the eigenvalues and stability of singular neutral delay-differential systems. The matrix pencil and linear operator methods are used to derive a new algebraic criterion to obtain imaginary axis eigenvalues. In addition, a criterion for the asymptotic stability is introduced.

4. Modelling and Control of Nonlinear Systems/Processes

In the paper entitled *“Adaptive fault detection for complex dynamic processes based on jit updated data set,”* J. Li et al. presents a fault detection technique for nonlinear complex industrial processes. Just-in-time and k-nearest neighbour methods are integrated to construct a flexible and adaptive detection scheme. The updating of the database is taken into account in the design of fault detection algorithm.

The paper, *“Computation of the added masses of an unconventional airship,”* authored by N. Azouz et al., presents a modelling method for an unmanned airship. An aerodynamic investigation is firstly presented and a mathematical analysis of the velocity potential flow of the shape of the airship is developed. The whole system, described by nonlinear equations subjected to boundary conditions, governing the interaction air-structure, is solved in an analytical setting.

In the paper entitled *“Nonlinearities in drug release process from polymeric microparticles: long-time-scale behavior,”* E. S. Bacaita et al. present a theoretical model of the drug release process from polymeric microparticles through dispersive fractal approximation of motion. The drug release process takes place through cnoidal oscillations modes of a normalized concentration field, which indicates that the drug particles assemble in a lattice of nonlinear oscillators occur macroscopically.

The paper *“Application of the poor man’s navier-stokes equations to real-time control of fluid flow,”* authored by J. M. Polly and J. M. McDonough, presents an application study of the Poor Man’s Navier Stokes (PMNS) equations to real time control of fluid flow. The PMNS equations comprise a discrete dynamic system that is algebraic so that they can be easily and rapidly solved. Theoretical analysis of the PMNS equations is discussed and a control force is added to control the fluid flow.

In the paper entitled *“Data fusion based hybrid approach for the estimation of urban arterial travel time,”* S. P. Anusha et al. present a data based approach to estimate the travel time in urban arterials in India. Three different approaches are proposed, namely, highway capacity manual based method, Kalman filter based method and hybrid method.

In the paper entitled, *“Adaptive control of a two-item inventory model with unknown demand rate coefficients,”* Alshamrani presents a multiitem inventory model with unknown demand coefficients. The output of the system is tracked by an adaptive control approach with a nonlinear feedback. The asymptotic stability of the adaptive controlled system is proven by the Lyapunov technique.

In the paper entitled, *"Optimal control of a spatio-temporal model for malaria: synergy treatment and prevention,"* M. Zorom et al. present a metapopulation model for malaria with two control variables, treatment and prevention, distributed between n different patches. A mathematical model with three patches is constructed using optimal control techniques. Based on the model, qualitative suggestions for treatment and prevention of Malaria are provided.

In the paper entitled *"Nonsmooth recursive identification of sandwich systems with backlash-like hysteresis,"* R. Dong et al. present a recursive gradient identification algorithm based on the bundle method for sandwich systems with backlash-like hysteresis. A dynamic parameter estimation scheme is developed to handle the nonsmooth problem caused by the backlash embedded in the system. The bundle method is used to estimate the search direction of the subgradient algorithm.

In the paper entitled *"The hybrid adaptive control of T-S fuzzy system based on niche,"* T. Zhao et al. presents a hybrid adaptive fuzzy control method with the function of continuous supervisory control. The method is based on the niche characteristics, whose control law is designed by tracking, continuous supervisory, and adaptive compensation. An adaptive compensator and adaptive laws of parameters are designed to embody system adaptability.

The paper entitled *"Stability and limit oscillations of a control event-based sampling criterion,"* authored by M. De La Sen et al., investigates the presence of limit oscillations in an adaptive sampling system. An extended sampling criterion is proposed and the limit oscillations are interpreted on the basis of adaptive sampling in nonlinear dynamic systems.

In the paper entitled *"Modeling optimal scheduling for pumping system to minimize operation cost and enhance operation reliability,"* Y. Luo et al. present an optimal scheduling plan for pumping system to minimize operation cost and enhance operation reliability. The operation reliability is explored by investigating vibration level and a model is built as a function of the capacity and rotation speed of the pump for the operation reliability.

In the paper entitled *"Improving the solution of least squares support vector machines with application to a blast furnace system,"* L. Jian et al. present a strategy to speed up the search of the solution of the least square support vector machine (LSSVM) by using minimal residual method to solve saddle point systems in LSSVM directly. The proposed method is validated by a blast furnace industrial process.

5. Modelling and Control of Systems with Time Delays

In the paper, entitled *"Modelling and analysis of epidemic diffusion within small-world network,"* M. Liu and Y. Xiao present two different models, Susceptible-Exposure-Infected-Recovered-Susceptible model and Susceptible-Exposure-Infected-Quarantine-Recovered-Susceptible model, to describe the rule of epidemic diffusion. The models are analysed within the small-world network. Theoretical properties such as condition for occurrence of disease diffusion, existence, and global stability of the disease free equilibrium are addressed.

In the paper entitled *"Synchronization of coupled networks with mixed delays by intermittent control,"* the synchronization of coupled networks with mixed delay is investigated by employing Lyapunov functional method and intermittent control. A sufficient condition is derived to ensure the global synchronization of coupled networks and an intermittent controller is designed to control the system.

In the paper entitled *"Type-K exponential ordering with application to delayed hopfield-type neural networks,"* B.-W. Wang presents order-preserving and convergent results for delay functional differential equations without quasi-monotone condition under type-K

exponential ordering. The delayed Hopfield-type neural networks with a type K monotone interconnection matrix are considered as an application and the attractor is obtained.

6. Modelling and Control of Stochastic Systems

For modelling and control of stochastic systems, the paper entitled "*Pareto design of decoupled sliding-mode controllers for nonlinear systems based on a multiobjective genetic algorithm*," authored by M. J. Mahmoodabadi et al., presents the pareto design of decoupled sliding-mode controller based on multiobjective genetic algorithm for several fourth-order coupled nonlinear systems. The decoupled sliding mode controller is applied to stabilize the fourth-order coupled systems at the equilibrium point. The optimal parameters of the decoupled sliding-mode controller are obtained using the multiobjective genetic algorithm.

In the paper entitled "*Mean square almost periodic solutions for impulsive stochastic differential equations with delays*," R. Zhang et al. present a result on existence and uniqueness of mean square almost periodic solutions for a class of impulsive stochastic differential equations with delay. This is an extension of the earlier works reported in the literature.

In the paper entitled "*Bank liquidity and the global financial crisis*," F. Gideon et al. present an analysis on the stochastic dynamics of bank liquidity parameters such as liquid assets and net cash outflow in the case of global financial crisis. Numerical results relating to bank behaviour are obtained and a theoretical-quantitative approach to bank liquidity provisioning is provided.

7. Modelling and Regulation of Gene Complex Network

The paper entitled "*Modules identification in gene positive networks of hepatocellular carcinoma using pearson agglomerative method and pearson cohesion coupling modularity*" is contributed by J. Hu and Z. Gao. The paper firstly addresses the concepts of gene network community and gene positive network on the basis of weighted undirected graphs. A module identification algorithm is developed to calculate the threshold values of gene positive networks. The fourteen strong modules and thirteen very strong modules are obtained and the relations between these modules are analysed. The biological significance of these modules is explained. The research may provide new clues and ideas for liver cancer treatment.

The special issue involves a variety of research issues for complex systems, which provides a platform for the research community to share their knowledge and experience in complex systems modelling, dynamic analyses, optimization, and control.

Acknowledgment

Finally the guest editors would like to take this opportunity to thank all the contributions from the authors, reviewers, and journal editorial staff.

Zhiwei Gao
Dexing Kong
Chuanhou Gao

Research Article

Interval Continuous Plant Identification from Value Sets

R. Hernández,¹ J. A. García,² and C. Mañoso¹

¹ Departamento de Sistemas de Comunicación y Control, UNED, c/Juan del Rosal 16, 28040 Madrid, Spain

² Departamento de Tecnología de Computadores y Comunicaciones, Universidad de Extremadura, 06800 Madrid, Spain

Correspondence should be addressed to R. Hernández, roberto@scc.uned.es

Received 13 April 2012; Revised 1 September 2012; Accepted 1 September 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 R. Hernández et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper shows how to obtain the values of the numerator and denominator Kharitonov polynomials of an interval plant from its value set at a given frequency. Moreover, it is proven that given a value set, all the assigned polynomials of the vertices can be determined if and only if there is a complete edge or a complete arc lying on a quadrant. This algorithm is nonconservative in the sense that if the value-set boundary of an interval plant is exactly known, and particularly its vertices, then the Kharitonov rectangles are exactly those used to obtain these value sets.

1. Introduction

In reference to the identification problem, these have been widely motivated and analysed over recent years [1]. Van Overschee and De Moor in [2] explains a subspace identification algorithm. In [3] the authors present a robust identification procedure for a priori classes of models in H_∞ ; the authors consider casual, linear time invariant, stable, both continuous or discrete time models, and only SISO systems.

Interval plants have been widely motivated and analysed over recent years. For further engineering motivation, among the numerous papers and books, [4–9] must be pointed out and the references thereof.

The identification problem using the interval plant framework, that is, to compute an interval plant from the frequency response, has not been completely solved. Interval plant identification was investigated by Bhattacharyya et al. [5], who developed a method in which identification is carried out for interval plants so that the numerator and denominator have the same degree, starting from the variation of the coefficient values of a nominal transfer

function at certain intervals. So, the identification of a nominal transfer function is carried out first, and then the intervals of variation of the coefficients are determined.

A different approach was developed by Hernández et al. [10] studying the problem from the extreme point results point of view. This was a first step for the identification of an interval plant, showing three main properties to characterize the value set lying on a quadrant. Then an algorithm for the identification of interval plants from the vertices of the value sets is obtained. However, this algorithm solves the identification problem when the value set contains at least five vertices in a quadrant.

This paper improves the results obtained in [10] and shows how to obtain the values of the numerator and denominator Kharitonov polynomials when the value sets have less than five vertices in the same quadrant. Identification with such an interval plant allows engineers predict the worst case performance and stability margins using the results on interval systems, particularly extreme point results.

2. Problem Statement

Let us consider a linear interval plant of real coefficients, of the form

$$P(s, a, b) = \frac{N_p(s, a)}{D_p(s, b)}, \quad (2.1)$$

where $N_p(s, a)$ and $D_p(s, b)$ are interval polynomials given as

$$\begin{aligned} N_p(s, a) &= a_m s^m + a_{m-1} s^{m-1} + \dots + a_0, \quad a \in A = \{a : a_i^- \leq a_i \leq a_i^+, i = 0, \dots, m\}, \\ D_p(s, b) &= b_n s^n + b_{n-1} s^{n-1} + \dots + b_0, \quad b \in B = \{b : b_i^- \leq b_i \leq b_i^+, i = 0, \dots, n\}, \end{aligned} \quad (2.2)$$

with $m \geq 1, n \geq 1, 0 \notin D_p(s, b)$, and where vectors $a = [a_0, a_1, \dots, a_m]$, $a_m \neq 0$, and $b = [b_0, b_1, \dots, b_n]$, $b_n \neq 0$ are the uncertainty parameters that lie in the hyperrectangles A and B , respectively.

Numerator and denominator polynomial families are characterized by their respective Kharitonov polynomials, and they can be expressed in terms of their even and odd parts, at $s = j\omega$, as follows:

Family $N_p(s)$:

$$\begin{aligned} k_{n1} &= p_{e \min}(j\omega) + jp_{o \min}(j\omega), & k_{n2} &= p_{e \max}(j\omega) + jp_{o \min}(j\omega), \\ k_{n3} &= p_{e \max}(j\omega) + jp_{o \max}(j\omega), & k_{n4} &= p_{e \min}(j\omega) + jp_{o \max}(j\omega), \end{aligned} \quad (2.3)$$

where

$$\begin{aligned} p_{e \min}(j\omega) &= a_0^- - a_2^+ \omega^2 + a_4^- \omega^4 - a_6^+ \omega^6 + \dots, & p_{e \max}(j\omega) &= a_0^+ - a_2^- \omega^2 + a_4^+ \omega^4 - a_6^- \omega^6 + \dots, \\ p_{o \min}(j\omega) &= a_1^- \omega - a_3^+ \omega^3 + a_5^- \omega^5 - a_7^+ \omega^7 + \dots, & p_{o \max}(j\omega) &= a_1^+ \omega - a_3^- \omega^3 + a_5^+ \omega^5 - a_7^- \omega^7 + \dots. \end{aligned} \quad (2.4)$$

Family $D_p(s)$:

$$\begin{aligned} k_{d1} &= q_{e \min}(j\omega) + jq_{o \min}(j\omega), & k_{d2} &= q_{e \max}(j\omega) + jq_{o \min}(j\omega), \\ k_{d3} &= q_{e \max}(j\omega) + jq_{o \max}(j\omega), & k_{d4} &= q_{e \min}(j\omega) + jq_{o \max}(j\omega), \end{aligned} \quad (2.5)$$

where

$$\begin{aligned} q_{e \min}(j\omega) &= b_0^- - b_2^+ \omega^2 + b_4^- \omega^4 - b_6^+ \omega^6 + \dots, & q_{e \max}(j\omega) &= b_0^+ - b_2^- \omega^2 + b_4^+ \omega^4 - b_6^- \omega^6 + \dots, \\ q_{o \min}(j\omega) &= b_1^- \omega - b_3^+ \omega^3 + b_5^- \omega^5 - b_7^+ \omega^7 + \dots, & q_{o \max}(j\omega) &= b_1^+ \omega - b_3^- \omega^3 + b_5^+ \omega^5 - b_7^- \omega^7 + \dots. \end{aligned} \quad (2.6)$$

As is well known, the values $G(j\omega)$ of the complex plane obtained for the transfer function $G(s)$ at a given frequency are denominated as a *value set*. The identification of the system consists in determining the transfer function coefficients from the value set.

As can be observed in [10], when the values $\{k_{n1}(j\omega), k_{n2}(j\omega), k_{n3}(j\omega), k_{n4}(j\omega)\}$ and $\{k_{d1}(j\omega), k_{d2}(j\omega), k_{d3}(j\omega), k_{d4}(j\omega)\}$ are known, then the system of equations given in [10, equation 14] can be solved and therefore the interval plant is identified (see [10] for details).

As is shown [10] the vertices of the value-set boundary of an interval plant can be assigned as

$$v_i = \frac{n_j}{d_k}, \quad (2.7)$$

where n_j , $j = 1, 2, 3, 4$ and d_k , $k = 1, 2, 3, 4$ are the assigned polynomials numerator and denominator, respectively. When they are in the same quadrant they are a *Sorted Set of Vertices* (SSV).

As is well known, the Kharitonov polynomials values can be obtained from

$$\begin{aligned} k_{n1}(j\omega) &= \min[\operatorname{Re}(n_1, n_3)] + j \min[\operatorname{Im}(n_1, n_3)], \\ k_{n2}(j\omega) &= \max[\operatorname{Re}(n_1, n_3)] + j \min[\operatorname{Im}(n_1, n_3)], \\ k_{n3}(j\omega) &= \max[\operatorname{Re}(n_1, n_3)] + j \max[\operatorname{Im}(n_1, n_3)], \\ k_{n4}(j\omega) &= \min[\operatorname{Re}(n_1, n_3)] + j \max[\operatorname{Im}(n_1, n_3)], \\ k_{d1}(j\omega) &= \min[\operatorname{Re}(d_1, d_3)] + j \min[\operatorname{Im}(d_1, d_3)], \\ k_{d2}(j\omega) &= \max[\operatorname{Re}(d_1, d_3)] + j \min[\operatorname{Im}(d_1, d_3)], \\ k_{d3}(j\omega) &= \max[\operatorname{Re}(d_1, d_3)] + j \max[\operatorname{Im}(d_1, d_3)], \\ k_{d4}(j\omega) &= \min[\operatorname{Re}(d_1, d_3)] + j \max[\operatorname{Im}(d_1, d_3)]. \end{aligned} \quad (2.8)$$

It must be pointed out that the results presented in [10] must be considered as the background necessary for this work. Thus, the geometry of the value set is described in [10] and the concepts necessary for its description are defined, (such as the successor, predecessor element, etc.) and the fundamental properties on which this work is based are proven.

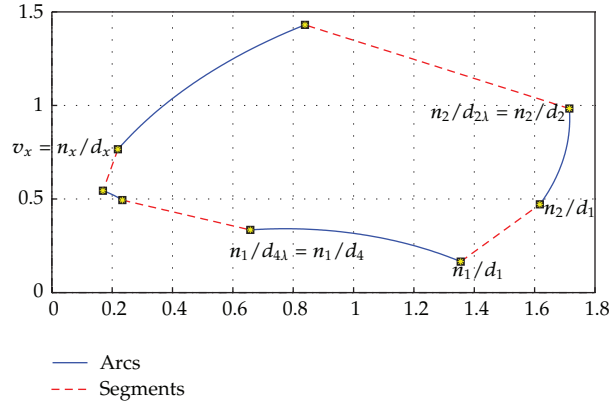


Figure 1: segment and complete arcs.

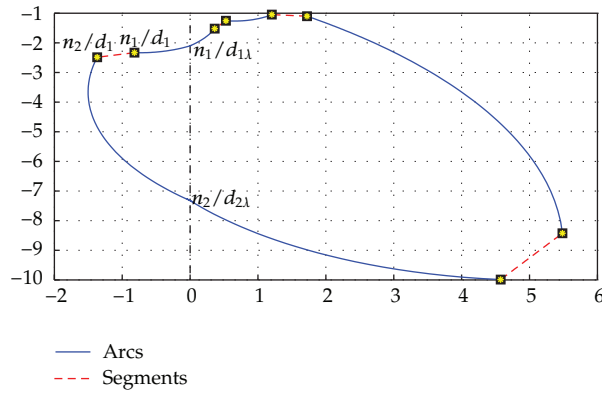


Figure 2: Segment and no complete arcs.

This paper is organized as follows. Section 3 shows how to determine the assigned polynomial with the only condition that there is a complete segment in a quadrant. Similarly Section 4 shows it when there is an arc in a quadrant. Section 5 illustrates the algorithm and examples. Finally, the conclusions are shown in Section 6.

3. Assigned Polynomial Determination When There Is a Complete Segment in a Quadrant

In order to determine the polynomials numerator and denominator associated to a vertex of the value set boundary with the minimum number of elements, the situation of a segment in a quadrant will be considered. So, let S_1 be a segment of the value-set boundary with vertices $v_1 = n_1/d_1$ and $v_2 = n_2/d_1$. Continuity segment-arc in a quadrant (see [10, Theorem 2]) implies that there will be a successor arc with vertices $v_2 = n_2/d_1$, $v_{2\text{succ}} = n_2/d_{2\lambda}$ counter-clockwise and a predecessor arc with vertices $v_{1\text{pred}} = n_1/d_{4\lambda}$ counter-clockwise. When these arcs are completed the denominators are vertices of the Kharitonov rectangle. Figures 1 and 2 show this situation.

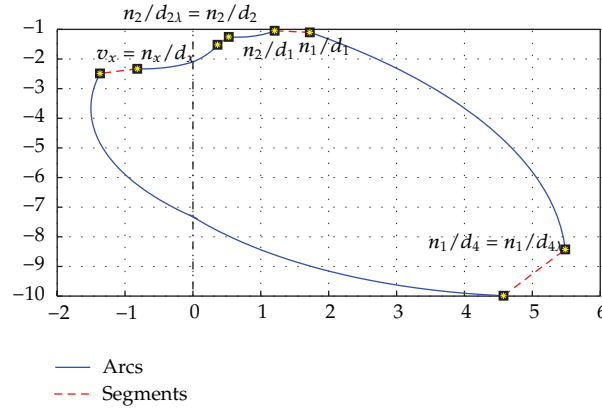


Figure 3: v_x vertex of two elements, arc-segment.

As was shown, the values of n_1 , n_2 , and d_1 can be calculated from the complete segment based on a normalization (see [10, Theorem 4]). The following normalization simplifies the nomenclature.

Lemma 3.1 (segment normalization). *Let S_1 be a complete segment of the value-set boundary with vertices $v_1 = n_1/d_1$ and $v_2 = n_2/d_1$ and the normalization $d_1 = \cos(\varphi(d_1)) + j \sin(\varphi(d_1))$, where $\varphi(d_1) = 360^\circ - \arg(v_2 - v_1)$ $\arg(v_2 - v_1)$ being the argument of the segment $v_2 - v_1$. Then $n_1 = v_1 d_1$, $n_2 = v_2 d_1$, $d_{2\lambda} = n_2/v_{2 \text{ succ}}$, and $d_{4\lambda} = n_1/v_{1 \text{ pred}}$, where $v_{2 \text{ succ}}$ ($v_{1 \text{ pred}}$) is any point of the next (previous) arc of the segment S_1 .*

Proof. It is trivial. This normalization is one of the infinite possible solutions [10] for a value set. This normalization implies fitting d_1 with modulus $|d_1| = 1$ and angle so that the segment of the Kharitonov polynomial numerator with vertices n_1 and n_2 will be parallel to the real axis counter-clockwise. Thus, from the information with a complete segment in a quadrant the values of d_1 , n_1 , n_2 , $d_{2\lambda}$, and $d_{4\lambda}$ can be calculated. \square

This paper deals with the general case where $n_{2R} \neq 0$, $n_{2I} \neq 0$, $n_{1R} \neq 0$, and $n_{1I} \neq 0$.

Given a vertex $v_x = n_x/d_x$ in a quadrant, the target is to determine the polynomials n_x and d_x . The vertex v_x belongs to a part of a segment and a part of an arc, due to the continuity segment-arc in a quadrant. So, v_x will be the vertex of two elements, arc-segment (Figure 3) or segment-arc (Figure 4).

The following Lemma shows the necessary conditions on the denominator d_x to be a solution of $v_x = n_x/d_x$.

Lemma 3.2 (denominator condition). *Let S_1 be a complete segment in a quadrant and let d_x be the denominator of a vertex $v_x = n_x/d_x$ in a quadrant. Then it is a necessary condition that d_x satisfies one of the following conditions:*

- (1) $(d_{1R} < d_{2\lambda R} \text{ and } d_{1I} < d_{4\lambda I}) \text{ and } \{(d_{xR} = d_{1R} \text{ and } d_{xI} = d_{1I}) [d_x = d_1] \text{ or } (d_{xR} = d_{1R} \text{ and } d_{xI} \geq d_{1I}) [d_x = d_4] \text{ or } (d_{xI} = d_{1I} \text{ and } d_{xR} \geq d_{1R}) [d_x = d_2] \text{ or } (d_{xR} > d_{1R} \text{ and } d_{xI} > d_{1I}) [d_x = d_3]\}$,

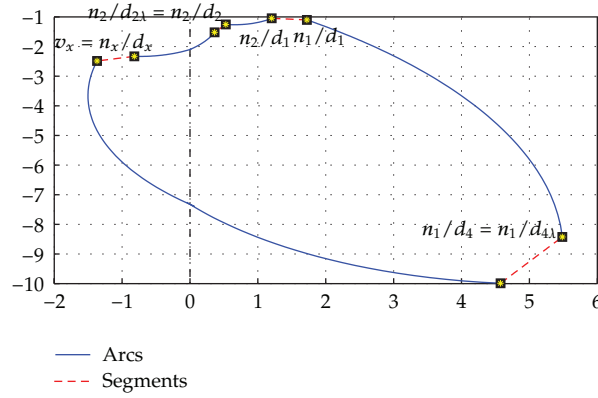


Figure 4: v_x vertex of two elements, segment-arc.

- (2) ($d_{1R} > d_{4\lambda R}$ and $d_{1I} < d_{2\lambda I}$) and $\{(d_{xR} = d_{1R}$ and $d_{xI} = d_{1I})$ [$d_x = d_1$] or ($d_{xR} = d_{1R}$ and $d_{xI} \geq d_{1I}$) [$d_x = d_2$] or ($d_{xI} = d_{1I}$ and $d_{xR} \leq d_{1R}$) [$d_x = d_4$] or ($d_{xR} < d_{1R}$ and $d_{xI} > d_{1I}$) [$d_x = d_3$]],
- (3) ($d_{1R} > d_{2\lambda R}$ and $d_{1I} > d_{4\lambda I}$) and $\{(d_{xR} = d_{1R}$ and $d_{xI} = d_{1I})$ [$d_x = d_1$] or ($d_{xR} = d_{1R}$ and $d_{xI} \leq d_{1I}$) [$d_x = d_4$] or ($d_{xI} = d_{1I}$ and $d_{xR} \leq d_{1R}$) [$d_x = d_2$] or ($d_{xR} < d_{1R}$ and $d_{xI} < d_{1I}$) [$d_x = d_3$]],
- (4) ($d_{1R} < d_{4\lambda R}$ and $d_{1I} > d_{2\lambda I}$) and $\{(d_{xR} = d_{1R}$ and $d_{xI} = d_{1I})$ [$d_x = d_1$] or ($d_{xR} = d_{1R}$ and $d_{xI} \leq d_{1I}$) [$d_x = d_2$] or ($d_{xI} = d_{1I}$ and $d_{xR} \geq d_{1R}$) [$d_x = d_4$] or ($d_{xR} > d_{1R}$ and $d_{xI} < d_{1I}$) [$d_x = d_3$]],

where d_{iR} is the real part of d_i and d_{iI} is the imaginary part of d_i , and the corresponding assigned denominator is shown between brackets.

Proof. The proof is obtained directly from the information of a complete segment in a quadrant and the properties of the Kharitonov rectangle. So, from the complete segment and the normalization (Lemma 3.1), the values of d_1 , $d_{2\lambda}$, and $d_{4\lambda}$ are known. Then, d_1 can be established as k_{d1} , k_{d2} , k_{d3} , or k_{d4} .

- (1) If ($d_{1R} < d_{2\lambda R}$ and $d_{1I} < d_{4\lambda I}$) then d_1 is k_{d1} . Given a value d_x , it will be a vertex of the Kharitonov rectangle denominator only if $d_{xR} = d_{1R}$ and $d_{xI} = d_{1I}$ (d_x is $d_1 = k_{d1}$) or $d_{xR} = d_{1R}$ and $d_{xI} > d_{1I}$ (d_x is $d_4 = k_{d4}$) or $d_{xI} = d_{1I}$ and $d_{xR} > d_{1R}$ (d_x is $d_2 = k_{d2}$) or $d_{xR} > d_{1R}$ and $d_{xI} > d_{1I}$ (d_x is $d_3 = k_{d3}$). (Figures 5(a), 5(b), 5(c), and 5(d)).

Note that if any of these conditions is not satisfied, then d_x cannot be a solution. For example, if $d_{xR} = d_{1R}$ and $d_{xI} < d_{1I}$, d_x does not belong to the rectangle with vertex d_1 , $d_{2\lambda}$, and $d_{4\lambda}$ are elements of the successor and predecessor edges. Figure 6 shows these considerations.

- (2) Similarly, if ($d_{1R} > d_{4\lambda R}$ and $d_{1I} < d_{2\lambda I}$) then d_1 is k_{d2} . Given a value d_x , it will be a vertex of the Kharitonov rectangle denominator only if $d_{xR} = d_{1R}$ and $d_{xI} = d_{1I}$ (d_x is $d_1 = k_{d2}$) or $d_{xR} = d_{1R}$ and $d_{xI} > d_{1I}$ (d_x is $d_2 = k_{d3}$) or $d_{xI} = d_{1I}$ and $d_{xR} < d_{1R}$ (d_x is $d_4 = k_{d1}$) or $d_{xR} < d_{1R}$ and $d_{xI} > d_{1I}$ (d_x is $d_3 = k_{d4}$).
- (3) If $d_{1R} > d_{2\lambda R}$ and $d_{1I} > d_{4\lambda I}$ then d_1 is k_{d3} . Given a value d_x , it will be a vertex of the Kharitonov rectangle denominator only if $d_{xR} = d_{1R}$ and $d_{xI} = d_{1I}$ (d_x is

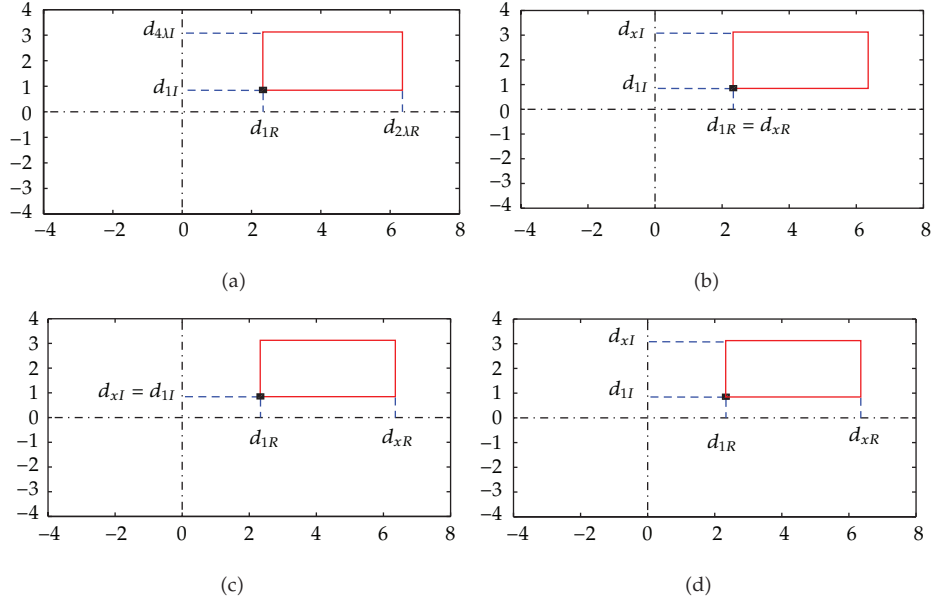


Figure 5: Cases where d_x is a vertex of the kharitonov rectangle denominator.

$d_1 = k_{d3}$) or $d_{xR} = d_{1R}$ and $d_{xI} < d_{1I}$ (d_x is $d_4 = k_{d2}$) or $d_{xI} = d_{1I}$ and $d_{xR} < d_{1R}$ (d_x is $d_2 = k_{d4}$) or $d_{xR} < d_{1R}$ and $d_{xI} < d_{1I}$ (d_x is $d_3 = k_{d1}$).

- (4) Finally, if $d_{1R} < d_{4LR}$ and $d_{1I} > d_{2LI}$ then d_1 is k_{d4} . Given a value d_x , it will be a vertex of the Kharitonov rectangle denominator only if $d_{xR} = d_{1R}$ and $d_{xI} = d_{1I}$ (d_x is $d_1 = k_{d4}$) or $d_{xR} = d_{1R}$ and $d_{xI} < d_{1I}$ (d_x is $d_2 = k_{d1}$) or $d_{xI} = d_{1I}$ and $d_{xR} > d_{1R}$ (d_x is $d_4 = k_{d3}$) or $d_{xR} > d_{1R}$ and $d_{xI} < d_{1I}$ (d_x is $d_3 = k_{d2}$). \square

On the other hand, the behaviour of a segment on the complex plane when divided by a complex number is well known. The following property shows this behaviour.

Property 1. Let $S_x = S/d_x$ be a segment on the complex plane with vertices v_{x1} and v_{x2} counter-clockwise where S is a segment with vertices n_a and n_b counter-clockwise. Let d_x be a complex number with argument $\arg(d_x)$. Let $\varphi(S_x)$ be $\varphi(S_x) \equiv \arg(v_{x2} - v_{x1})$. Then the relation between the argument of d_x and $\varphi(S_x)$, is given by

- (1) $\arg(d_x) = -\varphi(S_x)$ if and only if $\arg(n_b - n_a) = 0^\circ$,
- (2) $\arg(d_x) = 90^\circ - \varphi(S_x)$ if and only if $\arg(n_b - n_a) = 90^\circ$,
- (3) $\arg(d_x) = 180^\circ - \varphi(S_x)$ if and only if $\arg(n_b - n_a) = 180^\circ$,
- (4) $\arg(d_x) = 270^\circ - \varphi(S_x)$ if and only if $\arg(n_b - n_a) = 270^\circ$.

The following Theorem shows how to characterize and calculate the polynomials n_x and d_x associated with a vertex $v_x = n_x/d_x$ from the information of the boundary with a segment S_x in a quadrant, $v_x = n_x/d_x$ belonging to a segment-arc.

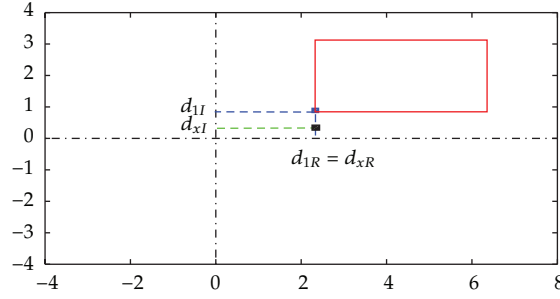


Figure 6: Cases where d_x is not a vertex of the kharitonov rectangle denominator.

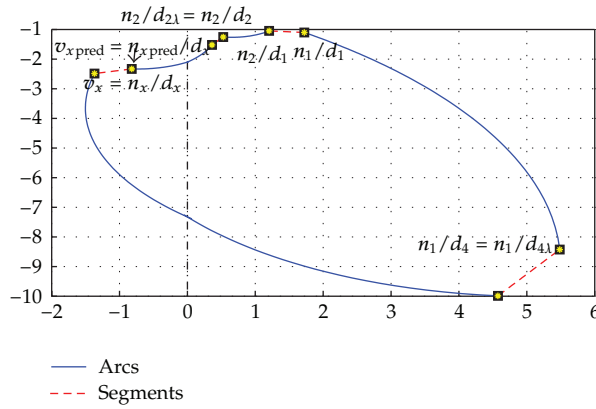


Figure 7: Vertices for the conditions of the Theorem 3.3.

Theorem 3.3 (predecessor). Let S_1 be a complete segment of the value-set boundary with vertices $v_1 = n_1/d_1$ and $v_2 = n_2/d_1$, the successor arc with vertices $v_2 = n_2/d_1$, $v_{2\text{succ}} = n_2/d_{2\lambda}$ counter-clockwise, and the predecessor arc with vertices $v_{1\text{pred}} = n_1/d_{4\lambda}$, $v_1 = n_1/d_1$ counter-clockwise. Let S_x be a segment with vertices $v_{x\text{pred}} = n_{x\text{pred}}/d_x$ and $v_x = n_x/d_x$ counter-clockwise, where v_x belongs to the intersection of S_x and an arc of the boundary (Figure 7). Then

- (1) $\arg(v_x/v_2) = \arg(d_1) + \varphi(S_x)$ (condition C1) and the denominator d_x of v_x defined by n_2/v_x satisfies the denominator condition (Lemma 3.2), if and only if $n_x = n_2$ and cannot be any other assigned polynomial,
- (2) when $n_x \neq n_2$, $\arg(v_x/v_1) = \arg(d_1) + \varphi(S_x) + 90^\circ$ (condition C2) and the denominator d_x of v_x defined by n_1/v_x satisfies the denominator condition (Lemma 3.2) if and only if $n_x = n_1$ and cannot be any other assigned polynomial,
- (3) when $n_x \neq n_1$ and $n_x \neq n_2$, $\tan(\arg(v_x) - \varphi(S_x) + 90^\circ)n_{2R} > n_{2I}$ (condition C3), and the denominator d_x of v_x defined by $n_{2R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 90^\circ)]/v_x$ satisfies the denominator condition (Lemma 3.2) if and only if $n_x = n_3 = n_{2R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 90^\circ)]$ and cannot be any other assigned polynomial,
- (4) when $n_x \neq n_1$, $n_x \neq n_2$, and $n_x \neq n_3$, $\tan(\arg(v_x) - \varphi(S_x) + 180^\circ)n_{1R} > n_{1I}$ (condition C4), and the denominator d_x of v_x defined by $n_{1R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 180^\circ)]/v_x$ satisfies the denominator condition (Lemma 3.2) if and only if $n_x = n_4 = n_{1R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 180^\circ)]$.

Proof. From the complete segment S_1 using the normalization (Lemma 3.1) the values of $d_1, n_1 = v_1 d_1, n_2 = v_2 d_1, d_{2\lambda} = n_2/v_{2\text{succ}},$ and $d_{4\lambda} = n_1/v_{1\text{pred}}$ are known. Obviously the value v_x is known.

(1) \Leftarrow If $n_x = n_2$ the value of $d_x = n_2/v_x$ can be calculated and the denominator condition (Lemma 3.2) is satisfied. On the other hand, the quotient of the vertices $v_x = n_2/d_x$ and $v_2 = n_2/d_1$ is $v_x/v_2 = d_1/d_x,$ and $\arg(v_x/v_2) = \arg(d_1) - \arg(d_x).$ $S_x = S_2/d_x,$ where S_2 is part of the segment with vertices n_1 and $n_2,$ then $\arg(n_2 - n_1) = 0^\circ$ (normalization). Thus $\arg(d_x) = -\varphi(S_x)$ (Property 1) and $\arg(v_x/v_2) = \arg(d_1) + \varphi(S_x);$ Theorem 3.3(C1) is satisfied.

\Rightarrow In order to demonstrate the “only if” part, it must be proven that if Theorem 3.3(C1) and the denominator condition are satisfied then the solution $d_x = n_2/v_x, n_x = n_2$ is unique. It must be noted that Theorem 3.3(C1) can be satisfied when (a) $n_x = n_3,$ (b) $n_x = n_4$ or (c) $n_x = n_1$ and in all the cases, the value of d_x determined, verify the denominator condition.

Let d_x be the denominator of v_x determined by $n_2/v_x,$ verifying Theorem 3.3(C1), and denominator condition, and let $S_x = S_2/d_x$ where S_2 is part of the segment with vertices n_1 and $n_2, \arg(n_2 - n_1) = 0^\circ.$

(a) Let d_x^* be the denominator of v_x determined by $n_3/v_x.$ Then $S_x = S_3/d_x^*$ where S_3 is part of the segment with vertices n_2 and $n_3, \arg(n_3 - n_2) = 90^\circ$ (normalization) and using Property 1 $\arg(d_x^*) = 90^\circ - \varphi(S_x).$ As v_x is the same vertex, then $\arg(n_3/d_x^*) = \arg(n_2/d_x),$ and $\arg(n_3) = \arg(n_2) + 90^\circ. n_x = n_3$ verify Theorem 3.3(C1), because

$$\arg\left(\frac{v_x}{v_2}\right) = \arg(n_2) + 90^\circ - \arg(n_2) + \arg(d_1) - 90^\circ + \varphi(S_x) = \arg(d_1) + \varphi(S_x). \quad (3.1)$$

Let $\alpha = \arg(n_2)$ with $\tan(\alpha) = n_{2I}/n_{2R}.$ Then $\arg(n_3) = \alpha + 90^\circ$ and $\tan(\alpha + 90^\circ) = n_{3I}/n_{3R} = n_{3I}/n_{2R}$ (by normalization $n_{3R} = n_{2R}).$ Thus $n_3 = n_{2R} + j n_{3I} = n_{2R} + j \tan(\alpha + 90^\circ) n_{2R} = n_{2R} - j(n_{2R}^2/n_{2I}).$ Moreover $\arg(d_x^*) = 90^\circ + \arg(d_x),$ and if $d_x = d_{xR} + j d_{xI}$ then $d_x^* = \rho e^{j(\pi/2)} d_x = -\rho d_{xI} + j \rho d_{xR}.$ As $v_x = n_2/d_x$ and $v_x = n_3/d_x^*,$ then $n_2 d_x^* = n_3 d_x$ and they have equal real and imaginary parts.

$\text{Re}[n_2 d_x^*] = \text{Re}[n_3 d_x]$ then

$$\begin{aligned} -\rho d_{xI} n_{2R} - \rho d_{xR} n_{2I} &= n_{2R} d_{xR} + \frac{n_{2R}^2}{n_{2I}} d_{xI}, \\ -\rho d_{xI} n_{2R} n_{2I} - \rho d_{xR} n_{2I}^2 &= n_{2R} n_{2I} d_{xR} + n_{2R}^2 d_{xI}, \\ -(\rho n_{2I} + n_{2R}) d_{xR} n_{2I} &= (n_{2R} + \rho n_{2I}) d_{xI} n_{2R}. \end{aligned} \quad (3.2)$$

Thus $d_{xI}/d_{xR} = -n_{2I}/n_{2R}$

$\text{Im}[n_2 d_x^*] = \text{Im}[n_3 d_x]$ then

$$\begin{aligned} \rho d_{xR} n_{2R} - \rho d_{xI} n_{2I} &= d_{xI} n_{2R} - d_{xR} \frac{n_{2R}^2}{n_{2I}}, \\ \rho d_{xR} n_{2R} n_{2I} - \rho d_{xI} n_{2I}^2 &= d_{xI} n_{2R} n_{2I} - d_{xR} n_{2R}^2, \\ (n_{2I} \rho + n_{2R}) n_{2R} d_{xR} &= d_{xI} n_{2I} (n_{2R} + \rho n_{2I}). \end{aligned} \quad (3.3)$$

Thus $d_{xI}/d_{xR} = n_{2R}/n_{2I}.$

Taking into account both conditions, $n_{2R}/n_{2I} = -n_{2I}/n_{2R} \Leftrightarrow n_{2R}^2 < 0$. This relation is impossible. Therefore, if d_x is a solution then d_x^* is not, and $n_x = n_3$ is not a solution.

(b) Let d_x^* be the denominator of v_x determined by n_4/v_x . Then $S_x = S_4/d_x^*$ where S_4 is part of the segment with vertices n_3 and n_4 , $\arg(n_4 - n_3) = 180^\circ$ (normalization) and using Property 1 $\arg(d_x^*) = 180^\circ - \varphi(S_x)$. As v_x is the same vertex, then $\arg(n_4/d_x^*) = \arg(n_2/d_x)$ and $\arg(n_4) = \arg(n_2) + 180^\circ$. $n_x = n_4$ verify Theorem 3.3(C1), because

$$\arg\left(\frac{v_x}{v_2}\right) = \arg(n_2) + 180^\circ - \arg(n_2) + \arg(d_1) - 180^\circ + \varphi(S_x) = \arg(d_1) + \varphi(S_x). \quad (3.4)$$

In this case the demonstration is trivial noting that $\arg(d_x^*) = 180^\circ + \arg(d_x)$. This is not possible because the Kharitonov polynomial denominator cannot contain the zero.

(c) Let d_x^* be the denominator of v_x determined by n_1/v_x . Then $S_x = S_1/d_x^*$ where S_1 is part of the segment with vertices n_4 and n_1 , $\arg(n_1 - n_4) = 270^\circ$ (normalization) and using Property 1 $\arg(d_x^*) = 270^\circ - \varphi(S_x)$. As v_x is the same vertex, then $\arg(n_1/d_x^*) = \arg(n_2/d_x)$, and $\arg(n_1) = \arg(n_2) + 270^\circ$. $n_x = n_1$ verify Theorem 3.3(C1), because

$$\arg\left(\frac{v_x}{v_2}\right) = \arg(n_2) + 270^\circ - \arg(n_2) + \arg(d_1) - 270^\circ + \varphi(S_x) = \arg(d_1) + \varphi(S_x). \quad (3.5)$$

Let $\alpha = \arg(n_2)$ with $\tan(\alpha) = n_{2I}/n_{2R}$. Then $\arg(n_1) = \alpha + 270^\circ$ and $\tan(\alpha + 270^\circ) = n_{1I}/n_{1R} = n_{2I}/n_{1R}$ (by normalization $n_{3R} = n_{2R}$). Thus $n_1 = n_{1R} + jn_{2I} = (n_{2I}/\tan(\alpha + 270^\circ)) + jn_{2I} = -(n_{2I}^2/n_{2R}) + jn_{2I}$. Moreover $\arg(d_x^*) = 270^\circ + \arg(d_x)$, and if $d_x = d_{xR} + jd_{xI}$ then $d_x^* = \rho e^{j3(\pi/2)} d_x = \rho d_{xI} - j\rho d_{xR}$. As $v_x = n_2/d_x$ and $v_x = n_1/d_x^*$, then $n_2 d_x^* = n_1 d_x$ and they have equals real and imaginary parts.

$\text{Re}[n_2 d_x^*] = \text{Re}[n_1 d_x]$ then

$$\begin{aligned} +\rho d_{xI} n_{2R} + \rho d_{xR} n_{2I} &= -\frac{n_{2I}^2}{n_{2R}} d_{xR} - n_{2I} d_{xI}, \\ (n_{2I} + n_{2R}\rho) d_{xI} n_{2R} &= -(n_{2R}\rho + n_{2I}) d_{xR} n_{2I}. \end{aligned} \quad (3.6)$$

Thus $d_{xI}/d_{xR} = -n_{2I}/n_{2R}$.

$\text{Im}[n_2 d_x^*] = \text{Im}[n_1 d_x]$ then

$$\begin{aligned} -\rho d_{xR} n_{2R} + \rho d_{xI} n_{2I} &= -d_{xI} \frac{n_{2I}^2}{n_{2R}} + d_{xR} n_{2I}, \\ -\rho d_{xR} n_{2R} n_{2R} + \rho d_{xI} n_{2I} n_{2R} &= -d_{xI} n_{2I}^2 + d_{xR} n_{2I} n_{2R}, \\ (n_{2I} + \rho n_{2R}) d_{xI} n_{2I} &= d_{xR} n_{2R} (\rho n_{2R} + n_{2I}). \end{aligned} \quad (3.7)$$

Thus $d_{xI}/d_{xR} = n_{2R}/n_{2I}$.

Taking into account both conditions, $n_{2R}/n_{2I} = -n_{2I}/n_{2R}$. This relation is impossible. Therefore, if d_x is a solution, d_x^* is not and $n_x = n_1$ cannot be a solution.

(2) \Leftarrow If $n_x = n_1$ the value of $d_x = n_1/v_x$ can be calculated and the denominator condition (Lemma 3.2) is satisfied. On the other hand, the quotient of the vertices $v_x = n_1/d_x$ and

$v_2 = n_1/d_1$ is $v_x/v_1 = d_1/d_x$, and $\arg(v_x/v_1) = \arg(d_1) - \arg(d_x)$. $S_x = S_1/d_x$ where S_1 is part of the segment with vertices n_4 and n_1 , then $\arg(n_1 - n_4) = 270^\circ$ (normalization). Thus $\arg(d_x) = 270^\circ - \varphi(S_x)$ (Property 1) and $\arg(v_x/v_1) = \arg(d_1) + \varphi(S_x) + 90^\circ$; Theorem 3.3(C2) is satisfied.

\Rightarrow In order to demonstrate the “only if” part, it must be proven that if Theorem 3.3(C2) and the denominator condition are satisfied then the solution $d_x = n_1/v_x$, $n_x = n_1$ is unique. It must be noted that Theorem 3.3(C2) can be satisfied when (a) $n_x = n_3$ or (b) $n_x = n_4$ and in all the cases, the value of d_x determined, verify the denominator condition.

Let d_x be the denominator of v_x determined by n_1/v_x , verifying Theorem 3.3(C2), and denominator condition, and let $S_x = S_1/d_x$ where S_1 is part of the segment with vertices n_4 and n_1 , $\arg(n_2 - n_1) = 270^\circ$.

(a) Let d_x^* be the denominator of v_x determined by n_3/v_x . Then $S_x = S_3/d_x^*$ where S_3 is part of the segment with vertices n_2 and n_3 , $\arg(n_3 - n_2) = 90^\circ$ (normalization) and using Property 1 $\arg(d_x^*) = 90^\circ - \varphi(S_x)$. As v_x is the same vertex, then $\arg(n_3/d_x^*) = \arg(n_1/d_x)$ and $\arg(n_3) = \arg(n_1) + 180^\circ$. $n_x = n_3$ verify Theorem 3.3(C2), because

$$\arg\left(\frac{v_x}{v_1}\right) = \arg(n_1) + 180^\circ - \arg(n_1) + \arg(d_1) - 90^\circ + \varphi(S_x) = \arg(d_1) + \varphi(S_x) + 90^\circ. \quad (3.8)$$

In this case the demonstration is trivial noting that $\arg(d_x^*) = -180^\circ + \arg(d_x)$. This is not possible because the Kharitonov polynomial denominator cannot contain the zero.

(b) Let d_x^* be the denominator determined by n_4/v_x . Then $S_x = S_4/d_x^*$ where S_4 is part of the segment with vertices n_3 and n_4 , $\arg(n_4 - n_3) = 180^\circ$ (normalization) and using Property 1 $\arg(d_x^*) = 180^\circ - \varphi(S_x)$. As v_x is the same vertex, then $\arg(n_4/d_x^*) = \arg(n_1/d_x)$, and $\arg(n_4) = \arg(n_1) + 270^\circ$. $n_x = n_4$ verify Theorem 3.3(C2), because

$$\arg\left(\frac{v_x}{v_1}\right) = \arg(n_1) + 270^\circ - \arg(n_1) + \arg(d_1) - 180^\circ + \varphi(S_x) = \arg(d_1) + \varphi(S_x) + 90^\circ. \quad (3.9)$$

Let $\alpha = \arg(n_1)$ with $\tan(\alpha) = n_{1I}/n_{1R}$. Then $\arg(n_1) = \alpha + 270^\circ$ and $\tan(\alpha + 270^\circ) = n_{4I}/n_{4R} = -n_{1R}/n_{1I}$ (by normalization $n_{1R} = n_{4R}$). Thus $n_4 = n_{4R} + jn_{4I} = n_{1R} + jn_{2R} \tan(\alpha + 270^\circ) = n_{1R} - j(n_{1R}^2/n_{1I})$. Moreover $\arg(d_x^*) = -90^\circ + \arg(d_x)$, and if $d_x = d_{xR} + jd_{xI}$ then $d_x^* = \rho e^{j3(\pi/2)} d_x = \rho d_{xI} - j\rho d_{xR}$. How $v_x = n_1/d_x$ and $v_x = n_4/d_x^*$, then $n_1 d_x^* = n_4 d_x$ and they have equals real and imaginary parts.

$$\operatorname{Re}[n_1 d_x^*] = \operatorname{Re}[n_4 d_x]$$

$$\begin{aligned} \rho d_{xI} n_{1R} + \rho d_{xR} n_{1I} &= + \frac{n_{1R}^2}{n_{1I}} d_{xI} + n_{1R} d_{xR}, \\ \rho d_{xI} n_{1R} n_{1I} + \rho d_{xR} n_{1I} n_{1I} &= + n_{1R}^2 d_{xI} + n_{1R} d_{xR} n_{1I}, \\ (\rho n_{1I} - n_{1R}) d_{xI} n_{1R} &= (n_{1R} - \rho n_{1I}) d_{xR} n_{1I}. \end{aligned} \quad (3.10)$$

Thus $d_{xI}/d_{xR} = -n_{1I}/n_{1R}$.

$$\text{Im}[n_1 d_x^*] = \text{Im}[n_4 d_x]$$

$$\begin{aligned} -\rho d_{xR} n_{1R} + \rho d_{xI} n_{1I} &= -d_{xR} \frac{n_{1R}^2}{n_{1I}} + d_{xI} n_{1R}, \\ -\rho d_{xR} n_{1R} n_{1I} + \rho d_{xI} n_{1I} n_{1I} &= -d_{xR} n_{1R}^2 + d_{xI} n_{1R} n_{1I}, \\ (-\rho n_{1I} + n_{1R}) d_{xR} n_{1R} &= (n_{1R} - \rho n_{1I}) d_{xI} n_{1I}. \end{aligned} \quad (3.11)$$

and finally $d_{xI}/d_{xR} = n_{1R}/n_{1I}$.

Taking into account both conditions, $-n_{1I}/n_{1R} = n_{1R}/n_{1I}$. This relation is impossible. Therefore, if d_x is a solution, d_x^* is not and $n_x = n_4$ is not a solution.

(3) \Leftarrow If $n_x = n_3$ then $d_x = n_3/v_x$ cannot be directly calculated because n_3 is not known. First, Theorem 3.3(C3) is developed. If $n_x = n_3$ then $S_x = S_3/d_x$ where S_3 is part of the segment with vertices n_2 and n_3 and $\arg(n_3 - n_2) = 90^\circ$. Thus $\arg(d_x) = 90^\circ - \varphi(S_x)$ (Property 1) and $\arg(n_3) = \arg(v_x) + \arg(d_x) = \arg(v_x) + 90^\circ - \varphi(S_x)$.

As $n_{2R} = n_{3R}$, then $n_3 = n_{3R} + j n_{3I} = n_{2R} + j n_{2R} \tan(\arg(v_x) + 90^\circ - \varphi(S_x))$. On the other hand, n_{3I} is greater than n_{2I} because it is counter-clockwise. Therefore $\tan(\arg(v_x) - \varphi(S_x) + 90^\circ) n_{2R} > n_{2I}$ (Theorem 3.3(C3)) is satisfied and d_x can be calculated by the expression $d_x = n_3/v_x = n_{2R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 90^\circ)]/v_x$.

\Rightarrow In order to demonstrate the “only if” part, it must be proven that if Theorem 3.3(C3) and the denominator condition are satisfied then the solution $d_x = n_3/v_x$, $n_x = n_3$ is unique. If $n_x \neq n_2$ and $n_x \neq n_1$, it must be noted that Theorem 3.3(C3) can be satisfied when $n_x = n_4$.

Let d_x be the denominator of v_x determined by n_3/v_x verifying Theorem 3.3(C3) and denominator condition. $S_x = S_3/d_x^*$ where S_3 is part of the segment with vertices n_2 and n_3 , $\arg(n_3 - n_2) = 90^\circ$.

Let d_x^* be the denominator of v_x determined by n_4/v_x . Then $S_x = S_4/d_x^*$ where S_4 is part of the segment with vertices n_3 and n_4 , $\arg(n_4 - n_3) = 180^\circ$ (normalization) and using Property 1 $\arg(d_x^*) = 180^\circ - \varphi(S_x) = \arg(d_x) + 90^\circ$. Thus $d_x^* = \rho e^{j(\pi/2)} d_x = -\rho d_{xI} + j \rho d_{xR}$.

As v_x is the same vertex, $\arg(n_4/d_x^*) = \arg(n_3/d_x)$, and then $\arg(n_4) = \arg(n_3) + 90^\circ$. Let $\alpha = \arg(n_3)$, then $\alpha + 90^\circ = \arg(n_4) = \arg(v_x) + \arg(d_x^*) = \arg(v_x) + 180^\circ - \varphi(S_x)$, and because $\arg(n_3)$ verifies $n_3 = n_{2R} \tan(\alpha) > n_{2I}$ (by normalization), Theorem 3.3(C3) is satisfied.

$n_3 = n_{2R} + j \tan(\alpha) n_{2R} = n_{2R} + j n_{2R} (n_{3R}/n_{3I})$. If $n_x = n_4$ then $n_4 = n_{1R} + j \tan(\alpha + 90^\circ) n_{1R} = n_{1R} - j n_{1R} (n_{3R}/n_{3I})$. As $v_x = n_3/d_x$ and $v_x = n_4/d_x^*$, then $n_2 d_x^* = n_3 d_x$ and they have equal real and imaginary parts.

$$\text{Re}[n_3 d_x^*] = \text{Re}[n_4 d_x]$$

$$\begin{aligned} -n_{2R} \rho d_{xI} - n_{3I} \rho d_{xR} &= n_{1R} d_{xR} + d_{xI} n_{1R} \frac{n_{2R}}{n_{3I}}, \\ -n_{2R} n_{3I} \rho d_{xI} - n_{3I} n_{3I} \rho d_{xR} &= n_{1R} n_{3I} d_{xR} + d_{xI} n_{1R} n_{3R}, \\ -(n_{3I} \rho + n_{1R}) n_{2R} d_{xI} &= (n_{1R} + n_{3I} \rho) n_{3I} d_{xR}, \end{aligned} \quad (3.12)$$

and finally $d_{xI}/d_{xR} = -n_{3I}/n_{3R}$.

$$\text{Im}[n_3 d_x^*] = \text{Im}[n_4 d_x]$$

$$\begin{aligned} -n_{3I}\rho d_{xI} + n_{2R}\rho d_{xR} &= d_{xI}n_{1R} - d_{xR}n_{1R}\frac{n_{2R}}{n_{3I}}, \\ -n_{3I}n_{3I}\rho d_{xI} + n_{3I}n_{2R}\rho d_{xR} &= d_{xI}n_{1R}n_{3I} - d_{xR}n_{1R}n_{2R}, \\ -(n_{3I}\rho + n_{1R})d_{xI}n_{3I} &= -(n_{3I}\rho + n_{1R})d_{xR}n_{2R} \end{aligned} \quad (3.13)$$

and finally $d_{xI}/d_{xR} = n_{3R}/n_{3I}$.

Taking into account both conditions, $-n_{3I}/n_{3R} = n_{3R}/n_{3I}$. This relation is impossible. Therefore, if d_x is a solution, d_x^* is not, and $n_x = n_3$ is not a solution.

(4) \Leftarrow If $n_x = n_4$ then $d_x = n_4/v_x$ cannot be directly calculated because n_4 is not known. First, Theorem 3.3(C4) is developed.

If $n_x = n_4$ then $S_x = S_4/d_x$ where S_4 is part of the segment with vertices n_3 and n_4 verifying that $\arg(n_4 - n_3) = 180^\circ$. Thus $\arg(d_x) = 180^\circ - \varphi(S_x)$ (Property 1) and $\arg(n_4) = \arg(v_x) + \arg(d_x) = \arg(v_x) + 180^\circ - \varphi(S_x)$. Moreover, $n_{1R} = n_{4R}$. Then $n_4 = n_{4R} + jn_{4I} = n_{1R} + jn_{1R} \tan(\arg(v_x) + 180^\circ - \varphi(S_x))$. On the other hand, n_{4I} is greater than n_{1I} because it is counter-clockwise.

Therefore the condition $\tan(\arg(v_x) - \varphi(S_x) + 180^\circ)n_{1R} > n_{1I}$ Theorem 3.3(C4) is satisfied and d_x can be calculated using the expression $d_x = n_4/v_x = n_{1R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 180^\circ)]/v_x$.

\Rightarrow If $n_x \neq n_2$, $n_x \neq n_1$ and $n_x \neq n_3$ it is $n_x = n_4$. \square

Remark 3.4. This theorem is used in the example of Section 5, for the value set III (frequency $\omega = 1.2$) in order to assign the second and fifth vertices.

The following Theorem is analogous to Theorem 3.3 when S_x is a segment with vertices $v_x = n_x/d_x$ and $v_{x\text{succ}} = n_{x\text{succ}}/d_x$ counter-clockwise, and belonging to an arc-segment.

Theorem 3.5 (successor). *Let S_1 be a complete segment of the value-set boundary with vertices $v_1 = n_1/d_1$ and $v_2 = n_2/d_1$, the successor arc to S_1 , with vertices $v_2 = n_2/d_1$, $v_{2\text{succ}} = n_2/d_{2\lambda}$ counter-clockwise, and the predecessor arc to S_1 with vertices $v_{1\text{pred}} = n_1/d_{4\lambda}$, $v_1 = n_1/d_1$ counter-clockwise. Let S_x be a boundary segment with vertices $v_x = n_x/d_x$ and $v_{x\text{succ}} = n_{x\text{succ}}/d_x$ counter-clockwise, where v_x belongs to the intersection of an arc of the boundary and S_x . Then*

- (1) $\arg(v_x/v_2) = \arg(d_1) + \varphi(S_x) - 90^\circ$ (condition C1) and the denominator d_x of v_x defined by n_2/v_x satisfies the denominator condition (Lemma 3.2), if and only if $n_x = n_2$ and cannot be any other assigned polynomial,
- (2) when $n_x \neq n_2$, $\arg(v_x/v_1) = \arg(d_1) + \varphi(S_x)$ (condition C2) and the denominator d_x of v_x defined by n_1/v_x satisfies the denominator condition (Lemma 3.2) if and only if $n_x = n_1$ and cannot be any other assigned polynomial,
- (3) when $n_x \neq n_1$ and $n_x \neq n_2$, $\tan(\arg(v_x) - \varphi(S_x) + 180^\circ)n_{2R} > n_{2I}$ (condition C3), and the denominator d_x of v_x defined by $n_{2R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 180^\circ)]/v_x$ satisfies the denominator condition (Lemma 3.2) if and only if $n_x = n_3 = n_{2R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 180^\circ)]$ and cannot be any other assigned polynomial,

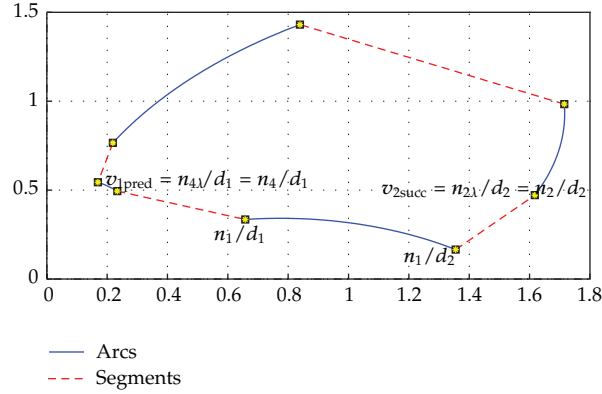


Figure 8: Arc and two complete segments.

- (4) when $n_x \neq n_1$, $n_x \neq n_2$, and $n_x \neq n_3$, $\tan(\arg(v_x) - \varphi(S_x) + 270^\circ)n_{1R} > n_{1I}$ (condition C4), and the denominator d_x of v_x defined by $n_{1R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 270^\circ)]/v_x$ satisfies the denominator condition (Lemma 3.2) if and only if $n_x = n_4 = n_{1R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 270^\circ)]$.

Proof. Analogous to Theorem 3.3. □

Remark 3.6. This theorem is used in the example of Section 5, for the value set III (frequency $\omega = 1.2$) in order to assign the third, fifth, and sixth vertices.

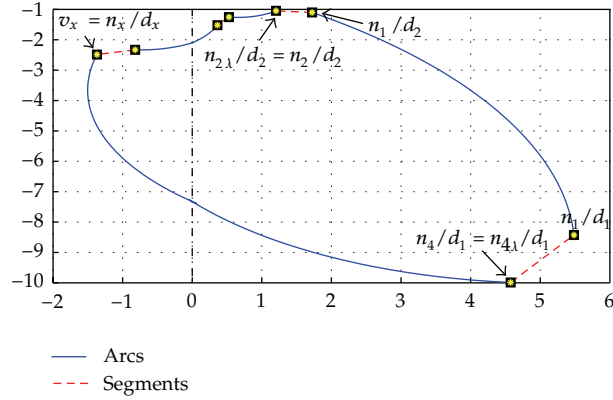
4. Assigned Polynomial Determination When There Is a Complete Arc in a Quadrant

In order to determine the polynomials numerator and denominator associated to a vertex of the value set boundary with the minimum number of elements, the situation of an arc in a quadrant will be considered. So, let A_1 be an arc of the value-set boundary with vertices $v_1 = n_1/d_1$ and $v_2 = n_1/d_2$. A continuity arc-segment in a quadrant (see [10, Theorem 2]) implies that there will be a successor segment with vertices $v_2 = n_1/d_2$, $v_{2succ} = n_{2\lambda}/d_2$ counter-clockwise and a predecessor segment with vertices $v_1 = n_1/d_1$ and $v_{1pred} = n_{4\lambda}/d_1$ counter-clockwise.

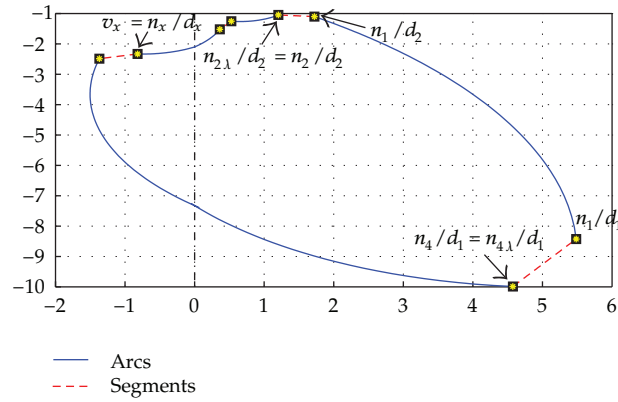
When these segments are completed the denominators are vertices of the Kharitonov rectangle. Figure 8 shows this situation.

As was shown, the values of d_1 , d_2 , and n_1 can be calculated from the complete arc based on a normalization (see [10, Theorem 5]). The following normalization simplifies the nomenclature.

Lemma 4.1 (arc normalization). *Let A_1 be a complete arc of the value-set boundary with vertices $v_1 = n_1/d_1$ and $v_2 = n_1/d_2$, the normalization $n_1 = \cos(\varphi(n_1)) + j \sin(\varphi(n_1))$, where $\varphi(n_1) = 360^\circ - \arg(1/v_2 - 1/v_1)$, $\arg(1/v_2 - 1/v_1)$ being the argument of the segment $1/v_2 - 1/v_1$. Then $d_1 = n_1/v_1$, $d_2 = n_1/v_2$, $n_{4\lambda} = d_1 v_{1pred}$, and $n_{2\lambda} = d_2 v_{2succ}$, where v_{2succ} (v_{1pred}) is any point of the next (previous) segment of the arc A_1 .*



(a)



(b)

Figure 9: (a) v_x vertex of two elements, segment-arc. (b) v_x vertex of two elements, arc-segment.

Proof. It is trivial. This normalization is one of the infinite possible solutions for a value set. This normalization implies fitting n_1 with modulus $|n_1| = 1$ and angle so that the segment of the Kharitonov polynomial denominator with vertices d_1 and d_2 will be parallel to the real axis counter-clockwise. Thus, from the information with a complete arc in a quadrant the values of d_1 , d_2 , n_1 , $n_{2\lambda}$, and $n_{4\lambda}$ can be calculated. \square

This paper deals with the general case where $d_{2R} \neq 0$, $d_{2I} \neq 0$, $d_{1R} \neq 0$, and $d_{1I} \neq 0$.

Given a vertex $v_x = n_x/d_x$ in a quadrant, the target is to determine the polynomials n_x and d_x . The vertex v_x belongs to a part of an arc and a part of a segment, due to the continuity arc-segment in a quadrant. So, v_x will be the vertex of two elements, segment-arc (Figure 9(a)) or arc-segment (Figure 9(b)).

The following Lemma shows the necessary conditions on the denominator d_x to be a solution of $v_x = n_x/d_x$.

Lemma 4.2 (numerator condition). *Let A_1 be a complete arc in a quadrant and let n_x be the numerator of a vertex $v_x = n_x/d_x$ in a quadrant. Then it is a necessary condition that n_x satisfies one of the following conditions:*

- (1) $(n_{1R} < n_{2\lambda R} \text{ and } n_{1I} < n_{4\tau I})$ and $\{(n_{xR} = n_{1R} \text{ and } n_{xI} = n_{1I}) [n_x = n_1] \text{ or } (n_{xR} = n_{1R} \text{ and } n_{xI} \geq n_{1I}) [n_x = n_4] \text{ or } (n_{xI} = n_{1I} \text{ and } n_{xR} \geq n_{1R}) [n_x = n_2] \text{ or } (n_{xR} > n_{1R} \text{ and } n_{xI} > n_{1I}) [n_x = n_3]\}$,
- (2) $(n_{1R} > n_{4\lambda R} \text{ and } n_{1I} < n_{2\tau I})$ and $\{(n_{xR} = n_{1R} \text{ and } n_{xI} = n_{1I}) [n_x = n_1] \text{ or } (n_{xR} = n_{1R} \text{ and } n_{xI} \geq n_{1I}) [n_x = n_2] \text{ or } (n_{xI} = n_{1I} \text{ and } n_{xR} \leq n_{1R}) [n_x = n_4] \text{ or } (n_{xR} < n_{1R} \text{ and } n_{xI} > n_{1I}) [n_x = n_3]\}$,
- (3) $(n_{1R} > n_{2\lambda R} \text{ and } n_{1I} > n_{4\tau I})$ and $\{(n_{xR} = n_{1R} \text{ and } n_{xI} = n_{1I}) [n_x = n_1] \text{ or } (n_{xR} = n_{1R} \text{ and } n_{xI} \leq n_{1I}) [n_x = n_4] \text{ or } (n_{xI} = n_{1I} \text{ and } n_{xR} \leq n_{1R}) [n_x = n_2] \text{ or } (n_{xR} < n_{1R} \text{ and } n_{xI} < n_{1I}) [n_x = n_3]\}$,
- (4) $(n_{1R} < n_{4\lambda R} \text{ and } n_{1I} > n_{2\tau I})$ and $\{(n_{xR} = n_{1R} \text{ and } n_{xI} = n_{1I}) [n_x = n_1] \text{ or } (n_{xR} = n_{1R} \text{ and } n_{xI} \leq n_{1I}) [n_x = n_2] \text{ or } (n_{xI} = n_{1I} \text{ and } n_{xR} \geq n_{1R}) [n_x = n_4] \text{ or } (n_{xR} > n_{1R} \text{ and } n_{xI} < n_{1I}) [n_x = n_3]\}$,

where n_{iR} is the real part of n_i and n_{iI} is the imaginary part of n_i , and the corresponding assigned numerator is shown between brackets.

Proof. The proof is obtained directly from the information of a complete arc in a quadrant and the properties of the Kharitonov rectangle. So, from the complete arc and the normalization (Lemma 3.2), the values of n_1 , $n_{2\lambda}$, and $n_{4\lambda}$ are known. Then, n_1 can be established as k_{n1} , k_{n2} , k_{n3} , or k_{n4} .

- (1) If $(n_{1R} < n_{2\lambda R} \text{ and } n_{1I} < n_{4\tau I})$ then n_1 is k_{n1} . Given a value n_x , it will be a vertex of the Kharitonov rectangle numerator only if $n_{xR} = n_{1R}$ and $n_{xI} = n_{1I}$ (n_x is $n_1 = k_{n1}$) or $n_{xR} = n_{1R}$ and $n_{xI} > n_{1I}$ (n_x is $n_4 = k_{n4}$) or $n_{xI} = n_{1I}$ and $n_{xR} > n_{1R}$ (n_x is $n_2 = k_{n2}$) or $n_{xR} > n_{1R}$ and $n_{xI} > n_{1I}$ (n_x is $n_3 = k_{n3}$). Note that if any of these conditions is not satisfied, then n_x cannot be a solution. For example, if $n_{xR} = n_{1R}$ and $n_{xI} < n_{1I}$, n_x does not belong to the rectangle with vertex n_1 , $n_{2\lambda}$, and $n_{4\lambda}$ are elements of the successor and predecessor edge.
- (2) Similarly, if $(n_{1R} > n_{4\lambda R} \text{ and } n_{1I} < n_{2\tau I})$ then n_1 is k_{n2} . Given a value n_x , it will be a vertex of the Kharitonov rectangle numerator only if $n_{xR} = n_{1R}$ and $n_{xI} = n_{1I}$ (n_x is $n_1 = k_{n2}$) or $n_{xR} = n_{1R}$ and $n_{xI} > n_{1I}$ (n_x is $n_2 = k_{n3}$) or $n_{xI} = n_{1I}$ and $n_{xR} < n_{1R}$ (n_x is $n_4 = k_{n1}$) or $n_{xR} < n_{1R}$ and $n_{xI} > n_{1I}$ (n_x is $n_3 = k_{n4}$).
- (3) If $n_{1R} > n_{2\lambda R}$ and $n_{1I} > n_{4\tau I}$ then n_1 is k_{n3} . Given a value n_x , it will be a vertex of the Kharitonov rectangle numerator only if $n_{xR} = n_{1R}$ and $n_{xI} = n_{1I}$ (n_x is $n_1 = k_{n3}$) or $n_{xR} = n_{1R}$ and $n_{xI} < n_{1I}$ (n_x is $n_4 = k_{n2}$) or $n_{xI} = n_{1I}$ and $n_{xR} < n_{1R}$ (n_x is $n_2 = k_{n4}$) or $n_{xR} < n_{1R}$ and $n_{xI} < n_{1I}$ (n_x is $n_3 = k_{n1}$).
- (4) Finally, if $n_{1R} < n_{4\lambda R}$ and $n_{1I} > n_{2\tau I}$ then n_1 is k_{n4} . Given a value n_x , it will be a vertex of the Kharitonov rectangle numerator only if $n_{xR} = n_{1R}$ and $n_{xI} = n_{1I}$ (n_x is $n_1 = k_{n4}$) or $n_{xR} = n_{1R}$ and $n_{xI} < n_{1I}$ (n_x is $n_2 = k_{n1}$) or $n_{xI} = n_{1I}$ and $n_{xR} > n_{1R}$ (n_x is $n_4 = k_{n3}$) or $n_{xR} > n_{1R}$ and $n_{xI} < n_{1I}$ (n_x is $n_3 = k_{n2}$). \square

On the other hand, the behaviour of an arc on the complex plane when it is divided by a complex number is well known. The following property shows this behaviour.

Property 2. Let $A_x = n_x/S$ be an arc on the complex plane with vertices v_{x1} and v_{x2} counter-clockwise where S is a segment with vertices d_a and d_b counter-clockwise. Let n_x be a complex number with argument $\arg(n_x)$. Let $\varphi(A_x)$ be $\varphi(A_x) \equiv \arg(1/v_{x2} - 1/v_{x1})$. Then the relation between the argument of n_x and $\varphi(A_x)$, is given by

- (1) $\arg(n_x) = -\varphi(A_x)$ if and only if $\arg(d_b - d_a) = 0^\circ$,
- (2) $\arg(n_x) = 90^\circ - \varphi(A_x)$ if and only if $\arg(d_b - d_a) = 90^\circ$,
- (3) $\arg(n_x) = 180^\circ - \varphi(A_x)$ if and only if $\arg(d_b - d_a) = 180^\circ$,
- (4) $\arg(n_x) = 270^\circ - \varphi(A_x)$ if and only if $\arg(d_b - d_a) = 270^\circ$.

The following Theorem shows how to characterize and calculate the polynomials n_x and d_x associated with a vertex $v_x = n_x/d_x$ from the information of the boundary with an arc A_x in a quadrant, belonging to an arc-segment.

Theorem 4.3 (predecessor). *Let A_1 be an arc of the value-set boundary with vertices $v_1 = n_1/d_1$ and $v_2 = n_1/d_2$, the successor segment with vertices $v_2 = n_1/d_2$, $v_{2 \text{ succ}} = n_{2\lambda}/d_2$ counter-clockwise, and the predecessor segment with vertices $v_{1 \text{ pred}} = n_{4\lambda}/d_1$, $v_1 = n_1/d_1$ counter-clockwise. Let A_x be an arc with vertices $v_{x \text{ pred}} = n_x/d_{x \text{ pred}}$ and $v_x = n_x/d_x$ counter-clockwise. Then*

- (1) $\arg(v_2/v_x) = \arg(n_1) + \varphi(A_x)$ (condition C1) and n_x satisfies the numerator condition, where $n_x = d_2 v_x$, if and only if $d_x = d_2$ and cannot be any other assigned polynomial,
- (2) when $d_x \neq d_2$, $\arg(v_1/v_x) = \arg(n_1) + \varphi(A_x) + 90^\circ$ (condition C2) and $n_x = d_1 v_x$ satisfies the numerator condition if and only if $d_x = d_1$ and cannot be any other assigned polynomial,
- (3) when $d_x \neq d_1$ and $d_x \neq d_2$, $\tan(\arg(1/v_x) - \varphi(A_x) + 90^\circ) d_{2R} > d_{2I}$ (condition C3), and $n_x = d_{2R}[1 + j \tan(\arg(1/v_x) - \varphi(A_x) + 90^\circ)] v_x$ satisfies the numerator condition if and only if $d_x = d_3 = d_{2R}(1 + j \tan(\arg(1/v_x) - \varphi(A_x) + 90^\circ))$ and cannot be any other assigned polynomial,
- (4) when $d_x \neq d_1$, $d_x \neq d_2$, and $d_x \neq d_3$, $\tan(\arg(1/v_x) - \varphi(A_x) + 180^\circ) d_{1R} > d_{1I}$ (condition C4), and $n_x = d_{1R}[1 + j \tan(\arg(1/v_x) - \varphi(A_x) + 180^\circ)] v_x$ satisfies the numerator condition if and only if $d_x = d_4 = d_{1R}(1 + j \tan(\arg(1/v_x) - \varphi(A_x) + 180^\circ))$.

Proof. Analogous to Theorem 3.3. □

Remark 4.4. This theorem is used in the example of Section 5, for the value set I (frequency $\omega = 1.0$) in order to assign the fifth and seventh vertices, and for the value set II (frequency $\omega = 1.1$) to assign the third, fifth, and seventh vertices.

The following theorem is analogous to Theorem 4.3 when A_x is an arc with vertices $v_x = n_x/d_x$ and $v_{x \text{ succ}} = n_x/d_{x \text{ succ}}$ counter-clockwise, and belonging to a segment-arc.

Theorem 4.5 (successor). *Let A_1 be a complete arc of the value-set boundary with vertices $v_1 = n_1/d_1$ and $v_2 = n_1/d_2$, the successor segment with vertices $v_2 = n_1/d_2$, $v_{2 \text{ succ}} = n_{2\lambda}/d_2$ counter-clockwise and the predecessor segment with vertices $v_{1 \text{ pred}} = n_{4\lambda}/d_1$, $v_1 = n_1/d_1$ counter-clockwise. Let A_x be an arc with vertices $v_{x \text{ succ}} = n_x/d_{x \text{ succ}}$ and $v_x = n_x/d_x$ counter-clockwise*

Then

- (1) $\arg(v_2/v_x) = \varphi(A_x) + \arg(n_1) - 90^\circ$ (condition C1) and n_x satisfies the numerator condition, where $n_x = d_2 v_x$, if and only if $d_x = d_2$ and cannot be any other assigned polynomial,

- (2) when $d_x \neq d_2$, $\arg(v_1/v_x) = \varphi(A_x) + \arg(n_1)$ (condition C2) and $n_x = d_1 v_x$ satisfies the numerator condition if and only if $d_x = d_1$ and cannot be any other assigned polynomial,
- (3) when $d_x \neq d_1$ and $d_x \neq d_2$, $\tan(\arg(1/v_x) - \varphi(A_x) + 180^\circ) d_{2R} > d_{2I}$ (condition C3), and $n_x = d_{2R}[1 + j \tan(\arg(1/v_x) - \varphi(A_x) + 180^\circ)] v_x$ satisfies the numerator condition if and only if $d_x = d_3 = d_{2R}(1 + j \tan(\arg(1/v_x) - \varphi(A_x) + 180^\circ))$ and cannot be any other assigned polynomial,
- (4) when $d_x \neq d_1$, $d_x \neq d_2$, and $d_x \neq d_3$, $\tan(\arg(1/v_x) - \varphi(A_x) + 270^\circ) d_{1R} > d_{1I}$ (condition C4), and $n_x = d_{1R}[1 + j \tan(\arg(1/v_x) - \varphi(A_x) + 270^\circ)] v_x$ satisfies the numerator condition if and only if $d_x = d_4 = d_{1R}(1 + j \tan(\arg(1/v_x) - \varphi(A_x) + 270^\circ))$.

Proof. Analogous to Theorem 3.3. □

Remark 4.6. This theorem is used in the example of Section 5, for the value set I (frequency $\omega = 1.0$) in order to assign the third, fourth, and sixth vertices, and for the value set II (frequency $\omega = 1.1$) to assign the fourth and sixth vertices.

Finally, the following theorem points out the necessary and sufficient condition.

Theorem 4.7. *Given a value set, all the assigned polynomials of the vertices can be determined if and only if there is a complete edge or a complete arc lying on a quadrant when the normalized edge satisfies $n_{2R} \neq 0$, $n_{2I} \neq 0$, $n_{1R} \neq 0$, and $n_{1I} \neq 0$ or the normalized arc satisfies $d_{2R} \neq 0$, $d_{2I} \neq 0$, $d_{1R} \neq 0$, and $d_{1I} \neq 0$.*

Proof. It is obvious from Theorems 3.3–4.5. □

5. Algorithm and Examples

Algorithm 5.1. Given a value set with a complete segment or a complete arc in a quadrant, to obtain the Kharitonov polynomials the following.

- (1) If there is a complete segment in a quadrant, S_1 , with vertices $v_1 = n_1/d_1$ and $v_2 = n_2/d_1$, the successor arc with vertices $v_2 = n_2/d_1$, $v_{2\text{succ}} = n_2/d_{2\lambda}$ counter-clockwise and the predecessor arc with vertices $v_{1\text{pred}} = n_1/d_{4\lambda}$, $v_1 = n_1/d_1$ counter-clockwise then for all vertex $v_x = n_x/d_x$:
 - (a) if $v_x = n_x/d_x$ is a vertex intersection of a segment and an arc counter-clockwise, then the assigned polynomials numerator and denominator, n_x and d_x , determine applying Theorem 3.3,
 - (b) if $v_x = n_x/d_x$ is a vertex intersection of an arc and a segment counter-clockwise, then the assigned polynomials numerator and denominator, n_x and d_x , determine applying Theorem 3.5.
- (2) If there is a complete arc in a quadrant, A_1 , with vertices $v_1 = n_1/d_1$ and $v_2 = n_1/d_2$, the successor segment with vertices $v_2 = n_1/d_2$, $v_{2\text{succ}} = n_{2\lambda}/d_2$ counter-clockwise and the predecessor segment with vertices $v_{1\text{pred}} = n_{4\lambda}/d_1$, $v_1 = n_1/d_1$ counter-clockwise, then given a vertex $v_x = n_x/d_x$:
 - (a) if $v_x = n_x/d_x$ is a vertex intersection of an arc and a segment counter-clockwise, then the assigned polynomials numerator and denominator, n_x and d_x , determine applying Theorem 4.3,

Table 1: Value set boundary information.

$\omega = 1.0$			$\omega = 1.1$			$\omega = 1.2$		
(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
v_1	$1.5676 + 2.5946j$	0	v_1	$-2.8422 + 2.9830j$	0	v_1	$6.1015 + 5.2779j$	1
v_2	$2.0000 + 8.0000j$	1	v_2	$-0.9808 + 2.4599j$	1	v_2	$6.5135 + 6.8573j$	0
v_3	$0.8000 + 10.4000j$	0		$0 + 3.0420j$	0		$0 + 8.5560j$	0
	$0 + 10.0000j$	1	v_3	$0.4996 + 3.0386j$	1	v_3	$-3.0339 + 6.1294j$	1
v_4	$-4.8000 + 7.6000j$	0	v_4	$2.3317 + 3.0261j$	0	v_4	$-2.2110 + 5.1007j$	0
v_5	$-3.5862 + 1.0345j$	1	v_5	$5.1859 + 6.6181j$	1	v_5	$-0.4710 + 3.4462j$	1
v_6	$2.5517 + 0.6207j$	0	v_6	$5.2164 + 8.6623j$	0		$0 + 3.6463j$	0
v_7	$-1.3443 + 1.2131j$	1		$0 + 8.7404j$	0	v_6	$1.4690 + 3.4428j$	1
	$0 + 2.3336j$	1	v_7	$-3.8291 + 3.7385j$	1	v_7	$2.9559 + 3.2369j$	0

(a): Vertex (v_i) or cut point (blank) with an axis. (b): Value of the vertex or cut point.

(c): Edge (2.1) or arc (0) between this element and the next element. If the element is the last, the next element is the first.

(b) if $v_x = n_x/d_x$ is a vertex intersection of a segment and an arc counter-clockwise, then the assigned polynomials numerator and denominator, n_x and d_x , determine applying Theorem 4.5.

(3) Calculate the values of the assigned polynomials n_j, d_k , solving the equation system (2.7):

$$v_i = \frac{n_j}{d_k}. \quad (5.1)$$

(4) Calculate the numerator and denominator rectangles with Kharitonov polynomial values $N = (k_{n1}(j\omega), k_{n2}(j\omega), k_{n3}(j\omega), k_{n4}(j\omega))$, $D = (k_{d1}(j\omega), k_{d2}(j\omega), k_{d3}(j\omega), k_{d4}(j\omega))$ applying (2.8).

Example 5.2. Figure 10 shows three value sets of an interval plant. The necessary information (Table 1) is

- (i) the vertices,
- (ii) the intersections with the axis,
- (iii) the shape of the boundary's elements: arc or segment.

This example illustrates how to obtain the assigned polynomials and the numerator and denominator rectangles for each value set, and remarks the theorem used in each step.

5.1. Value Set at Frequency $\omega = 1.0$

The complete arc with vertices $v_1 = n_1/d_1 = 1.5676 + 2.5946j$ and $v_2 = n_1/d_2 = 2.0000 + 8.0000j$ is taken as initial element. Then Theorems 4.3 and 4.5 will be applied. So

$$v_{2\text{succ}} = \frac{n_{2\lambda}}{d_2} = 0.8000 + 10.4000j, \quad v_{1\text{pred}} = \frac{n_{4\lambda}}{d_1} = 2.3336j. \quad (5.2)$$

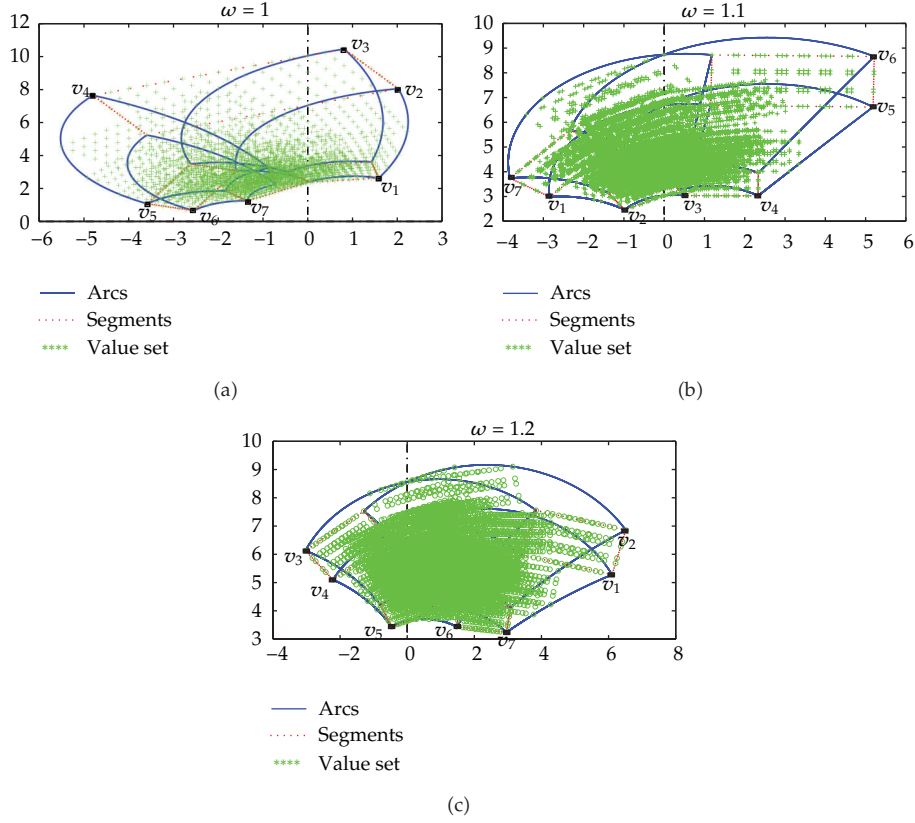


Figure 10: Three value sets of an interval plant.

Applying the arc normalization (Lemma 4.1) the following data are obtained

$$\begin{aligned} \varphi(n_1) &= 229.40, \quad n_1 = -0.6508 - 0.7592j, \quad d_1 = -0.3254 + 0.0542j, \quad d_2 = -0.1085 + 0.0542j, \\ n_{4\lambda} &= -0.1266 - 0.7594j, \quad n_{2\lambda} = -0.6508 - 1.0846j. \end{aligned} \quad (5.3)$$

Then, all the other vertices are assigned as follows.

(1) Vertex $v_3 = v_x = n_x/d_x = 0.8000 + 10.4000j$. Then $v_{x\text{pred}} = n_x/d_{x\text{pred}} = 2.0000 + 8.0000j$. These are the vertices of an edge, and Theorem 4.5 is applied, $v_{x\text{succ}} = 10.0000j$, $\varphi(A_x) = 210.97$.

Case 1. Theorem 4.5(C1) is satisfied: $\arg(v_2/v_x) = \varphi(A_x) + \arg(n_1) - 90 = 350.36$ and $n_x = d_2 v_x = -0.6508 - 1.0846j$ satisfies the Numerator Condition (Lemma 4.2(4), $n_x = n_2$):

$$\begin{aligned} (n_{1R} = -0.6508 < n_{4\lambda R} = -0.1266, \quad n_{1I} = -0.7592 > n_{2\lambda I} = -1.0846), \\ (n_{xR} = n_{1R} = -0.6508, \quad n_{xI} = -1.0846 \leq n_{1I} = -0.7592). \end{aligned} \quad (5.4)$$

Then $d_x = d_2 = -0.1085 + 0.0542j$. Therefore $v_3 = v_x = n_2/d_2$.

(2) Vertex $v_4 = v_x = n_x/d_x = -4.8000 + 7.6000j$. Then $v_{xpred} = 10j$. These are the vertices of an edge, and Theorem 4.5 is applied: $v_{xsucc} = -3.5862 + 1.0345j$, $\varphi(A_x) = 174.29$.

Case 1. Theorem 4.5(C1) is satisfied: $\arg(v_2/v_x) = \varphi(A_x) + \arg(n_1) - 90 = 313.69$ and $n_x = d_2 v_x = 0.1084 - 1.0847j$ satisfies the Numerator Condition (Lemma 4.2(4), $n_x = n_3$). Then $d_x = d_2 = -0.1085 + 0.0542j$. Therefore $v_4 = v_x = n_3/d_2$.

(3) Vertex $v_5 = v_x = n_x/d_x = -3.5862 + 1.0345j$. Then $v_{xpred} = n_x/d_{xpred} = -4.8000 + 7.6000j$. These are the vertices of an arc, and Theorem 4.3 is applied: $\varphi(A_x) = 174.29$.

Case 1. Theorem 4.3(C1) is not satisfied: $\arg(v_2/v_x) = 272.06 \neq \arg(n_1) + \varphi(A_x) = 43.69$.

Case 2. Theorem 4.3(C2) is not satisfied: $\arg(v_1/v_x) = 254.95 \neq \arg(n_1) + \varphi(A_x) + 90 = 133.69$.

Case 3. Theorem 4.3(C3) is satisfied: $\tan(\arg(1/v_x) - \varphi(A_x) + 90)d_{2R} = 0.2712 > d_{2I} = 0.0542$ and $n_x = 0.1085 - 1.0846j$ satisfies the Numerator Condition (Lemma 4.2(4)) $n_x = n_3$. Then $d_x = d_3 = -0.1085 + 0.2712j$, $v_5 = v_x = n_3/d_3$.

(4) Vertex $v_6 = v_x = n_x/d_x = -2.5517 + 0.6207j$. Then $v_{xpred} = n_x/d_{xpred} = -3.5862 + 1.0345j$. These are the vertices of an edge, and Theorem 4.5 is applied: $v_{xsucc} = -1.3443 + 1.2131j$, $\varphi(A_x) = 261.87$.

Case 1. Theorem 4.5(C1) is not satisfied: $\arg(v_2/v_x) = 269.64 \neq \varphi(A_x) + \arg(n_1) - 90 = 41.27$.

Case 2. Theorem 4.5(C2) is not satisfied: $\arg(v_1/v_x) = 252.53 \neq \varphi(A_x) + \arg(n_1) = 131.27$.

Case 3. Theorem 4.5(C3) is satisfied: $\tan(\arg(1/v_x) - \varphi(A_x) + 180)d_{2R} = 0.2712 > d_{2I} = 0.0542$ and $n_x = 0.1085 - 0.7592j$ satisfies the Numerator Condition (Lemma 4.2(3)) $n_x = n_4$; then $d_x = d_3 = -0.1085 + 0.2712j$ and $v_6 = v_x = n_4/d_3$.

(5) Vertex $v_7 = v_x = n_x/d_x = -1.3443 + 1.2131j$. Then $v_{xpred} = n_x/d_{xpred} = -2.5517 + 0.6207j$. These are the vertices of an arc, and Theorem 4.3 is applied: $\varphi(A_x) = 261.87$.

Case 1. Theorem 4.3(C1) is not satisfied: $\arg(v_2/v_x) = 298.03 \neq \arg(n_1) + \varphi(A_x) = 131.27$.

Case 2. Theorem 4.3(C2) is not satisfied: $\arg(v_1/v_x) = 280.93 \neq \arg(n_1) + \varphi(A_x) + 90 = 221.27$.

Case 3. Theorem 4.3(C3) is not satisfied: $\tan(\arg(1/v_x) - \varphi(A_x) + 90)d_{2R} = -0.1302 < d_{2I} = 0.0542$.

Case 4. Theorem 4.3(C4) is satisfied: $\tan(\arg(1/v_x) - \varphi(A_x) + 180)d_{1R} = 0.2712 > d_{1I} = 0.0542$ and $n_x = 0.1085 - 0.7592j$ satisfies the Numerator Condition (Lemma 4.2(4)) $n_x = n_4$. Then $d_x = d_4 = -0.3254 + 0.2712j$; $v_7 = v_x = n_4/d_4$.

In summary, the assigned polynomials are

$$v_1 = \frac{n_1}{d_1}, \quad v_2 = \frac{n_1}{d_2}, \quad v_3 = \frac{n_2}{d_2}, \quad v_4 = \frac{n_3}{d_2}, \quad v_5 = \frac{n_3}{d_3}, \quad v_6 = \frac{n_4}{d_3}, \quad v_7 = \frac{n_4}{d_4}, \quad (5.5)$$

and the values can be calculated: from normalization,

$$n_1 = -0.6508 - 0.7592j, \quad d_1 = -0.3254 + 0.0542j, \quad d_2 = -0.1085 + 0.0542j, \quad (5.6)$$

and from the vertices,

$$\begin{aligned}
 v_3 : n_2 &= -0.6508 - 1.0846j, & d_2 &= -0.1085 + 0.0542j, \\
 v_4 : n_3 &= 0.1084 - 1.0847j, & d_2 &= -0.1085 + 0.0542j, \\
 v_5 : n_3 &= 0.1085 - 1.0846j, & d_3 &= -0.8464 + 2.0152j, \\
 v_6 : n_4 &= 0.1085 - 0.7592j, & d_3 &= -0.1085 + 0.2712j, \\
 v_7 : n_4 &= 0.1085 - 0.7593j, & d_4 &= -0.3254 + 0.2712j.
 \end{aligned} \tag{5.7}$$

Then

$$\begin{aligned}
 k_{n1}(j\omega) &= -0.6508 - 1.0847j, & k_{n2}(j\omega) &= 0.1085 - 1.0847j, \\
 k_{n3}(j\omega) &= 0.1085 - 0.7592j, & k_{n4}(j\omega) &= -0.6508 - 0.7592j, \\
 k_{d1}(j\omega) &= -0.3254 + 0.0542j, & k_{d2}(j\omega) &= -0.1085 + 0.0542j, \\
 k_{d3}(j\omega) &= -0.1085 + 0.2712j, & k_{d4}(j\omega) &= -0.3254 + 0.2712j.
 \end{aligned} \tag{5.8}$$

Table 2 shows the results of the algorithm for the value set at frequency $\omega = 1.0$.

From these Kharitonov rectangles the value set given in Figure 11(a) is directly obtained.

5.2. Value Set at Frequency $\omega = 1.1$

The complete arc with vertices $v_1 = n_1/d_1 = -2.8422 + 2.9830j$ and $v_2 = n_1/d_2 = -0.9808 + 2.4599j$ is taken as initial element. Then Theorems 4.3 and 4.5 will be applied. So

$$v_{2\text{succ}} = \frac{n_{2\lambda}}{d_2} = 3.0420j, \quad v_{1\text{pred}} = \frac{n_{4\lambda}}{d_1} = -3.8291 + 3.7385j. \tag{5.9}$$

Applying the arc normalization (Lemma 4.1) the following data are obtained:

$$\begin{aligned}
 \varphi(n_1) &= 360 - \arg\left(\frac{1}{v_2} - \frac{1}{v_1}\right) = 81.05, & n_1 &= 0.1556 + 0.9878j, \\
 d_1 &= \frac{n_1}{v_1} = 0.1475 - 0.1927j;
 \end{aligned} \tag{5.10}$$

$$\begin{aligned}
 d_2 &= \frac{n_1}{v_2} = 0.3247 - 0.1927j, & n_{4\lambda} &= d_1 v_{1\text{pred}} = 0.1556 + 1.2895j, \\
 n_{2\lambda} &= d_2 v_{2\text{succ}} = 0.5862 + 0.9878j.
 \end{aligned} \tag{5.11}$$

Then, all the other vertices are assigned as follows.

(1) Vertex $v_3 = v_x = n_x/d_x = 0.4996 + 3.0386j$. Then $v_{x\text{pred}} = n_x/d_{x\text{pred}} = 3.0420j$. These are the vertices of an arc, and Theorem 4.3 is applied: $\varphi(A_x) = 8.95$.

Table 2: Results of the algorithm for the value set at frequency $\omega = 1.0$.

	$v_1 - v_2$ arc	v_3	v_4	v_5	v_6	v_7	Kharitonov rectangles calculated
		$0.8000 + 10.4000j$	$-4.8000 + 7.6000j$	$-3.5862 + 1.0345j$	$-2.5517 + 0.6207j$	$-1.3443 + 1.2131j$	
		Theorem 4.5	Theorem 4.5	Theorem 4.3	Theorem 4.5	Theorem 4.3	
v_1	$1.5676 + 2.5946j$	Theorem applied					$k_{n1}(j\omega) = -0.6508 - 1.0847j$
v_2	$2.0000 + 8.0000j$	$v_{x \text{ pred}}$	$2.0000 + 8.0000j$	$-4.8000 + 7.6000j$	$-3.5862 + 1.0345j$	$-2.5517 + 0.6207j$	$k_{n2}(j\omega) = 0.1085 - 1.0847j$
$v_{2 \text{ succ}}$	$0.8000 + 10.4000j$	$v_{x \text{ succ}}$	$10.0000j$	$-2.5517 + 0.6207j$	$-1.3443 + 1.2131j$	$2.3336j$	$k_{n3}(j\omega) = 0.1085 - 0.7592j$
$v_{1 \text{ pred}}$	$0 + 2.3336j$	$\varphi(A_x)$	210.97	174.29	261.87	261.87	$k_{n4}(j\omega) = -0.6508 - 0.7592j$
$\varphi(n_1)$	229.40	Condition verified	Theorem 4.5(C1)	Theorem 4.3(C3)	Theorem 4.5(C3)	Theorem 4.3(C4)	$k_{d1}(j\omega) = -0.3254 + 0.0542j$
n_1	$-0.6508 - 0.7592j$	n_x	$0.1084 - 1.0847j$	$0.1085 - 1.0846j$	$0.1085 - 0.7592j$	$0.1085 - 0.7592j$	$k_{d2}(j\omega) = -0.1085 + 0.0542j$
d_1	$-0.3254 + 0.0542j$	d_x	$-0.1085 + 0.0542j$	$-0.1085 + 0.2712j$	$-0.1085 + 0.2712j$	$-0.3254 + 0.2712j$	$k_{d3}(j\omega) = -0.1085 + 0.2712j$
d_2	$-0.1085 + 0.0542j$	v_x	n_2/d_2	n_3/d_3	n_4/d_3	n_4/d_4	$k_{d4}(j\omega) = -0.3254 + 0.2712j$
n_{41}	$-0.1266 - 0.7594j$						
n_{21}	$-0.6508 - 1.0846j$						

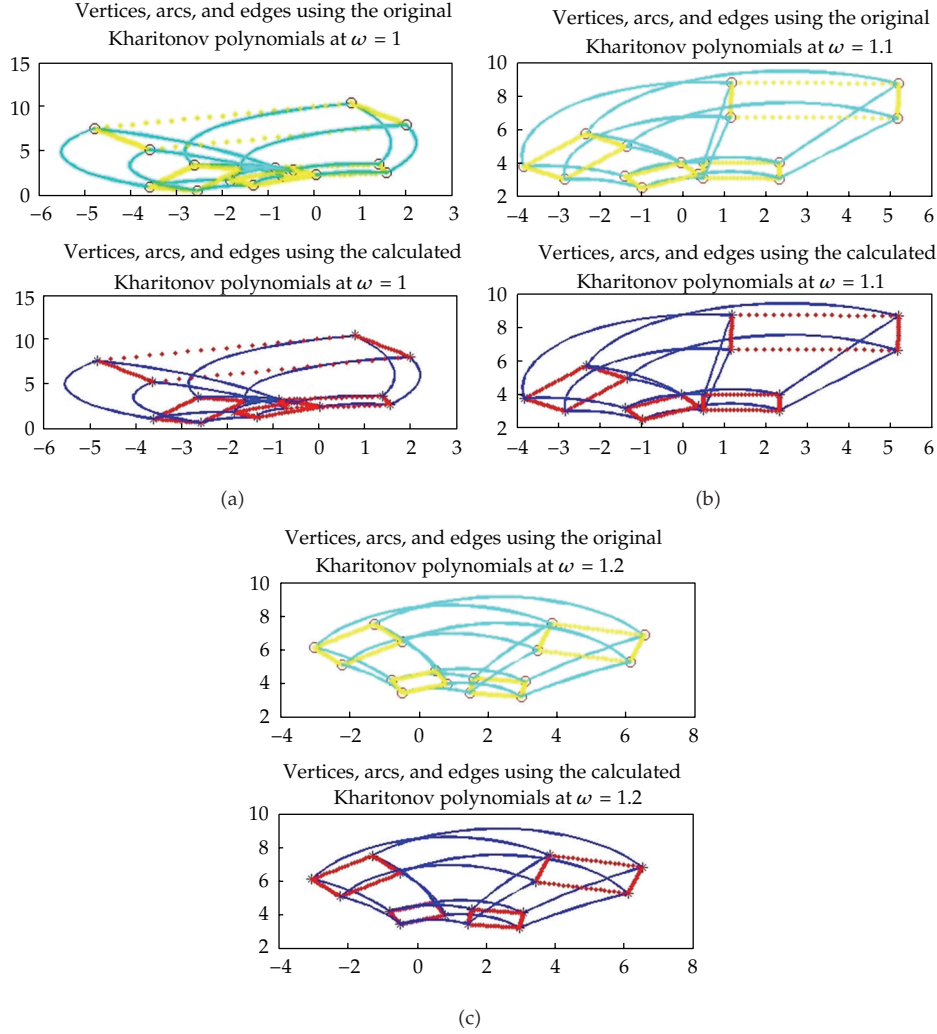


Figure 11

Cases 1 and 2. Theorem 4.3(C1) and (C2) are not satisfied.

Case 3. Theorem 4.3(C3) is satisfied: $\tan(\arg(1/v_x) - \varphi(A_x) + 90)d_{2R} = 0.0022 > d_{2I} = -0.1927$ and $n_x = 0.1555 + 0.9878j$ satisfies the Numerator Condition (Lemma 4.2(1)) $n_x = n_1$. Then $d_x = d_3 = 0.3247 + 0.0022j$ $v_3 = v_x = n_1/d_3$.

(2) Vertex $v_4 = v_x = n_x/d_x = 2.3317 + 3.0261j$. Then $v_{xpred} = n_x/d_{xpred} = 0.4996 + 3.0386j$. These are the vertices of an edge, and Theorem 4.5 is applied: $v_{xsucc} = 5.1859 + 6.6181j$ and $\varphi(A_x) = 127.23$.

Cases 1 and 2. Theorem 4.5(C1) and (C2) are not satisfied.

Case 3. Theorem 4.5(C3) is satisfied: $\tan(\arg(1/v_x) - \varphi(A_x) + 180)d_{2R} = 0.0022 > d_{2I} = -0.1927$ and $n_x = 0.7505 + 0.9878j$ satisfies the Numerator Condition (Lemma 4.2(1)) $n_x = n_2$. Then $d_x = d_3 = 0.3247 + 0.0022j$. $v_4 = v_x = n_2/d_3$.

(3) Vertex $v_5 = v_x = n_x/d_x = 5.1859 + 6.6181j$. Then $v_{xpred} = n_x/d_{xpred} = 2.3317 + 3.0261j$. These are the vertices of an arc, and Theorem 4.3 is applied: $\varphi(A_x) = 127.23$.

Cases 1, 2, and 3. Theorem 4.3(C1), (C2), and (C3) are not satisfied.

Case 4. Theorem 4.3(C4) is satisfied: $\tan(\arg(1/v_x) - \varphi(A_x) + 180)d_{1R} = 0.0022 > d_{1I} = -0.1927$ and $n_x = 0.7505 + 0.9878j$ satisfies the Numerator Condition (Lemma 4.2(1)) $n_x = n_2$. Then $d_x = d_4 = 0.1475 + 0.0022j$ $v_5 = v_x = n_2/d_4$.

(4) Vertex $v_6 = v_x = n_x/d_x = 5.2164 + 8.6623j$. Then $v_{xpred} = n_x/d_{xpred} = 5.1859 + 6.6181j$. These are the vertices of an edge, and Theorem 4.5 is applied: $v_{xsucc} = 8.7404j$, $\varphi(A_x) = 210.20$.

Cases 1, 2, and 3. Theorem 4.5(C1), (C2), and (C3) are not satisfied.

Case 4. Theorem 4.5(C4) is satisfied: $\tan(\arg(1/v_x) - \varphi(A_x) + 270)d_{1R} = 0.0022 > d_{1I} = -0.1927$ and $n_x = 0.7505 + 1.2895j$ satisfies the Numerator Condition (Lemma 4.2(1)) $n_x = n_3$: then $d_x = d_4 = 0.1475 + 0.0022j$, $v_6 = v_x = n_3/d_4$.

(5) Vertex $v_7 = v_x = n_x/d_x = -3.8291 + 3.7385j$. Then $v_{xpred} = n_x/d_{xpred} = 8.7404j$. These are the vertices of an arc, and Theorem 4.3 is applied: $\varphi(A_x) = 186.88$.

Case 1. Theorem 4.3(C1) is not satisfied.

Case 2. Theorem 4.3(C2) is satisfied: $\arg(v_1/v_x) = \arg(n_1) + \varphi(A_x) + 90 = 357.93$ and $n_x = d_1v_x = 0.1556 + 1.2895j$ satisfies the Numerator Condition (Lemma 4.2(1)) $n_x = n_4$. Then $d_x = d_1 = 0.1475 - 0.1927j$, $v_7 = v_x = n_4/d_1$.

In summary, the assigned polynomials are

$$v_1 = \frac{n_1}{d_1}, \quad v_2 = \frac{n_1}{d_2}, \quad v_3 = \frac{n_1}{d_3}, \quad v_4 = \frac{n_2}{d_3}, \quad v_5 = \frac{n_2}{d_4}, \quad v_6 = \frac{n_3}{d_4}, \quad v_7 = \frac{n_4}{d_1} \quad (5.12)$$

and the values can be calculated: from normalization,

$$n_1 = 0.1556 + 0.9878j, \quad d_1 = 0.1475 - 0.1927j, \quad d_2 = 0.3247 - 0.1927j, \quad (5.13)$$

and from the vertices,

$$\begin{aligned} v_3 : n_1 &= 0.1555 + 0.9878j, & d_3 &= 0.3247 + 0.0022j, \\ v_4 : n_2 &= 0.7505 + 0.9878j, & d_3 &= 0.3247 + 0.0022j, \\ v_5 : n_2 &= 0.7505 + 0.9878j, & d_4 &= 0.1475 + 0.0022j, \\ v_6 : n_3 &= 0.7505 + 1.2895j, & d_4 &= 0.1475 + 0.0022j, \\ v_7 : n_4 &= 0.1556 + 1.2895j, & d_1 &= 0.1475 - 0.1927j. \end{aligned} \quad (5.14)$$

Then

$$\begin{aligned}
 k_{n1}(j\omega) &= 0.1555 + 0.9878j, & k_{n2}(j\omega) &= 0.7505 + 0.9878j, \\
 k_{n3}(j\omega) &= 0.7505 + 1.2895j, & k_{n4}(j\omega) &= 0.1556 + 1.2895j, \\
 k_{d1}(j\omega) &= 0.1475 - 0.1927j, & k_{d2}(j\omega) &= 0.3247 - 0.1927j, \\
 k_{d3}(j\omega) &= 0.3247 + 0.0022j, & k_{d4}(j\omega) &= 0.1475 + 0.0022j.
 \end{aligned} \tag{5.15}$$

Table 3 shows the results of the algorithm for the value set at frequency $\omega = 1.1$.

From these Kharitonov rectangles the value set given in Figure 11(b) is directly obtained.

5.3. Value Set at Frequency $\omega = 1.2$

The complete edge with vertices $v_1 = n_1/d_1 = 6.1015 + 5.2779j$ and $v_2 = n_2/d_1 = 6.5135 + 6.8573j$ is taken as initial element. Then Theorems 3.3 and 3.5 will be applied. So

$$v_{2\text{succ}} = \frac{n_2}{d_{2\lambda}} = 8.5560j, \quad v_{1\text{pred}} = \frac{n_1}{d_{4\lambda}} = 2.9559 + 3.2369j. \tag{5.16}$$

Applying the edge normalization (Lemma 3.1) the following data are obtained:

$$\begin{aligned}
 \phi(d_1) &= 360 - \arg(v_2 - v_1) = 284.62, & d_1 &= \cos(\phi(d_1)) + j \sin(\phi(d_1)) = 0.2524 - 0.9676j, \\
 n_1 &= v_1 d_1 = 6.6471 - 4.5717j, & n_2 &= v_2 d_1 = 8.2793 - 4.5717j, \\
 d_{2\lambda} &= \frac{n_2}{v_{2\text{succ}}} = -0.5343 - 0.9677j, & d_{4\lambda} &= \frac{n_1}{v_{1\text{pred}}} = 0.2524 - 1.8230j.
 \end{aligned} \tag{5.17}$$

Then, all the other vertices are assigned as follows.

(1) Vertex $v_3 = v_x = n_x/d_x = -3.0339 + 6.1294j$. Then $v_{x\text{pred}} = n_x/d_{x\text{pred}} = 8.5560j$. These are the vertices of an arc, and Theorem 3.5 is applied: $v_{x\text{succ}} = -2.2110 + 5.1007j$ and $\varphi(S_x) = \arg(v_{x\text{succ}} - v_x) = 308.66$.

Cases 1 and 2. Theorem 3.5(C1) and (C2) are not satisfied.

Case 3. Theorem 3.5(C3) is satisfied: $\tan(\arg(v_x) - \varphi(S_x) + 180)n_{2R} = -1.8087 > n_{2I} = -4.571$ and $d_x = n_{2R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 180)]/v_x = -0.7740 - 0.9676j$ satisfies the Denominator Condition (Lemma 3.2(3)) $d_x = d_2$: then $n_x = n_3 = n_{2R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 180)] = 8.2793 - 1.8087j$; $v_3 = v_x = n_3/d_2$.

(2) Vertex $v_4 = v_x = n_x/d_x = -2.211 + 5.1007j$. Then $v_{x\text{pred}} = n_x/d_{x\text{pred}} = -3.0339 + 6.1294j$. These are the vertices of an edge, and Theorem 3.3 is applied: $\varphi(S_x) = \arg(v_x - v_{x\text{pred}}) = 308.66$.

Cases 1 and 2. Theorem 3.3(C1) and (C2) are not satisfied.

Case 3. Theorem 3.3(C3) is satisfied: $\tan(\arg(v_x) - \varphi(S_x) + 90)n_{2R} = 30.4258 > n_{2I} = -4.5717$ but $d_x = n_{2R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 90)]/v_x = 4.4292 - 3.5431j$ does not satisfy the Denominator

Table 3: Results of the algorithm for the value set at frequency $\omega = 1.1$.

	$v_1 - v_2$ arc	v_3	v_4	v_5	v_6	v_7	Kharitonov rectangles calculated
v_1	$-2.8422 + 2.9830j$	$0.4996 + 3.0386j$	$2.3317 + 3.0261j$	$5.1859 + 6.6181j$	$5.2164 + 8.6623j$	$-3.8291 + 3.7385j$	
v_2	$-0.9808 + 2.4599j$	Theorem 4.3 3.0420j	Theorem 4.5 0.4996 + 3.0386j	Theorem 4.3 2.3317 + 3.0261j	Theorem 4.5 5.1859 + 6.6181j	Theorem 4.3 8.7404j	$k_{n1}(j\omega) = 0.1555 + 0.9878j$
$v_{2\text{succ}}$	$0 + 3.0420j$	$2.3317 + 3.0261j$	$5.1859 + 6.6181j$	$5.2164 + 8.6623j$	$8.7404j$	$-2.8422 + 2.9830j$	$k_{n2}(j\omega) = 0.7505 + 0.9878j$
$v_{1\text{pred}}$	$-3.8291 + 3.7385j$	8.95	127.23	127.23	210.20	186.88	$k_{n3}(j\omega) = 0.7505 + 1.2895j$
$\varphi(n_1)$	81.05	Condition verified	Theorem 4.3(C3)	Theorem 4.5(C3)	Theorem 4.3(C4)	Theorem 4.5(C4)	$k_{n4}(j\omega) = 0.1556 + 1.2895j$
n_1	$0.1556 + 0.9878j$	$-0.1555 + 0.9878j$	$0.7505 + 0.9878j$	$0.7505 + 0.9878j$	$0.7505 + 1.2895j$	$0.1556 + 1.2895j$	$k_{d1}(j\omega) = 0.1475 - 0.1927j$
d_1	$0.1475 - 0.1927j$	$0.3247 + 0.0022j$	$0.3247 + 0.0022j$	$0.1475 + 0.0022j$	$0.1475 + 0.0022j$	$0.1475 - 0.1927j$	$k_{d2}(j\omega) = 0.3247 - 0.1927j$
d_2	$0.3247 - 0.1927j$	n_1/d_3	n_2/d_3	n_2/d_4	n_3/d_4	n_4/d_1	$k_{d3}(j\omega) = 0.3247 + 0.0022j$
$n_{4\lambda}$	$0.1556 + 1.2895j$						$k_{d4}(j\omega) = 0.1475 + 0.0022j$
$n_{2\lambda}$	$0.5862 + 0.9878j$						

Condition: ($d_{1R} = 0.2524 > d_{2\lambda R} = -0.5343$ and $d_{1I} = -0.9676 > d_{4\lambda I} = -1.8230$) (Case 3) but ($d_{xR} = 4.4292 \neq d_{1R} = 0.2524$) then $d_x \neq d_1$ and $d_x \neq d_4$ ($d_{xI} = -3.5431 \neq d_{1I} = -0.9676$) then $d_x \neq d_2$ ($d_{xR} = 4.4292 > d_{1R} = 0.2524$ and $d_{xI} = -3.5431 < d_{1I} = -0.9676$) then $d_x \neq d_3$.

Case 4. Theorem 3.3(C4) is satisfied: $\tan(\arg(v_x) - \varphi(S_x) + 180)n_{1R} = -1.8088 > n_{1I} = -4.5717$ and $d_x = n_{1R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 180)]/v_x = -0.7741 - 0.9676j$ satisfies the Denominator Condition (Lemma 3.2(3)) $d_x = d_2$:

$$\begin{aligned} (d_{1R} = 0.2524 > d_{2\lambda R} = -0.5343, d_{1I} = -0.9676 > d_{4\lambda I} = -1.8230), \\ (d_{xI} = d_{1I} = -0.9676, d_{xR} = -0.7740 \leq d_{1R} = 0.2524). \end{aligned} \quad (5.18)$$

Then $n_x = n_4 = n_{1R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 180)] = 6.6471 - 1.8088j$, $v_4 = v_x = n_4/d_2$.

(3) Vertex $v_5 = v_x = n_x/d_x = -0.47099 + 3.4462j$. Then $v_{xpred} = n_x/d_{xpred} = -2.2110 + 5.1007j$. These are the vertices of an arc, and Theorem 3.5 is applied: $v_{xsucc} = 3.6463j$ and $\varphi(S_x) = \arg(v_{xsucc} - v_x) = 23.01$.

Cases 1, 2, and 3. Theorems 3.5(C1) and (C2) are not satisfied. Theorem 3.5(C3) is satisfied but $d_x = 8.3397 - 3.5422j$ does not satisfy the Denominator Condition (Lemma 3.2(3)).

Case 4. Theorem 3.5(C4) is satisfied: $\tan(\arg(v_x) - \varphi(S_x) + 270)n_{1R} = -1.8098 > n_{1I} = -4.5717$ and $d_x = n_{1R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 270)]/v_x = -0.7743 - 1.8230j$ satisfies the Denominator Condition (Lemma 3.2(3)).

Then $n_x = n_4 = n_{1R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 270)] = 6.6471 - 1.8098j$, $v_5 = v_x = n_4/d_3$.

(4) Vertex $v_6 = v_x = n_x/d_x = 1.469 + 3.4428j$. Then $v_{xpred} = n_x/d_{xpred} = 3.6463j$. These are the vertices of an arc, and Theorem 3.5 is applied: $v_{xsucc} = 2.9559 + 3.2369j$ and $\varphi(S_x) = \arg(v_{xsucc} - v_x) = 352.12$.

Cases 1, 2, and 3. Theorem 3.5(C1) and (C2) are not satisfied. Theorem 3.5(C3) is satisfied but $d_x = 8.3440 + 1.1554j$ does not satisfy the Denominator Condition (Lemma 3.2(3)).

Case 4. Theorem 3.5(C4) is satisfied: $\tan(\arg(v_x) - \varphi(S_x) + 270)n_{1R} = -1.8089 > n_{1I} = -4.5717$ and $d_x = 0.2524 - 1.8230j$ satisfies the Denominator Condition (Lemma 3.2(3)) $d_x = d_4$.

Then $n_x = n_4 = n_{1R}[1 + j \tan(\arg(v_x) - \varphi(S_x) + 270)] = 6.6471 - 1.8089j$, $v_6 = v_x = n_4/d_4$.

(5) Vertex $v_7 = v_x = n_x/d_x = 2.9559 + 3.2369j$. Then $v_{xpred} = n_x/d_{xpred} = 1.4690 + 3.4428j$. These are the vertices of an edge, and Theorem 3.3 is applied: $\varphi(S_x) = \arg(v_x - v_{xpred}) = 352.12$.

Case 1. Theorem 3.3(C1) is not satisfied.

Case 2. Theorem 3.3(C2) is satisfied: $\arg(v_x/v_1) = \arg(d_1) + \varphi(S_x) + 90 = 6.74$ and $d_x = n_1/v_x = 0.2524 - 1.8230j$ satisfies the Denominator Condition (Lemma 3.2(3)) $d_x = d_4$. Then $n_x = n_1 = 6.6471 - 4.5717j$, $v_7 = v_x = n_1/d_4$.

In summary, the assigned polynomials are

$$v_1 = \frac{n_1}{d_1}, \quad v_2 = \frac{n_2}{d_1}, \quad v_3 = \frac{n_3}{d_2}, \quad v_4 = \frac{n_4}{d_2}, \quad v_5 = \frac{n_4}{d_3}, \quad v_6 = \frac{n_4}{d_4}, \quad v_7 = \frac{n_1}{d_4}, \quad (5.19)$$

Table 4: Results of the algorithm for the value set at frequency $\omega = 1.2$.

	$v_1 - v_2$ edge	v_3	v_4	v_5	v_6	v_7	Kharitonov rectangles calculated
		$-3.0339 + 6.1294j$	$-2.211 + 5.1007j$	$-0.4710 + 3.4462j$	$-1.469 + 3.4428j$	$2.9559 + 3.2369j$	
		Theorem 3.5	Theorem 3.3	Theorem 3.5	Theorem 3.5	Theorem 3.3	
v_1	$6.1015 + 5.2779j$	Theorem applied	Theorem 3.3	Theorem 3.5	Theorem 3.5	Theorem 3.3	$k_{n1}(j\omega) = -6.6471 - 4.5717j$
v_2	$6.5135 + 6.8573j$	$8.5560j$	$-3.0339 + 6.1294j$	$-2.211 + 5.1007j$	$-3.6463j$	$1.4690 + 3.4428j$	$k_{n2}(j\omega) = 8.2793 - 4.5717j$
$v_{2\text{succ}}$	$8.5560j$	$-2.211 + 5.1007j$	$-0.4710 + 3.4462j$	$3.6463j$	$2.9559 + 3.2369j$	$6.1015 + 5.2779j$	$k_{n3}(j\omega) = 8.2793 - 1.8087j$
$v_{1\text{pred}}$	$2.9559 + 3.2369j$	308.66	308.66	23.01	352.12	352.12	$k_{n4}(j\omega) = 6.6471 - 1.8087j$
$\varphi(d_1)$	284.62	Condition verified	Theorem 3.3(C4)	Theorem 3.5(C4)	Theorem 3.5(C4)	Theorem 3.3(C2)	$k_{d1}(j\omega) = -0.7743 - 1.8230j$
d_1	$0.2524 - 0.9676j$	$-0.7741 - 0.9676j$	$-0.7741 - 0.9676j$	$-0.7743 - 1.8230j$	$0.2524 - 1.8230j$	$0.2524 - 1.8230j$	$k_{d2}(j\omega) = 0.2524 - 1.8230j$
n_1	$6.6471 - 4.5717j$	$8.2793 - 1.8088j$	$6.6471 - 1.8088j$	$6.6471 - 1.8098j$	$6.6471 - 1.8089j$	$6.6471 - 4.5717j$	$k_{d3}(j\omega) = 0.2524 - 0.9676j$
n_2	$8.2793 - 4.5717j$	n_3/d_2	n_4/d_2	n_4/d_3	n_4/d_4	n_1/d_4	$k_{d4}(j\omega) = -0.7743 - 0.9676j$
$d_{4\lambda}$	$0.2524 - 1.8230j$						
$d_{2\lambda}$	$-0.5343 - 0.9677j$						

and the values can be calculated: from normalization,

$$d_1 = 0.2524 - 0.9676j, \quad n_1 = 6.6471 - 4.5717j, \quad n_2 = 8.2793 - 4.5717j, \quad (5.20)$$

and from the vertices,

$$\begin{aligned} v_3 : n_3 &= 8.2793 - 1.8087j, & d_2 &= -0.7741 - 0.9676j, \\ v_4 : n_4 &= 6.6471 - 1.8087j, & d_2 &= -0.7741 - 0.9676j, \\ v_5 : n_4 &= 6.6471 - 1.8087j, & d_3 &= -0.7743 - 1.8230j, \\ v_6 : n_4 &= 6.6471 - 1.8087j, & d_4 &= 0.2524 - 1.8230j, \\ v_7 : n_1 &= 6.6471 - 4.5717j, & d_4 &= 0.2524 - 1.8230j. \end{aligned} \quad (5.21)$$

Then

$$\begin{aligned} k_{n1}(j\omega) &= 6.6471 - 4.5717j, & k_{n2}(j\omega) &= 8.2793 - 4.5717j, \\ k_{n3}(j\omega) &= 8.2793 - 1.8087j, & k_{n4}(j\omega) &= 6.6471 - 1.8087j, \\ k_{d1}(j\omega) &= -0.7743 - 1.8230j, & k_{d2}(j\omega) &= 0.2524 - 1.8230j, \\ k_{d3}(j\omega) &= 0.2524 - 0.9676j, & k_{d4}(j\omega) &= -0.7743 - 0.9676j. \end{aligned} \quad (5.22)$$

Table 4 shows the results of the algorithm for the value set at frequency $\omega = 1.2$.

From these kharitonov rectangles the value set given in Figure 11(c) is directly obtained.

Finally, solving the equation system [10, equation (16)], the interval plant is obtained:

$$G_p(s) = \frac{[10 \ 11]s^3 + [7 \ 8]s^2 + [6 \ 6.5]s + [5 \ 7.5]}{[0.75 \ 1.25]s^3 + [2 \ 2.5]s^2 + [1.5 \ 2]s + [1 \ 1.5]}. \quad (5.23)$$

Applying $G_p(s = j\omega)$ at $\omega = 1.0$, $\omega = 1.1$ and $\omega = 1.2$ the value sets given in Figure 12 are obtained.

6. Conclusions

This paper shows how to obtain the values of the numerator and denominator Kharitonov polynomials of an interval plant from its value set at a given frequency. Moreover, it is proven that given a value set, all the assigned polynomials of the vertices can be determined if and only if there is a complete edge or a complete arc lying on a quadrant, that is, if there are two vertices in a quadrant. This necessary and sufficient condition is not restrictive and practically all the value sets satisfy it. Finally, the interval plant can be identified solving the equation system between the Kharitonov rectangles and the parameters of the plant.

The algorithm has been formulated using the frequency domain properties of linear interval systems. The identification procedure of multilinear (affine, polynomial) systems will be studied using the results in [11].

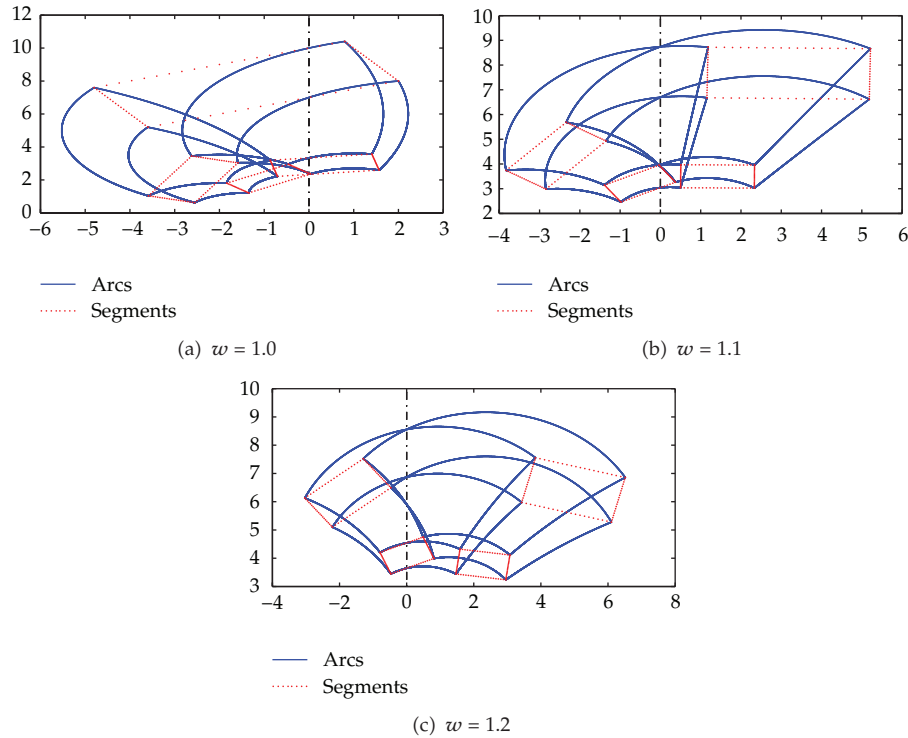


Figure 12: Value sets obtained at $w = 1.0$, $w = 1.1$, and $w = 1.2$.

Acknowledgments

The authors would like to express their gratitude to Dr. José Mira and to Dr. Ana Delgado for their example of ethics and professionalism, without which this work would not have been possible. Also, the authors are very grateful to the Editor-in-Chief, Zhiwei Gao, and the referees, for their suggestions and comments that very much enhanced the presentation of this paper.

References

- [1] Z. Gao, X. Dai, T. Breikin, and H. Wang, "Novel parameter identification by using a high-gain observer with application to a gas turbine engine," *IEEE Transactions on Industrial Informatics*, vol. 4, no. 4, pp. 271–279, 2008.
- [2] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems*, Kluwer Academic, Boston, Mass, USA, 1996.
- [3] R. S. Sánchez-Peña and M. Sznaier, *Robust Systems Theory and Applications*, John Wiley & Sons, 1998.
- [4] B. R. Barmish, *New Tools for Robustness of Linear Systems*, MacMillan, New York, NY, USA, 1993.
- [5] S. P. Bhattacharyya, H. Chapellat, and L. H. Keel, *Robust Control: The Parametric Approach*, Prentice-Hall, 1995.
- [6] A. C. Bartlett, C. V. Hollot, and H. Lin, "Root locations of an entire polytope of polynomials: it suffices to check the edges," *Mathematics of Control, Signals, and Systems*, vol. 1, no. 1, pp. 61–71, 1988.
- [7] M. Fu and B. R. Barmish, "Polytopes of polynomials with zeros in a prescribed set," *IEEE Transactions on Automatic Control*, vol. 34, no. 5, pp. 544–546, 1989.

- [8] B. R. Barmish, "A generalization of Kharitonov's four-polynomial concept for robust stability problems with linearly dependent coefficient perturbations," *IEEE Transactions on Automatic Control*, vol. 34, no. 2, pp. 157–165, 1989.
- [9] L. Ljung, *System Identification: Theory for the User*, PTR Prentice Hall, 1999.
- [10] R. Hernández, J. A. García, and A. P. de Madrid, "Interval plant identification from value sets with five vertices in a quadrant," *International Journal of Robust and Nonlinear Control*, vol. 21, no. 1, pp. 21–43, 2011.
- [11] N. Tan, "Computation of the frequency response of multilinear affine systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 10, pp. 1691–1696, 2002.

Research Article

Improving the Solution of Least Squares Support Vector Machines with Application to a Blast Furnace System

Ling Jian, Shuqian Shen, and Yunquan Song

College of Science, China University of Petroleum, Qingdao 266580, China

Correspondence should be addressed to Ling Jian, bebetter@upc.edu.cn

Received 4 May 2012; Revised 23 August 2012; Accepted 20 September 2012

Academic Editor: Chuanhou Gao

Copyright © 2012 Ling Jian et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The solution of least squares support vector machines (LS-SVMs) is characterized by a specific linear system, that is, a saddle point system. Approaches for its numerical solutions such as conjugate methods Sykens and Vandewalle (1999) and null space methods Chu et al. (2005) have been proposed. To speed up the solution of LS-SVM, this paper employs the minimal residual (MINRES) method to solve the above saddle point system directly. Theoretical analysis indicates that the MINRES method is more efficient than the conjugate gradient method and the null space method for solving the saddle point system. Experiments on benchmark data sets show that compared with mainstream algorithms for LS-SVM, the proposed approach significantly reduces the training time and keeps comparable accuracy. To heel, the LS-SVM based on MINRES method is used to track a practical problem originated from blast furnace iron-making process: changing trend prediction of silicon content in hot metal. The MINRES method-based LS-SVM can effectively perform feature reduction and model selection simultaneously, so it is a practical tool for the silicon trend prediction task.

1. Introduction

As one kernel method, SVM works by embedding the input data $\mathbf{x}, \mathbf{z} \in X$ into a Hilbert space \mathcal{H} by a high-dimensional mapping $\Phi(\cdot)$, and then trying to find a linear relation among the high-dimensional embedded data points [1, 2]. This process is implicitly performed by specifying a kernel function which satisfies $k(\mathbf{x}, \mathbf{z}) = \Phi(\mathbf{x})^T \Phi(\mathbf{z})$, that is, the inner product of the embedded points. Given observed samples $\{\mathbf{x}_i, y_i\}_{i=1}^n$ with size n , SVM formulates the learning problem as a variational problem of finding a decision function f that minimizes the regularized risk functional [3, 4]

$$\min_{f \in \mathcal{H}} R[f] = \frac{1}{n} \sum_{i=1}^n V(y_i, f(\mathbf{x}_i)) + \lambda \|f\|_{\mathcal{H}}^2, \quad (1.1)$$

where $V(\cdot, \cdot)$ is called a loss function, λ is the so-called regularization parameter to trade off the empirical risk with the complexity of f , that is, $\|f\|_{\mathcal{H}}$, the norm in a reproducing kernel Hilbert space \mathcal{H} . By the representer theorem [3, 5], the optimal decision function f satisfying (1.1) has the form

$$f(\mathbf{x}) = \sum_{i=1}^n \alpha_i k(\mathbf{x}_i, \mathbf{x}) + b, \quad (1.2)$$

where $\alpha_i \in \mathcal{R}$ for $i = 1, \dots, n$, $b \in \mathcal{R}$. This equation can be easily used to tackle a practical problem if the kernel function is specified. To overcome the high computational complexity of traditional SVM, an interesting variant of the standard SVM, least squares support vector machines, has been proposed by Suykens and Vandewalle [6]. In the case of LS-SVM, the inequality constraints in ℓ_2 soft margin SVM are converted into equality constraints. The model training process of LS-SVM is performed by solving a specific linear equations, that is, a saddle point system which can be efficiently solved by iterative methods instead of a quadratic programming problem. Besides computational superiority extensive empirical studies have shown that LS-SVM is comparable to SVM in terms of generalization performance [7]; these features make LS-SVM an attractive algorithm and also a successful alternative to SVM. For the training of the LS-SVM, Van Gestel et al. [7] proposed to reformulate the $n + 1$ order saddle point system into two n order symmetric positive definite systems which can be solved in turn by the conjugate gradient (CG) algorithm. To speed up the training of LS-SVM, Chu et al. [8] employed the null space method to transform the saddle point system into a reduced $n - 1$ order symmetric positive definite system which was solved with the CG algorithm also. The minimal residual (MINRES) method proposed by Paige and Saunders is a specialized method for solving a nonsingular symmetric system [9]. This method can avoid the LU factorization and does not suffer from break-down, so it is an efficient numerical method for solving symmetric but indefinite systems. The Karush-Kuhn-Tucker system of LS-SVM is a specified linear system, that is, a saddle point system. Considering the above point, to speed up the solution of LS-SVM model we employ the MINRES method to solve the linear system directly. The main contribution of this paper is to provide a potential alternative to the solution of LS-SVM model. Theoretical analysis of the three numerical algorithms for the solution of LS-SVM model indicates that the MINRES method is the optimal choice. Experiments on benchmark data sets show that compared with the CG method proposed by Suykens et al. and the null space method proposed by Chu et al., the MINRES solver significantly improves the computational efficiency and at the same time keeps almost the same generalized performance with the above two methods. To heel, the MINRES method-based LS-SVM model is constructed and further employed to identify blast furnace (BF) iron-making process, a complex nonlinear system. Practical application to a typical real BF indicates that the established MINRES method-based LS-SVM model is a good candidate to predict the changing trend of the silicon content in BF hot metal with low time cost. The possible application of this work is to aid the BF operators to judge the inner state of BF getting hot or chilling in time properly, which can provide a guide for them to determine the direction of controlling BF in advance. The rest of this paper is organized as follows. In Section 2, we give a review for LS-SVM. Section 3 presents three numerical solutions for LS-SVM. It is followed by extensive experimental validations of the proposed method in Section 4. Section 5 concludes the paper and points out the possible future research.

2. Formulation of LS-SVM

The primal problem of LS-SVM can be formulated following unified format:

$$\begin{aligned} \min_{\mathbf{w}, b, \mathbf{e}} \quad & \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{C}{2} \sum_{i=1}^n e_i^2 \\ \text{s.t.} \quad & y_i = \mathbf{w}^T \Phi(\mathbf{x}_i) + b + e_i, \quad i = 1, \dots, n, \end{aligned} \quad (2.1)$$

for both regression analysis and pattern classification. In (2.1) n is the total number of training samples, \mathbf{x}_i is the i th input vector, y_i is the i th output value/label for regression/classification problem, e_i is the i th error variable, $C > 0$ is the regularization parameter, and b is the bias term. The Lagrangian of (2.1) is given below:

$$L(\mathbf{w}, b, \mathbf{e}; \boldsymbol{\alpha}) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{C}{2} \sum_{i=1}^n e_i^2 - \sum_{i=1}^n \alpha_i (\mathbf{w}^T \Phi(\mathbf{x}_i) + b + e_i - y_i), \quad (2.2)$$

where α_i is the i th Lagrange multiplier. For the convex program (2.1), it is obvious that the Slater constraint qualification holds. Therefore, the optimal solution of (2.1) satisfies its Karush-Kuhn-Tucker system

$$\begin{aligned} \frac{\partial L}{\partial \mathbf{w}} = 0 &\longrightarrow \mathbf{w} = \sum_{i=1}^n \alpha_i \Phi(\mathbf{x}_i), \\ \frac{\partial L}{\partial b} = 0 &\longrightarrow \sum_{i=1}^n \alpha_i = 0, \\ \frac{\partial L}{\partial e_i} = 0 &\longrightarrow \alpha_i = C e_i, \quad i = 1, \dots, n, \\ \frac{\partial L}{\partial \alpha_i} = 0 &\longrightarrow y_i = \mathbf{w}^T \Phi(\mathbf{x}_i) + b + e_i, \quad i = 1, \dots, n. \end{aligned} \quad (2.3)$$

After eliminating variables \mathbf{w} and \mathbf{e} the Karush-Kuhn-Tucker system (2.3) can be reformulated following saddle point system [10]:

$$\begin{bmatrix} K + \frac{1}{C} I & \mathbf{1}_n \\ \mathbf{1}_n^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} \\ b \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix}, \quad (2.4)$$

where $K_{ij} := k(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i)^T \Phi(\mathbf{x}_j)$, I stands for unit matrix, $\mathbf{1}_n$ denotes an n -dimensional vector of all ones, and $\mathbf{y} = (y_1, \dots, y_n)^T$.

3. Solution of LS-SVM

In this section, we give a brief review and some analysis of the three mentioned numerical algorithms for solution of LS-SVM.

3.1. Conjugate Gradient Methods

The kernel matrix K is a symmetric positive semidefinite matrix and the diagonal term $1/C$ is positive, so the matrix $H := K + (1/C)I$ is symmetric and positive definite. Through the following matrix transformation

$$M^T \begin{bmatrix} K + \frac{1}{C}I & \mathbf{1}_n \\ \mathbf{1}_n^T & 0 \end{bmatrix} M M^{-1} \begin{bmatrix} \boldsymbol{\alpha} \\ b \end{bmatrix} = M^T \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix}, \quad (3.1)$$

where

$$M = \begin{bmatrix} I & H^{-1}\mathbf{1}_n \\ 0 & 1 \end{bmatrix}, \quad (3.2)$$

the saddle point system (2.4) can be factorized into a positive definite system [11]

$$\begin{bmatrix} H & 0 \\ 0 & \mathbf{1}_n^T H^{-1} \mathbf{1}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} + H^{-1} \mathbf{1}_n b \\ b \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ \mathbf{1}_n^T H^{-1} \mathbf{y} \end{bmatrix}. \quad (3.3)$$

Suykens et al. suggested the use of the CG method for the solution of (3.3) and proposed to solve two n order positive definite systems. More exactly, their algorithm can be described as follows.

Step 1. Employ the CG algorithm to solve the linear equations $H\boldsymbol{\eta} = \mathbf{1}_n$ and get the intermediate variable $\boldsymbol{\eta}$.

Step 2. Solve the intermediate variable $\boldsymbol{\mu}$ from $H\boldsymbol{\mu} = \mathbf{y}$ by the CG method.

Step 3. Obtain Lagrange dual variables $\boldsymbol{\alpha} = \boldsymbol{\mu} - b\boldsymbol{\eta}$ and bias term $b = \mathbf{1}_n^T \boldsymbol{\mu} / \mathbf{1}_n^T \boldsymbol{\eta}$.

The output of any new data \mathbf{x} can subsequently be deduced by computing the decision function $f(\mathbf{x}) = \mathbf{w}^T \Phi(\mathbf{x}) + b = \sum_{i=1}^n \alpha_i k(\mathbf{x}_i, \mathbf{x}) + b$.

3.2. Null Space Methods

In what was mentioned previously, to get the intermediate variable $\boldsymbol{\eta}$ and $\boldsymbol{\mu}$ two n order positive definite systems need to be solved by CG methods. Chu et al. [8] proposed an interesting method to the numerical solution of LS-SVM by solving one $n - 1$ order reduced system of linear equations. The improved method suggested by Chu et al. can be seen as one kind of null space method. The saddle point system (2.4) can be written as

$$H\boldsymbol{\alpha} + \mathbf{1}_n b = \mathbf{y}, \quad \mathbf{1}_n^T \boldsymbol{\alpha} = 0. \quad (3.4)$$

Chu et al. specified a particular solution of $\mathbf{1}_n \boldsymbol{\alpha} = 0$ as $\hat{\boldsymbol{\alpha}} = \mathbf{0}$ and the null space of $\mathbf{1}_n \boldsymbol{\alpha} = 0$ as

$$Z = \begin{bmatrix} I_{n-1} \\ -\mathbf{1}_{n-1}^T \end{bmatrix}. \quad (3.5)$$

Through solving the following reduced system of order $n - 1$ for the auxiliary unknown \mathbf{v} ,

$$Z^T H Z \mathbf{v} = Z^T \mathbf{y}, \quad (3.6)$$

the solution of the saddle point system (2.4) can be obtained as $\boldsymbol{\alpha} = Z \mathbf{v}$ and $b = (1/n) \mathbf{1}_n^T (\mathbf{y} - H \boldsymbol{\alpha})$.

3.3. Minimal Residual Methods

The vector sequences in the CG method correspond to a factorization of a tridiagonal matrix similar to the coefficient matrix. Therefore, a breakdown of the algorithm can occur corresponding to a zero pivot if the matrix is indefinite. Furthermore, for indefinite matrices the minimization property of the CG method is no longer well defined. The MINRES method proposed by Paige and Saunders [9] is a variant of the CG method that avoids the LU factorization and does not suffer from breakdown. It minimizes the residual in the ℓ_2 -norm which is an efficient numerical algorithm for solving symmetric but indefinite systems; the corresponding convergence behavior of the MINRES method for indefinite systems has been analyzed by Van der Vorst [12]. The purpose of this paper is to employ the MINRES method to solve the saddle point system (2.4) directly. Next we gave a brief review of the MINRES algorithm. Let \mathbf{x}_0 be an initial guess for the solution of the symmetric indefinite linear system $A\mathbf{x} = \mathbf{b}$. One can obtain the iterative sequence \mathbf{x}_m , $m = 1, 2, \dots$ such that

$$\|\mathbf{r}_m\|_2 = \|\mathbf{b} - A\mathbf{x}_m\|_2 = \min_{\mathbf{x} \in \mathbf{x}_0 + \mathcal{K}_m(A, \mathbf{r}_0)} \|\mathbf{b} - A\mathbf{x}\|_2, \quad (3.7)$$

where $\mathbf{r}_m = \mathbf{b} - A\mathbf{x}_m$ is the m th residual for $m = 1, 2, \dots$, and

$$\mathcal{K}_m(A, \mathbf{r}_0) = \text{span}\{\mathbf{r}_0, A\mathbf{r}_0, \dots, A^{m-1}\mathbf{r}_0\} \quad (3.8)$$

is the m th Krylov subspace. Lanczos methods can be used to generate an orthonormal basis of $\mathcal{K}_m(A, \mathbf{r}_0)$, and then only two basis vectors are needed to compute \mathbf{x}_m ; see, for example, [12]. The detailed implementation of the MINRES algorithm can be found in [12].

It has been shown that rounding errors are propagated to the approximate solution with a factor proportional to the square of the condition number of coefficient matrix [12]; one should be careful with the MINRES method for ill-conditioned systems.

3.4. Some Analysis on These Three Numerical Algorithms

The properties of short recurrences and optimization [12] make the CG method the first choice for the solution of a symmetric positive definite system. Suykens et al. transformed

the $n + 1$ order saddle point system (2.4) into two n order positive definite systems which are solved by CG methods. However, it is time consuming to solve two n order positive definite systems with large scales. To overcome this shortcoming, Chu et al. [8] transformed equivalently the original $n + 1$ order system into an n order symmetric positive definite system, and then the CG method can be used. This method can be seen as a null space method. Unfortunately, the transformation may destroy heavily the sparse structure and increase greatly the condition number of the original system. This can hugely slow down the convergence rate of the CG algorithm. Theoretical analysis about the influence of the transformation on the condition number is indispensable, but it is rather difficult. We leave it as an open problem. In this paper, the MINRES method is directly applied to solve the original saddle point problem of $n + 1$ order. Similar to the CG method, the MINRES method also has properties of short recurrences and optimization.

In light of the analysis mentioned above, the MINRES method should be the first choice for the solution of LS-SVM model, since it avoids solving two linear systems and destroying the sparse structure of the original saddle point system simultaneously.

4. Numerical Implementations

4.1. Experiments on Benchmark Data Sets

In this section we give the experimental test results on the accuracy and efficiency of our method. For comparison purpose, we implement the CG method proposed by Suykens and Vandewalle [6] and the null space method suggested by Chu et al. [8]. All experiments are implemented with MATLAB version 7.8 programming environment running on an IBM compatible PC under Window XP operating system, which is configured with Intel Core 2.1Ghz CPU and 2G RAM. The generalized used Gaussian RBF kernel $k(\mathbf{x}, \mathbf{z}) = \exp(-\|\mathbf{x} - \mathbf{z}\|^2 / \sigma^2)$ is selected as the kernel function. We use the default setting for kernel width σ^2 , that is, set kernel width as the dimension of inputs.

We first compare three algorithms on three benchmark data sets: Boston, Concrete, and Abalone, which are download from UCI [13]. Each data set is randomly partitioned into 70% training and 30% test sets. We also list the condition numbers of coefficients matrices solved by three methods for the analysis of the computing efficiencies. As shown in Tables 1–3 the condition number for the CG method is the least one and the condition number for the null space method significantly increases.

The columns of Cond in Tables 1, 2, and 3 show that compared with the CG method the condition number for the MINRES method increases a bit, but much less than the condition number of the null space method. The orders of linear equations solved by the CG method, the null space method, and the MINRES method are n , $n - 1$, and $n + 1$, respectively. The condition numbers for the CG method and the MINRES method are very close, but we have to solve two systems of $n - 1$ order using CG methods. Hence, the running time of the MINRES method should be less than that of the CG method. CPU column in Tables 1–3 shows that the MINRES method-based LS-SVM model costs much less running time than the CG method and the null space method-based LS-SVM model in all cases of setting C. So the MINRES method-based LS-SVM model is a preferable algorithm for solving LS-SVM model. In the next subsection, we will employ the MINRES method-based LS-SVM model to solve a practical problem.

Table 1: Experimental results of three methods on Boston data set.

$\log_2 C$	Boston data set, 506 samples, 13-d inputs, σ^2 equals 13								
	Conjugate gradient method			Null space method			MINRES method		
	Cond [†]	CPU [‡]	MSE*	Cond	CPU	MSE	Cond	CPU	MSE
-5	4	0.3281	49.1027	366.9451	0.8438	49.1027	45.6283	0.2500	49.1027
-4	8	0.4688	39.4132	369.0926	0.6250	39.4132	31.6150	0.3438	39.4132
-3	15	0.3438	29.7686	388.1770	0.7656	29.7686	26.3388	0.3125	29.7686
-2	28	0.4531	24.2532	460.1920	0.7656	24.2532	29.2813	0.3281	24.2532
-1	60	0.2500	21.0322	474.6254	1.0625	21.0322	61.3493	0.4219	21.0322
0	116	0.3438	15.5875	566.2504	1.2500	15.5875	119.071	0.1875	15.5875
1	234	0.7188	13.6449	946.4564	1.1250	13.6449	239.374	0.4375	13.6449
2	472	0.9531	13.0252	1945.300	1.0625	13.0252	482.447	0.6875	13.0252
3	924	0.9375	10.9810	2244.342	1.4063	10.9810	944.042	0.6406	10.9810
4	1734	1.3594	10.3168	5229.460	1.2500	10.3168	1776.31	0.8906	10.3168
5	3801	1.5469	10.2063	10785.92	1.4844	10.2063	3876.97	1.1406	10.2063
6	7530	2.0469	11.3937	24998.71	1.9063	11.3937	7682.07	1.2969	11.3937
7	14618	2.4531	11.7750	47781.41	2.2188	11.7750	14932.2	1.6875	11.7750
8	29769	3.0625	12.9925	61351.85	2.9844	12.9925	30382.8	2.3750	12.9925
9	58387	3.4063	14.0194	101181.8	3.5938	14.0194	59619.0	2.6875	14.0194
10	119285	4.0313	17.2330	285440.0	4.8281	17.2330	121708	3.5313	17.2330

Cond[†] denotes the condition number, CPU[‡] stands for running time, MSE* is mean square error.

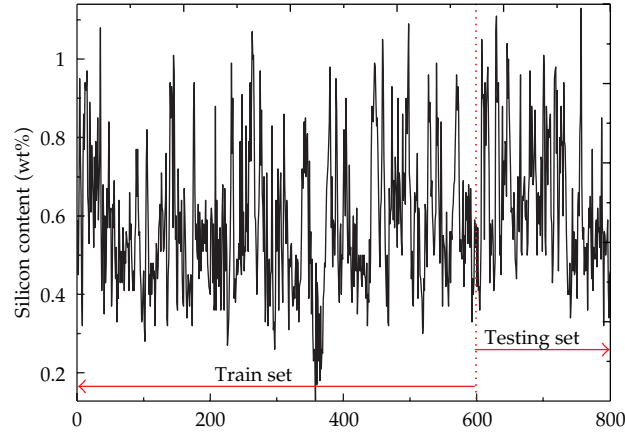
4.2. Application on Blast Furnace System

Blast furnace, one kind of metallurgical reactor used for producing pig iron, is often called hot metal. The chemical reactions and heat transport phenomena take place throughout the furnace as the solid materials move downwards and hot combustion gases flow upwards. The main principle involved in the BF iron-making process is the thermochemical reduction of iron oxide ore by carbon monoxide. During the iron-making period, a great deal of heat energy is produced which can heat up the BF temperature approaching 2000°C. The end products consisting of slag and hot metal sink to the bottom and are tapped periodically for the subsequent refining. It will take about 6–8 h for a cycle of iron-making [11]. BF iron-making process is a highly complex nonlinear process with the characteristics of high temperature, high pressure, concurrence of transport phenomena, and chemical reactions. The complexity of the BF and the occurrence of a variety of process disturbances have been obstacles for the adoption of modeling and control in the process. Generally speaking, to control a BF system often means to control the hot metal temperature and components, such as silicon content, sulfur content in hot metal, and carbon content in hot metal within acceptable bounds. Among these indicators, the silicon content often acts as a chief indicator to represent the thermal state of the BF, an increasing silicon content meaning a heating of the BF while a decreasing silicon content indicating a cooling of the BF [11, 14]. Thus, the silicon content is a reliable measure of the thermal state of the BF, and it becomes a key stage to predict the silicon content for regulating the thermal state of the BF. Therefore, it has been the active research issue to build silicon prediction model in the recent decades, including numerical prediction models [15] and trend prediction models [11].

In this subsection, the tendency prediction of silicon content in hot metal is transformed as a binary classification problem. Samples with increasing silicon content are

Table 2: Experimental results of three methods on Concrete data set.

$\log_2 C$	Concrete data set, 1030 samples, 8-d inputs, σ^2 equals 8								
	Conjugate gradient method			Null space method			MINRES method		
	Cond	CPU	MSE	Cond	CPU	MSE	Cond	CPU	MSE
-5	7	2.0781	140.8498	738.137223	3.6719	140.8498	51.5204280	1.6406	140.8498
-4	13	2.3594	111.2384	745.054714	3.6563	111.2384	39.7005872	1.8906	111.2383
-3	25	2.8125	89.3458	796.627246	3.8281	89.3458	31.7895802	2.0938	89.3459
-2	50	2.4844	74.4146	850.938881	3.9688	74.4146	51.4318149	2.0469	74.4146
-1	102	3.0000	60.2984	954.604170	4.3906	60.2984	104.293122	2.2969	60.2984
0	199	3.4219	50.4491	1397.13474	4.8438	50.4491	202.202543	2.8281	50.4490
1	399	4.0625	43.5416	1737.97400	5.7188	43.5416	406.110983	3.2500	43.5416
2	787	4.8750	41.5463	2369.65769	6.4219	41.5463	799.643656	3.8594	41.5463
3	1561	6.3125	36.5797	5375.87469	7.3750	36.5797	1586.70628	4.2500	36.5797
4	3197	8.0000	33.4861	7342.75323	8.7188	33.4861	3247.47638	5.0156	33.4861
5	6411	10.3281	33.1452	18274.6591	10.8438	33.1452	6510.73913	6.2188	33.1452
6	12530	13.8750	33.4936	37192.6189	12.9063	33.4936	12732.7611	8.2813	33.4936
7	25614	18.5781	33.8690	73008.8645	15.8594	33.8690	26010.1838	11.0156	33.8690
8	51260	25.1250	32.6925	126475.189	19.5938	32.6925	52056.7280	14.8906	32.6925
9	101053	33.9531	35.1044	249234.605	25.2969	35.1044	102657.615	19.9219	35.1043
10	199734	46.3125	40.4777	557864.123	32.9219	40.4777	202961.396	27.0625	40.4777

**Figure 1:** Evolution of silicon content in hot metal.

denoted by +1 whereas a decreasing silicon content is denoted by -1. In the present work, the experimental data is collected from a medium-sized BF with the inner volume of about 2500m^3 . The variables closely related to the silicon content are measured as the candidate inputs for modeling. Table 4 presents the variables information from the studied BF. There are totally 801 data points collected with the first 601 points as train set and the residual 200 points as testing set. The sampling interval is about 1.5 h for the current BF. Figure 1 illustrates the evolution of the silicon content in hot metal.

There are in total 15 candidate variables listed in Table 4 from which to select model inputs. Generally, too many input parameters will increase the complexity of model while too little inputs will reduce the accuracy of model. A tradeoff has to be taken between the

Table 3: Experimental results of three methods on Abalone data set.

$\log_2 C$	Abalone data set, 4177 samples, 7-d inputs, σ^2 equals 7								
	Conjugate gradient method			Null space method			MINRES method		
	Cond	CPU	MSE	Cond	CPU	MSE	Cond	CPU	MSE
-5	42.341	20.343	5.3623	2955.9655	39.7344	5.3623	369.9433	12.9531	5.3623
-4	84.059	22.984	5.1143	3028.4495	41.1406	5.1143	332.0862	14.2344	5.1143
-3	167.846	26.343	4.7978	3043.4615	42.6719	4.7978	306.5691	16.0156	4.7978
-2	337.691	33.281	4.6923	3567.1431	46.8125	4.6923	338.1679	18.4688	4.6923
-1	666.823	39.315	4.4360	4888.3227	50.8438	4.4360	667.7842	22.3125	4.4360
0	1327.351	47.531	4.4744	5355.5805	54.6563	4.4744	1329.291	26.2656	4.4744
1	2700.547	58.015	4.4217	9894.2450	59.8438	4.4217	2704.345	34.1719	4.4217
2	5275.703	74.859	4.3948	8239.2388	69.2031	4.3948	5283.506	42.5469	4.3948
3	10709.216	94.765	4.4169	18279.897	80.4219	4.4169	10724.46	54.7813	4.4169
4	21357.750	124.359	4.5053	24472.420	97.7500	4.5053	21388.43	71.3906	4.5053
5	42427.822	177.171	4.6144	105161.60	133.2656	4.6144	42489.70	103.6406	4.6144
6	85153.757	221.468	4.6857	185913.18	155.9219	4.6857	85276.97	129.3750	4.6857
7	171369.064	312.078	4.7145	212162.90	212.4531	4.7145	171614.1	181.8750	4.7145
8	344731.082	430.640	4.8621	705659.56	289.4531	4.8621	345216.0	260.5469	4.8621
9	681509.920	602.765	5.2294	1162595.7	395.6250	5.2294	682494.5	360.5625	5.2294
10	1363883.053	840.625	5.6517	3106655.0	549.4844	5.6517	1365853	488.6250	5.6517

Table 4: A list of input variables.

Variable name [unit]	Abbreviation	Range	F -score	Mean accuracy
Latest silicon content (wt%)	Si	0.13–1.13	0.1269	81.786%
Sulfur content (wt%)	S	0.012–0.077	0.0570	82.857%
Basicity of ingredients (wt%)	BI	0.665–1.609	0.0229	81.786%
Feed speed (mm/h)	FS	16.725–297.510	0.0132	83.214%
Blast volume (m^3/min)	BV	1454.30–5580.200	0.0054	83.747%
CO ₂ percentage in top gas (wt%)	CO ₂	7.921–22.892	0.0048	83.750%
Pulverized coal injection (ton)	PCI	0.230–98.533	0.0037	83.214%
CO percentage in top gas (wt%)	CO	9.267–27.374	0.0036	82.500%
Blast temperature (°C)	BT	1086.100–1239.700	0.0031	83.571%
Oxygen enrichment percentage (wt%)	OEP	−0.001–14.688	0.0019	83.393%
H ₂ percentage in top gas (wt%)	H ₂	2.564–4.065	0.0005	83.214%
Coke load of ingredients (wt%)	CLI	2.032–5.071	0.0004	82.857%
Furnace top temperature (°C)	TP	62.703–264.130	0.0002	82.679%
Blast pressure (kPa)	BP	59.585–367.780	0.0001	83.214%
Furnace top pressure (kPa)	TP	8.585–199.790	0.0001	82.679%

model complexity and accuracy when selecting the inputs. Therefore, it is necessary to screen out less important variables as inputs from these 15 candidate variables. Here, the inputs are screened out by an integrative way that combines F -score method [16] for variables ranking and cross-validation method for variables and model parameters selection.

F -score is an effective tool for feature selection in data mining and can give feature ranking by evaluating the discrimination of two sets with real values. For those 15 candidate variables in Table 4, their F -scores are defined as follows:

$$F_s(i) = \left(\frac{1}{N_+ - 1} \sum_{j=1}^{N_+} (x_{j,+}^{(i)} - \bar{x}_+^{(i)})^2 + \frac{1}{N_- - 1} \sum_{j=1}^{N_-} (x_{j,-}^{(i)} - \bar{x}_-^{(i)})^2 \right)^{-1} \cdot \left((\bar{x}_+^{(i)} - \bar{x}^{(i)})^2 + (\bar{x}_-^{(i)} - \bar{x}^{(i)})^2 \right), \quad i = 1, \dots, d, \quad (4.1)$$

where $\bar{x}^{(i)}$, $\bar{x}_+^{(i)}$ and $\bar{x}_-^{(i)}$ stand for the mean of the i th attribute of the whole training, positive and negative examples, respectively, while $\bar{x}_{j,+}^{(i)}$ and $\bar{x}_{j,-}^{(i)}$ are the i th variable of the j th positive and negative instance, respectively. Hence, a variable ranking can be achieved through F -score method. Table 4 gives the results of F -scores of all 15 variables, which are ranked according to the F -score values. As one kernel-based learning model, the kernel parameter σ^2 , and regularized parameter C play an important role in LS-SVM, so one should pay attention to selecting proper parameters. Grid search-based ten-fold cross-validation is executed on the train set for searching the optimal (σ^2, C) . The searching grid for model parameters is set as

$$\left[2^{-5}, 2^{-4}, \dots, 2^{10} \right] \times \left[2^{-5}, 2^{-4}, \dots, 2^{10} \right]. \quad (4.2)$$

Mean accuracy in Table 4 stands for the average accuracy under ten-fold cross-validation experiments of LS-SVM model on some grid points with the best performance. In the current work, we first select the variable with highest F -score as model input and then add variables one by one according to their F -scores. Mean accuracy under all kinds of input variables can be achieved and the results are shown in Table 4. The following are shown by the mean accuracy column: (1) at the beginning, the mean accuracy increases gradually as more candidate variables are taken as model inputs; (2) the largest mean accuracy appears when CO_2 is included within the input set; (3) when the mean accuracy is beyond the maximum, it will fluctuate as the residual variables are added by turns into the input set. These results indicate that, as the studied BF is concerned, the optimal input set is [Si, S, BI, FS, BV, CO_2] with the model parameters setting $(\sigma^2, C) = (2^9, 2^8)$. Table 5 lists the LS-SVM model accuracy including with/without feature and model selection versions on testing set. In the case of without feature and model selection version, all candidate variables are selected as inputs, and we use the default setting for LS-SVM model; that is, set kernel width σ^2 equal to the dimension of input variable and set regularized parameter C as 1. The information in the second row of this table, such as 34/42, denotes that there are 42 times predicted results that are ascending trend, and 34 times predictions are successful. The confidence level of the LS-SVM model without model and feature selection fluctuates severely between the ascending and descending prediction from 80.95% to 58.86%. The difference of confidence levels of LS-SVM model with model and feature selection between ascending and descending prediction is reduced to 2.19% indicating that model and feature selection procedure enhances the stability of the LS-SVM model obviously. As the last column of Table 5 shows, TSA of LS-SVM model with feature and model selection procedure is significantly improved compared with LS-SVM model without feature and model selection, so the selection procedure is indispensable for the current practical application. Table 6 lists the running time of three

Table 5: Predictive results of LS-SVM model with/without feature and model selection.

Inputs	(σ^2, C)	Ascend (99*)	Descend (101)	TSA [†]
15	(15, 1)	34/42 = 80.95%	93/158 = 58.86%	127/200 = 63.5%
6	(2 ⁹ , 2 ⁸)	73/94 = 77.66%	80/106 = 75.47%	153/200 = 76.5%

99* means 99 observations are ascending trend; TSA[†] stands for testing set accuracy.

Table 6: Running time of three numerical methods on model identification.

Algorithm	Conjugate gradient method	Null space method	MINRES method
CPU	1948	2800	1488

mentioned numerical algorithms when performing feature and model selection procedure. The cost time of the MINRES method is reduced significantly compared with the other algorithms. In a word, the feature and model selection procedure can be effectively performed for the MINRES method-based LS-SVM, and it is meaningful for practical using.

5. Conclusions and Points of Possible Future Research

In this paper, we have proposed an alternative, that is, the MINRES method, to the solution of LS-SVM model which is formulated as a saddle point system. Numerical experiments on UCI benchmark data sets show that the proposed numerical solution method of LS-SVM model is more efficient than the algorithms proposed by Suykens and Vandewalle [6] and Chu et al. [8]. To heel, the MINRES method-based LS-SVM model including feature selection from extensive candidate and model parameter selection is proposed and employed for the silicon content trend prediction task. The practical application to a typical real BF indicates that the proposed MINRES method-based LS-SVM model is a good candidate to predict the trend of silicon content in BF hot metal with low running time.

However, it should be pointed out that despite the MINRES method-based LS-SVM model displaying low running time, lack of metallurgical information may be the root to the limited accuracy of the current prediction model. So there is much work worth investigating in the future to further improve the model accuracy and increase the model transparency, such as constructing predictive model by integrating domain knowledge and extracting rules. The extracted rules can account for the output results with detailed and definite inputs information, which may further serve for the control purpose by linking the output results with controlled variables. These investigations are deemed to be helpful to further improve the efficiency of predictive model.

Acknowledgment

This work was partially supported by National Natural Science Foundation of China under Grant no. 11126084, Natural Science Foundation of Shandong Province under Grant no. ZR2011AQ003, Fundamental Research Funds for the Central Universities under Grant no. 12CX04082A, and Public Benefit Technologies R&D Program of Science and Technology Department of Zhejiang Province under Grant No. 2011C31G2010136.

References

- [1] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag, New York, NY, USA, 2nd edition, 2000.
- [2] B. Schölkopf and A. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, MIT Press, Cambridge, Mass, USA, 2002.
- [3] T. Evgeniou, M. Pontil, and T. Poggio, "Regularization networks and support vector machines," *Advances in Computational Mathematics*, vol. 13, no. 1, pp. 1–50, 2000.
- [4] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines*, Cambridge University, Cambridge, UK, 2000.
- [5] C. M. Bishop, *Pattern Recognition and Machine Learning*, vol. 4, Springer, New York, NY, USA, 2006.
- [6] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Processing Letters*, vol. 9, no. 3, pp. 293–300, 1999.
- [7] T. Van Gestel, J. A. K. Suykens, B. Baesens et al., "Benchmarking least squares support vector machine classifiers," *Machine Learning*, vol. 54, no. 1, pp. 5–32, 2004.
- [8] W. Chu, C. J. Ong, and S. S. Keerthi, "An improved conjugate gradient scheme to the solution of least squares SVM," *IEEE Transactions on Neural Networks*, vol. 16, no. 2, pp. 498–501, 2005.
- [9] C. C. Paige and M. A. Saunders, "Solutions of sparse indefinite systems of linear equations," *SIAM Journal on Numerical Analysis*, vol. 12, no. 4, pp. 617–629, 1975.
- [10] L. Jian, C. Gao, L. Li, and J. Zeng, "Application of least squares support vector machines to predict the silicon content in blast furnace hot metal," *ISIJ International*, vol. 48, no. 11, pp. 1659–1661, 2008.
- [11] C. Gao, L. Jian, and S. Luo, "Modeling of the thermal state change of blast furnace hearth with support vector machines," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 2, pp. 1134–1145, 2012.
- [12] H. A. van der Vorst, *Iterative Krylov Methods for Large Linear Systems*, vol. 13, Cambridge University Press, Cambridge, UK, 2003.
- [13] C. Blake and C. Merz, "Uci repository of machine learning databases," 1998.
- [14] C. Gao, J. L. Chen, J. Zeng, X. Liu, and Y. Sun, "A chaos-based iterated multistep predictor for blast furnace ironmaking process," *AIChE Journal*, vol. 55, no. 4, pp. 947–962, 2009.
- [15] L. Jian, C. Gao, and Z. Xia, "A sliding-window smooth support vector regression model for nonlinear blast furnace system," *Steel Research International*, vol. 82, no. 3, pp. 169–179, 2011.
- [16] Y. W. Chen and C. J. Lin, "Combining SVMs with various feature selection strategies," *Studies in Fuzziness and Soft Computing*, vol. 207, pp. 315–324, 2006.

Research Article

Adaptive Fault Detection for Complex Dynamic Processes Based on JIT Updated Data Set

Jinna Li,^{1,2} Yuan Li,³ Haibin Yu,²
Yanhong Xie,¹ and Cheng Zhang¹

¹ Department of Science, Shenyang University of Chemical Technology, Liaoning,
Shenyang 110142, China

² Lab of Industrial Control Networks and Systems, Shenyang Institute of Automation,
Chinese Academy of Sciences, Liaoning, Shenyang 110016, China

³ College of Information Engineering, Shenyang University of Chemical Technology, Liaoning,
Shenyang 110142, China

Correspondence should be addressed to Jinna Li, lijinna.721@yahoo.com.cn

Received 2 May 2012; Revised 1 July 2012; Accepted 2 July 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 Jinna Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A novel fault detection technique is proposed to explicitly account for the nonlinear, dynamic, and multimodal problems existed in the practical and complex dynamic processes. Just-in-time (JIT) detection method and k-nearest neighbor (KNN) rule-based statistical process control (SPC) approach are integrated to construct a flexible and adaptive detection scheme for the control process with nonlinear, dynamic, and multimodal cases. Mahalanobis distance, representing the correlation among samples, is used to simplify and update the raw data set, which is the first merit in this paper. Based on it, the control limit is computed in terms of both KNN rule and SPC method, such that we can identify whether the current data is normal or not by online approach. Noted that the control limit obtained changes with updating database such that an adaptive fault detection technique that can effectively eliminate the impact of data drift and shift on the performance of detection process is obtained, which is the second merit in this paper. The efficiency of the developed method is demonstrated by the numerical examples and an industrial case.

1. Introduction

Fault has been a constant topic of research for several decades [1–4]. Several fault detection methods have been developed to solve problems since there exists a growing need for fault detection in the real process engineering not only from the plant's safety perspective but also from considering the quality of the process products [5–7]. Moreover, the existing methods used to fault detection have been applied into a broad range of areas such as chemical process,

networked control systems and semiconduction process, and so forth [8–11]. The dynamic change, multiple mode, and nonlinearity exist objectively in the most of the aforementioned process, such as semiconduction process, tank reactors, and so forth [9–11], which has brought new challenges to the analysis and implementation of fault detection. Therefore, they must be taken into account carefully in developing a high-performance and adaptive fault identification method to detect the abnormal cases as early as possible.

As summarized in [12], the technologies of process analysis and operation are derived broadly into two categories: model-based approach and data-based approach. In the model-based approach, static or dynamic models are built for the process under normal operating situation. The difference between the actual process output and nominal model's output is monitored to determine whether any fault occurs or not [5, 13–18]. Noted that many control processes are data rich but information poor, which senses the data-based method is strongly needed to obtain a flexible and high-efficiency detection manager systems. Among the reported results in the literature, to mention a few, data-driven KNN fault detection was addressed in [9]. Based on it, [10] proposed an improved principal component analysis (PCA) KNN technique to implement the fault identification. Noted that data-based fault monitoring and identification methods were also investigated in [11, 18, 19].

It should be pointed out the nonlinearity, dynamics, and multimode are the inevitable obstacles for either model-based approach or data-based approach when process detection is required in the real-world applications. Various methods, including statistical process control (SPC), multivariate statistical process control (MSPC), qualitative knowledge-based methods, artificial intelligence, and various integrated methods, have been developed and performed in the available literature [20–23]. Well known that PCA is used in MSPC widely, the existing nonlinear PCA method [24, 25], dynamic PCA method [26], and independent component analysis PCA method [18] have been presented to address the nonlinearity, dynamics, and multimode faced by fault detection, respectively. Since JIT is inherently adaptive in nature, which is achieved by storing the current measured data in the database [27, 28], and is capable of detecting and diagnosing whether the query is normal or not by on-line and adaptive approach [19], as shown in Figure 1, data-based JIT method has attracted much more scholars' attention in the recent years (see [12]). Reference [12] proposed a data-based JIT-SPC detection and identification technique, in which the distance was calculated and checked every time when fault detection need to be conducted. References [9, 10] applied KNN rule to fault detection for semiconductor manufactory process, in which k -nearest neighbors were used to tackle the multimodal problem. However, though nonlinear PCA method [24, 25], dynamic PCA method [26], independent component analysis PCA method [18], data-based JIT [12, 18, 19], and KNN methods [9, 10] have been reported, the efficient method which can naturally handle nonlinear correlation among the variables, dynamic change of the systems and multimodal batch trajectories have not been fully investigated so far.

As a matter of fact, it is crucial for realizing the desired performance in detection to determine an appropriate normal operation data stored in database, since too much data will have heavy load on both storage cost and computation, while, less data naturally effect on the efficacy of detection technique. How to determine an appropriate raw data set so that cost, computation complexity, and performance of detection can be compromised is a challenge. Based on the Mahalanobis distance between samples, the simplification of raw data set is studied in [29], in which the simplifying procedure is terminated in terms of the desired number of samples. However, how to determine number of samples that is closely related to quality of detection was not investigated. Well known to all, data drift and shift exist

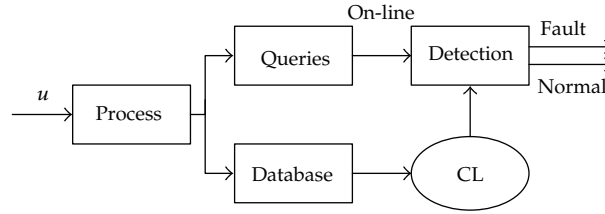


Figure 1: On-line detection scheme.

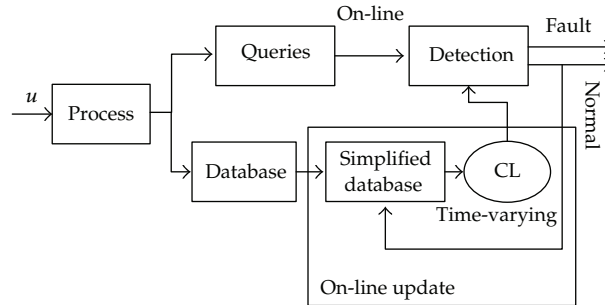


Figure 2: Simplified database-based on-line detection scheme.

in the practical complex dynamic process inevitably, which can be produced due to much more insuperable causes, such as aging of instrument, variation of temperature, effect of environment, and difference in the coming materials, and so forth [30]. In this case, some of queries with drift and shift that might be normal are mistaken for faults, or that might be fault are falsely identified as normalities. Then, it is of paramount importance to actively regulate the control limit (CL) by on-line approach, which will surely improve the quality of fault detection. However, to the best of authors' knowledge, how to simplify and update the raw samples set to light the computation load and realize high performance have not been investigated fully to date. Especially for data-driven fault detection with time-varying control limit, few results have been available in the literature so far, which motivates the present study.

The overall goal in this paper is to propose a flexible, adaptive JIT scheme of fault detection. By computing the Mahalanobis distance between normal samples, as well as Mahalanobis distance from the query to the normal operation data stored in the database, the raw data set is simplified and updated. Moreover, the updated database-based time-varying CL is derived in terms of KNN rule combined with SPC method, such that the judgment is conducted every time when fault detection is required as shown in Figure 2. Moreover, the integration of JIT method and KNN technique is well suited for the nonlinear, dynamic and multimodal fault detection process. Finally, simulation results and industrial case illustrate the efficiency of the method proposed.

2. Simplifying the Raw Data Set

In this section, we will describe the method of simplifying the raw data set.

In the practical detection of the process, the huge amounts of raw data bring the serious calculation load and cost spending for fault detection, identification, and diagnosis.

Since the closer Mahalanobis distance between two samples is, the more similar their basic features are, we take mean of them to retain the common characteristics, and this mean is put into the raw data set to replace the original two samples, which can keep the characteristics of raw data set to the greatest extent [29]. Let \mathbf{X} denotes the raw data matrix with n samples (rows) and m variables (columns). The specific procedure is as follows.

Algorithm 2.1.

Step 1. Let $\mathbf{Z}(n \times m) = \mathbf{X}(n \times m)$, and the covariance matrix of \mathbf{X} is defined as: $\mathbf{V} = (v_{ij}) = ((x_i - \bar{x}_i)'(x_j - \bar{x}_j)/(n-1))$. Where x_i, x_j ($i, j = 1, 2, \dots, m$) denote the stochastic variables, and \bar{x}_i and \bar{x}_j denote the mean of them.

Step 2. Calculate the Mahalanobis distance between every sample and all the other samples stored in data set \mathbf{Z} , which is denoted as md_{ij} ($i, j = 1, 2, \dots, n$) and is placed in Mahalanobis distance matrix $\mathbf{MD}(n \times n) = (md_{ij})$.

Step 3. Find out the minimum and nonzero element in each row in the matrix \mathbf{MD} , which is placed in row vector $\mathbf{v}(1 \times n)$, and the place (column number) of each minimum element in each row is recorded in row vector $\mathbf{p}(1 \times n)$. Based on it, finding out the minimum value in $\mathbf{v}(1 \times n)$, and if it's place in $\mathbf{v}(1 \times n)$ is i and No. i element is j in vector $\mathbf{p}(1 \times n)$, then md_{ij} is the minimum value in the matrix $\mathbf{MD}(n \times n)$, which presents the minimum Mahalanobis distance between the sample i and the sample j in raw data set.

Step 4. Let $\mathbf{M} = \mathbf{Z}$, and we use the mean of the sample i and the sample j to replace the sample i , and delete the sample j , then the row number of the matrix \mathbf{Z} is reduced a line. Define $\Gamma = ((z_i - \bar{z}_i)'(z_j - \bar{z}_j)/(l-1))$, where z_i, z_j ($i, j = 1, 2, \dots, m$), and l denotes the row number of the matrix \mathbf{Z} . If $1 \geq \text{Trace}(\Gamma)/\text{Trace}(\mathbf{V}) \geq \epsilon$, where $0 < \epsilon \leq 1$, return to Step 2; otherwise, \mathbf{M} is the simplified data matrix. Exit.

Pseudocode 1 is given.

Remark 2.2. Motivated by PCA [10, 19, 31], where combination of variables that capture the largest amount of information in data set is found, the inequality $1 \geq \text{Trace}(\Gamma)/\text{Trace}(\mathbf{V}) \geq \epsilon$ realizes the preservation and update of the original information to the greatest extent in Algorithm 2.1. Note that $\text{Trace}(\Gamma)$ and $\text{Trace}(\mathbf{V})$ denote the trace of covariance matrix of simplified or update data set and original data set, respectively, and well known that $\text{Trace}(\mathbf{V})$ is equal to the sum of all the eigenvalues of matrix \mathbf{V} . In this sense that the threshold ϵ is selected to maximize the retention of the originally statistical information. For example, $\epsilon = 0.98$ means that 98% of the variance in raw data set is represented by the new simplified data set \mathbf{Z} . Obviously, Algorithm 2.1 is a kind of logical and promising way for the data-driven fault detection.

Remark 2.3. Different from [29], in which though the raw data is reduced based on Mahalanobis distance between samples, the subjectively determined number of simplified data has an essential effect on high-performance detection process. In this paper, the threshold that limits the extent of reducing the raw data is given, which is a preferable way that compromises the accuracy of fault detection and the computation complexity as well as lower cost.

Remark 2.4. Apparently, simplification and update scheme of raw data set presented for fault detection also apply to data-driven fault monitoring, diagnosis, and isolation, since the proper number of raw data still need to be determined and data drift and shift still needs to be

```

for i = 1 : m
    for j = 1 : m
        cov(i,j) = sum((x(:,i) - mean(x(:,i))).
            *(x(:,j) - mean(x(:,j))))/(size(x,1) - 1);
    end
end
Z = X;
ss1 = eig(cov);
[sm1,sn1] = size(ss1);
lmd1 = 0
for i = 1 : sm1
    lmd1 = lmd1 + ss1(i);
end
while 1
    [p,q] = size(Z)
    for i = 1 : q
        for j = 1 : q
            cov(i,j) = sum((z(:,i) - mean(z(:,i))).
                *(z(:,j) - mean(z(:,j))))/(size(z,1) - 1);
        end
    end
    ss = eig(cov);
    [sm,sn] = size(ss)
    lmd = 0
    for i = 1 : sm
        lmd = lmd + ss(i);
    end
    if lmd/lmd1 < e||lmd/lmd1 > 1
        break
    end
    [n,m] = size(Z);
    d = zeros(n,n);
    M = Z;
    for i = 1 : n
        for j = 1 : n
            d(i,j) = d(i,j) + (z(i,:) - z(j,:)) * inv(cov)
                *(z(i,:) - z(j,:))^T
        end
    end
    md = zeros(n - 1, n - 1);
    for i = 1 : n
        for j = 1 : n
            if d(i,j) > 0
                md(i,j) = d(i,j);
            else
                md(i,j) = 10000;
            end
        end
    end
    [The_Min,The_Min_ID] = min(md')
    v = The_Min;
    p = The_Min_ID;
    [The_Min,The_Min_ID] = min(v);
    i = The_Min_ID;
    j = p(i);
    z(i,:) = (z(i,:) + z(j,:))/2;
    z(j,:) = [ ];
end

```

Pseudocode 1

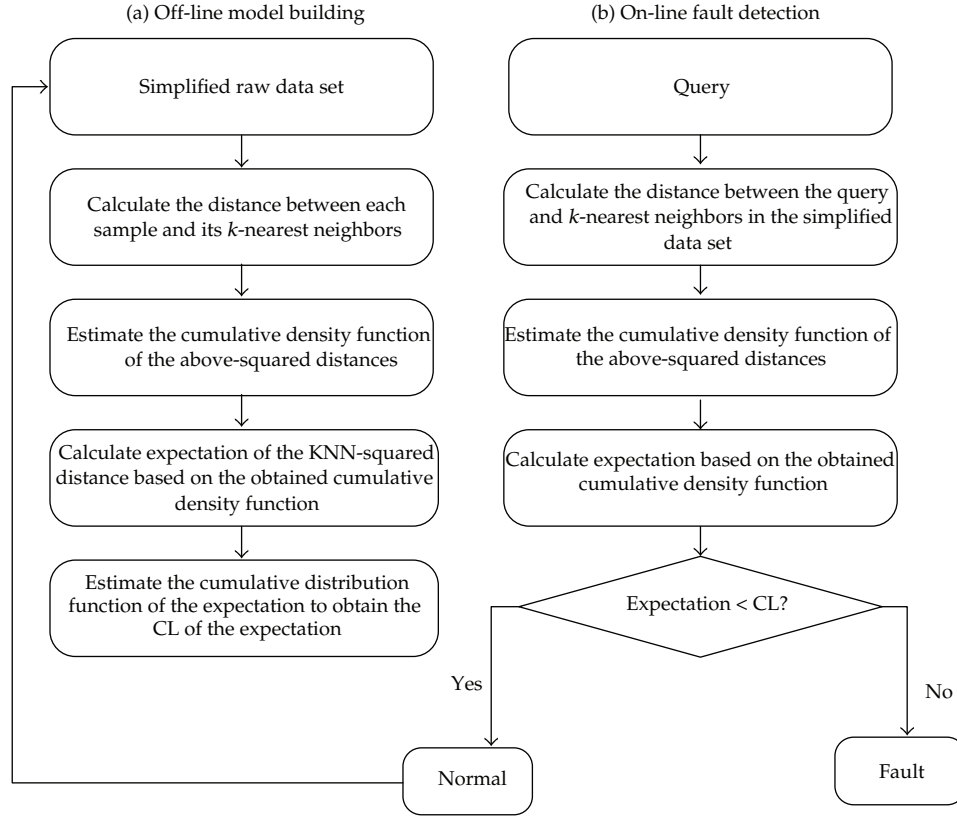


Figure 3: Flow chart of proposed detection method.

coped with. In this sense that the simplification and update technique suggested in this paper is quite general.

3. Detection Method

In this section, we will give the fault detection method including off-line and on-line cases as shown in Figure 3.

3.1. Off-Line Model Building

Algorithm 3.1.

Step 1. Set $i = 1$, and s is the number of simplified data set.

Step 2. Find k nearest neighbors (see [30]) for each sample i in the simplified data set and calculate the distances between sample i and its k -nearest neighbors.

Step 3. Estimate the cumulative density function of the above-squared distances by the function "ksdensity" in Matlab, which is denoted as $f_i(x)$.

Step 4. Calculate expectation of the KNN-squared distance based on the obtained cumulative density function in terms of the definition of expectation, and the expectation is denoted as \bar{X}_i .

Step 5. Set $i = i + 1$. If $i \leq s$, go to Step 2; otherwise, go to Step 6.

Step 6. Estimate the cumulative distribution function of the expectation to obtain the CL.

Remark 3.2. At Step 2, Euclidean distance is used, which is simple and easy, but any other distance is also suitable for the method proposed. The obtained CL in Algorithm 3.1 is based on the statistical test concept in the same way as SPC, in this sense off-line model is constructed by the KNN rule-based SPC approach.

Remark 3.3. In general, the estimation of probability density function in multidimensional space is difficultly derived [12]. To overcome this difficulty, we try to estimate the probability density functions of squared distance and expectation of squared distance in Step 3 and Step 4 from stochastic variable point of view. In addition, expectation of squared distance can be also obtained by taking an average over squared distances.

Remark 3.4. Since there are similar statistical characteristics for the normal samples and the distance between the fault sample and the nearest neighboring samples must be greater than the normal sample's distance to the nearest neighboring samples [9, 10, 12], setting CL to detect faults in terms of cumulative distribution function of the expectation is sound and effective. CL proposed in this paper means that the expectation values of vast majority distance for the normal samples do not exceed it. For example, 95% control limit means that the value within which 95% of population of normal operation data (expectation values) is included. Here, 95% is also called confidence level based on probability and statistical theory.

3.2. On-Line Fault Detection

Algorithm 3.5.

Step 1. Calculate the distance between the query and k nearest neighbors in the simplified data set.

Step 2. Estimate the cumulative density function of the above-squared distances.

Step 3. Calculate expectation based on the obtained cumulative density function.

Step 4. The query is abnormal if the expectation is beyond its CL, otherwise, this query is normal.

Step 5. If the query is normal, to update, it can be put into the normal samples database, which will be also simplified by using the technique described in Algorithm 2.1. In this case, the updated database is used to compute the new CL to continue to identify the next query in Step 1.

Remark 3.6. Compared with [9, 10, 12, 19], the technique that simplifies and updates raw data set is a main contribution in this paper. More importantly, the time-varying CL can be derived, such that adaptive fault detection can be implemented by on-line approach, which will eliminate the impact of data drift and shift on the quality of fault detection. Different from [18] wherein just-in-time-learning (JITL) along with two-step independent component analysis and principal component analysis was studied, in this paper, the database is updated and simplified. Moreover, note that the amount of database will not randomly increase when normal queries are added to it. The reason is that Algorithm 2.1 is implemented once the on-line detection process is completed. In other words, the updated database still can be simplified

Table 1: Control limits under the different thresholds.

Thresholds	Number of left samples	Maximum Mahalanobis distance	CL
0.99	292	$4.8747e - 005$	$8.1465e - 004$
0.95	271	$3.0298e - 004$	$8.7031e - 004$
0.6	124	0.0028	0.0026

by virtue of Algorithm 2.1. Obviously, high fault detection capability and low cost can be owned due to the usage of Algorithms 2.1–3.5.

Remark 3.7. In fact, the difficulties posed by nonlinearity, dynamics, and multiple modes of control process on fault detection have been addressed explicitly by the detection method proposed, which comes as no surprise, since the KNN technique, SPC method as well as on-line and update scheme are integrated.

4. Numerical Examples

In this section, two examples are given to show the effectiveness of the fault detection technique. The first example aims at the single modal case to show the efficacy of simplified data set based detection procedure presented in this paper. The second example is used in the multimodal case to compare with JIT method [12].

Example 4.1. Consider the following dominant nonlinear process mode [9, 10]:

$$x_1 = x_2^2 + \text{noise}. \quad (4.1)$$

Firstly, 30 normal runs are operated for verifying the method of simplifying raw data set. Here threshold $\epsilon = 0.95$ is set. By Algorithm 2.1, the 28 samples are left. Figures 4 and 5 show the raw data set and the simplified data set, respectively.

Continue to operate the system (4.1), we obtain 300 normal data used for the raw data, 5 normal runs used for validation, and 5 faults introduced, which is shown in Figure 6. The number of nearest neighbors k is set to be 10, and the confidence level is chosen as 99% to obtain the CL. Table 1 gives the number of left raw data set used for training, the maximum Mahalanobis distance, and the CL under the different thresholds ϵ , and the histogram of simplified raw data and fault detection is shown in Figure 7, where the percentages of left data to raw data and detected faults to total faults are clearly seen. Here, the maximum Mahalanobis distance means that the samples with smaller distance than it can be merged based on Algorithm 2.1. Correspondingly, the detection results under thresholds 0.99, 0.95, and 0.6 are also shown in Figures 8, 9, and 10, respectively. As illustrated in Figures 7–10, the amount of left data become gradually less and less and the CL is increasing with the decrease in threshold, consequently, the effect of fault detection becomes bad as expected.

Note that the threshold ϵ decided has a significant impact on the detection results, obviously, the bigger the threshold is, the more accurate detection operates. Simulation results presented illustrate that defection performance does not suffer degradation by virtue of the simplified data set. FD-KNN [9] is applied into this nonlinear case, and the detection result is shown in Figure 11. It should be pointed out though a better detection result is also obtained

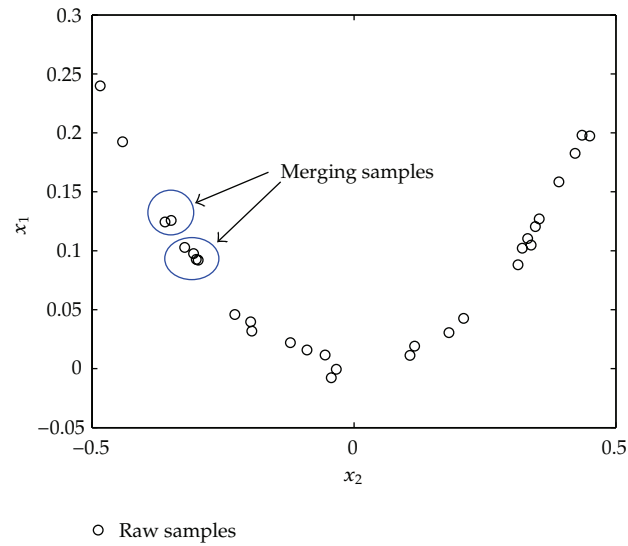


Figure 4: Raw data set.

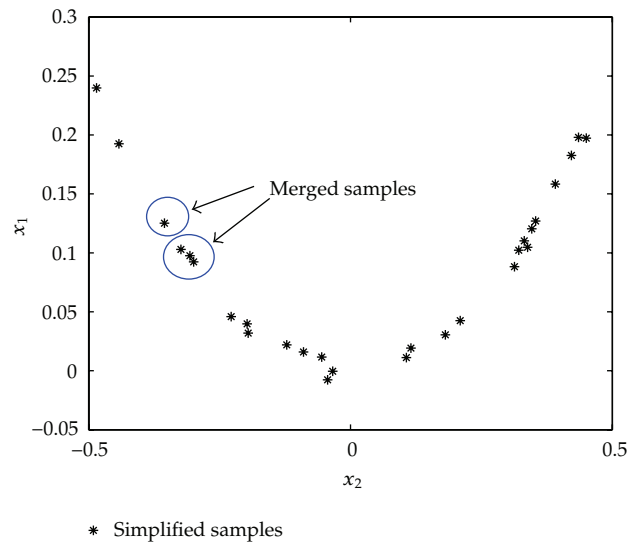


Figure 5: Simplified data set.

by using FD-KNN approach [9], the raw data set is simplified before implementing the detection in this paper, which will contribute to the saving storage space and reducing the computational complexity.

Example 4.2. Considering the following bimodal case [9, 10]:

$$\begin{aligned} A \quad y_1 &= 2y_2 + \text{noise}, \\ B \quad y_1 &= 1.5y_2 + 6 + \text{noise}. \end{aligned} \tag{4.2}$$

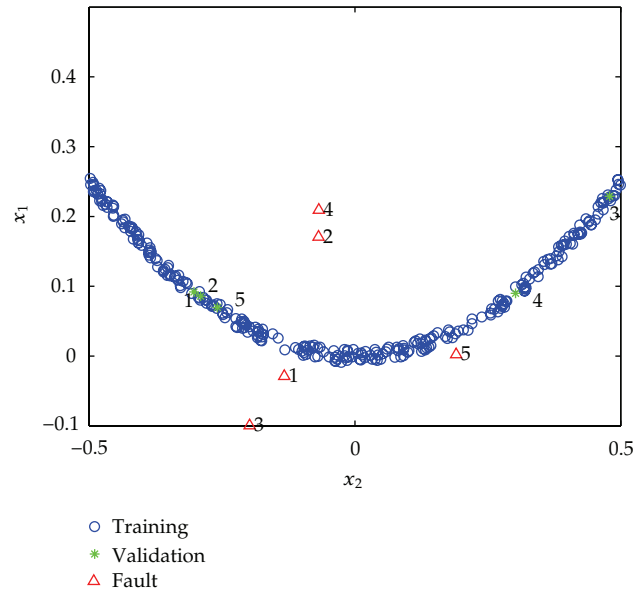


Figure 6: Scatter plot of samples.

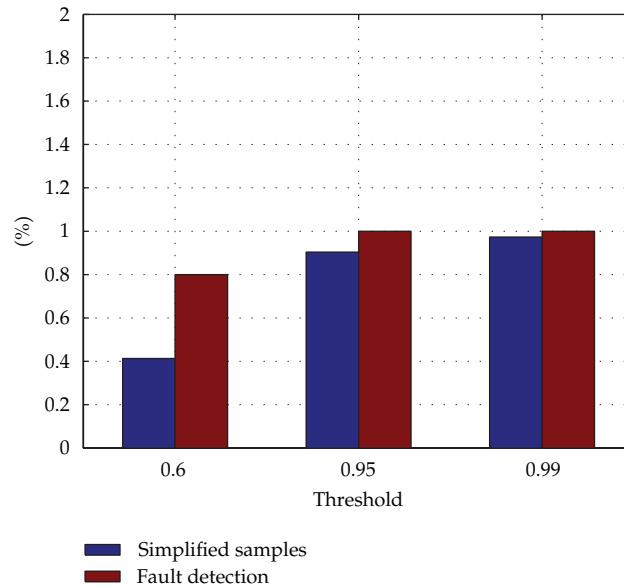


Figure 7: Histogram of detection results under different thresholds.

The above two cases are operated to produce 200 normal samples, respectively, and continue to be operated to produce 100 normal samples and 10 faults that are used for the validation and fault detection, respectively, which is given in Figure 12. For comparative analysis, 5 normal data and 10 fault data are marked. Similar to Example 4.1, the number of nearest neighbors k is set to be 10, and the confidence level is chosen as 99% to obtain the CL.

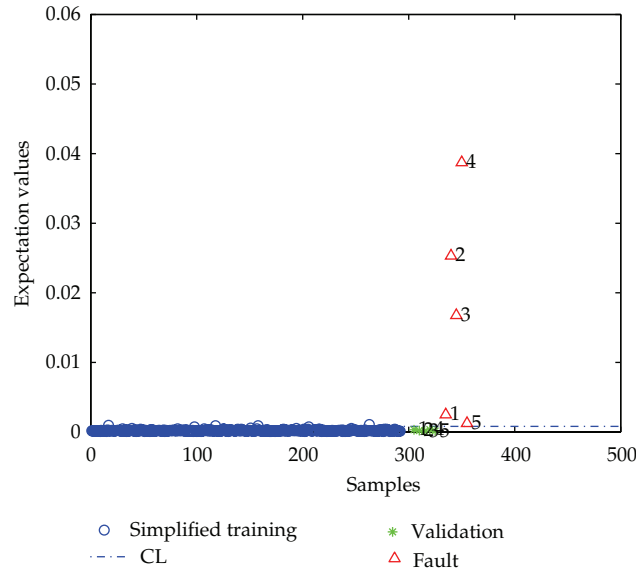


Figure 8: Detection result under the threshold $\epsilon = 0.99$.

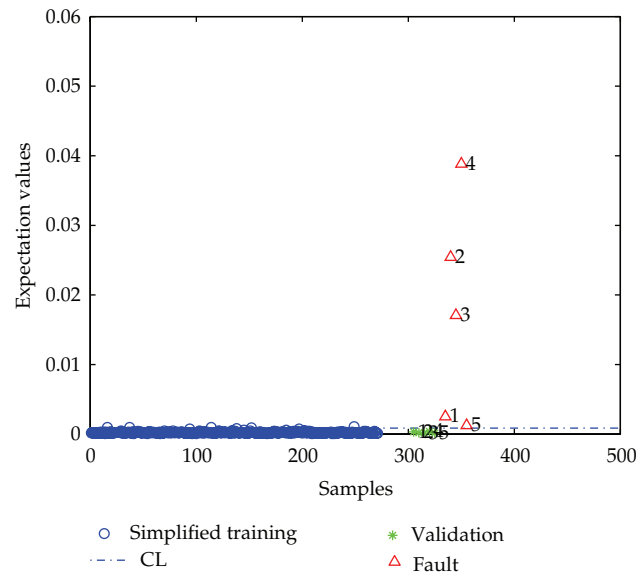
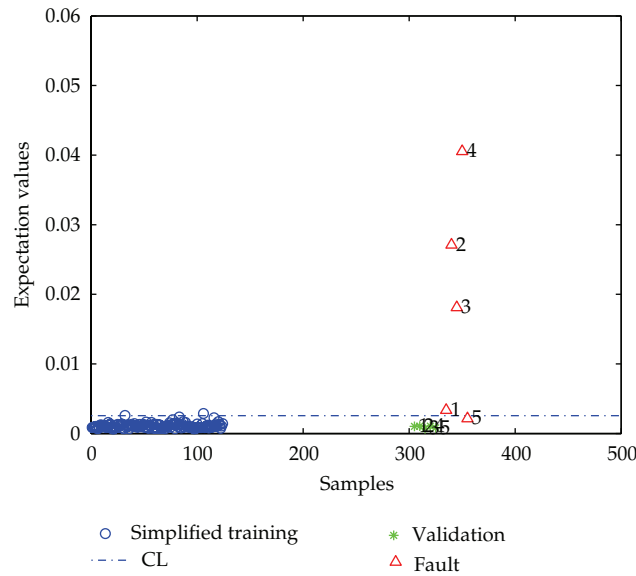


Figure 9: Detection result under the threshold $\epsilon = 0.95$.

By Algorithm 2.1, the raw data set is simplified. Moreover, the simplified data set is updated by on-line approach when the 100 normal data and 10 faults are detected. As shown in Table 2, at the end of the detection, the number of left raw data is 358, it is obvious that the amount of data set is not increased unlimitedly due to the threshold $\epsilon = 0.9$. There is no doubt that the on-line and update method proposed in this paper can surely reduce the cost of data storage and computation load. The detection results by the method in this paper and JIT method [12] are presented in Figures 13 and 14, respectively. Moreover, the embedded son

Table 2: Summary of parameters in Example 4.2.

Raw data	400
Threshold	0.99
Left data	358
Confidence level	99%
CL	0.0067
Normal data	100
Fault data	10
Updated data set	344

**Figure 10:** Detection result under the threshold $\epsilon = 0.6$.

figures in Figures 13 and 14 are used to emphasize the CL and the verification of training data and validation. From Figure 12, the CL is time-varying as the normal and fault data are identified. Note that since only normal data 33, 34, and 47 are mistaken for faults by the method proposed, whereas normal data 33, 34, 47, 6, and 62 are mistaken for faults by the JIT method [12], the better detection result is obtained in this paper using the original and update CL than the one in [12]. Here, the original CL means CL that is obtained in terms of simplified database by off-line approach. Comparatively speaking, the threshold determined subjectively in advance during calculating the sparse distance might partially degrade the performance of fault detection. Likewise, the tradeoff of storage cost and high detection performance is realized.

5. Case Study

In this section, an AL stack etch process was performed on a commercial scale Lam 9600 plasma etch tool at Texas Instrument, Inc. [9, 32, 33]. The data are taken from MACHINE_DATA, OES_DATA, and RFM_DATA during three experiments [33]. As pointed out by [9],

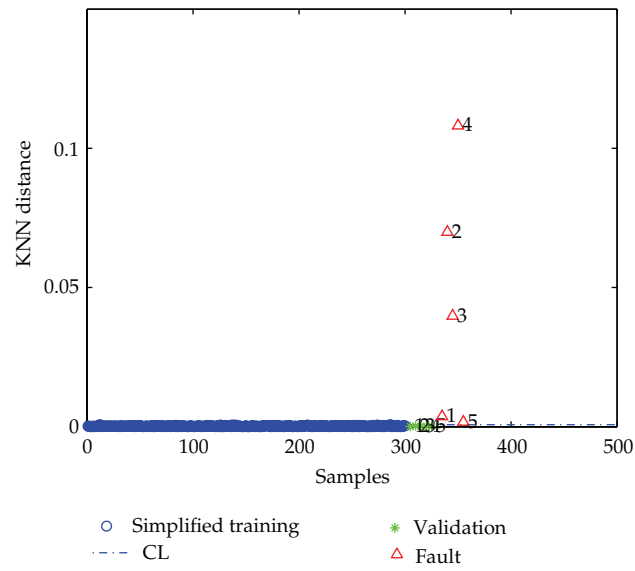


Figure 11: Fault detection by FD-KNN method.

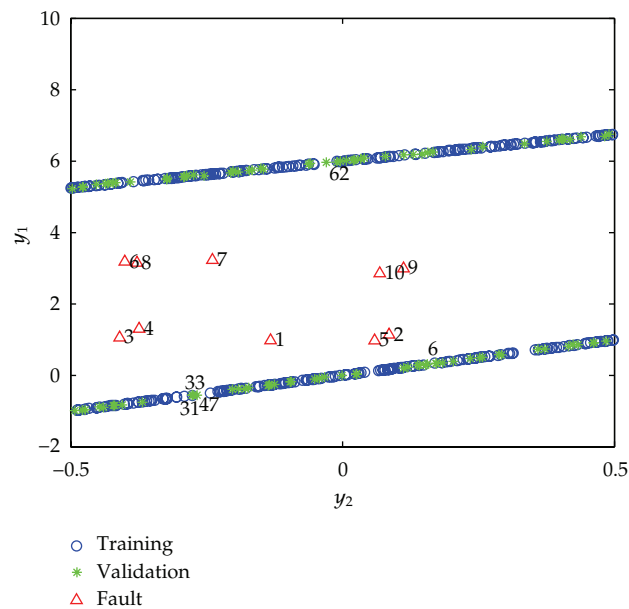


Figure 12: Scatter plot of samples.

the unique characteristics associated with semiconductor process different from other production processes include the unequal batch duration, unequal step duration, and process drift and shift, therefore, data from the different experiments of the same resource have different mean and different covariance structures, which can be seen more clearly in the case of

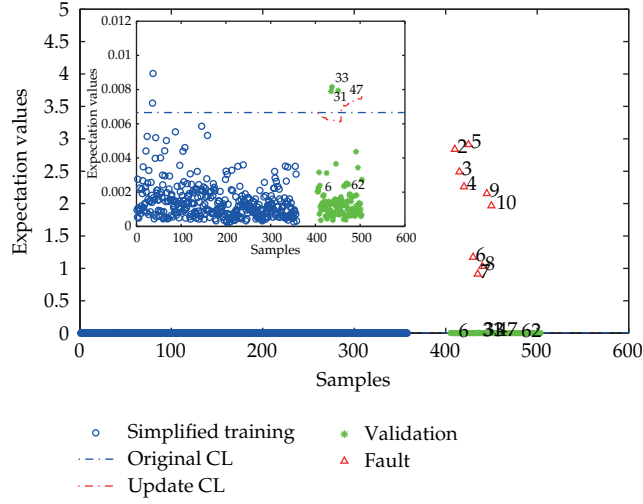


Figure 13: Fault detection by the method proposed in this paper.

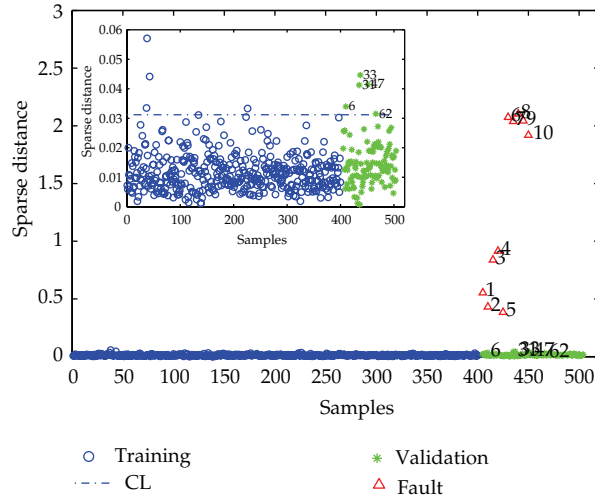


Figure 14: Fault detection by JIT method.

different resources. Due to the multimodal characteristic, the detection process based on the method proposed in this paper is as follows.

Step 1. Data preprocessing: we choose 10 normal batches and 10 process variables selected from the three data resources, such that the meaningful results can be obtained. Then, 30 normal batches are stored in the raw database. Further, the data is unfolded as the 2-D array by using the way in [9].

Step 2. The data set is simplified by Algorithm 2.1.

Step 3. The CL is computed by Algorithm 3.1.

Step 4. Another 15 normal batches are selected for the validation from the three different resources, in which 5 normal batches are used from each of resources. Moreover, 5 faults are also chosen from the intentionally induced during the experiments in the above

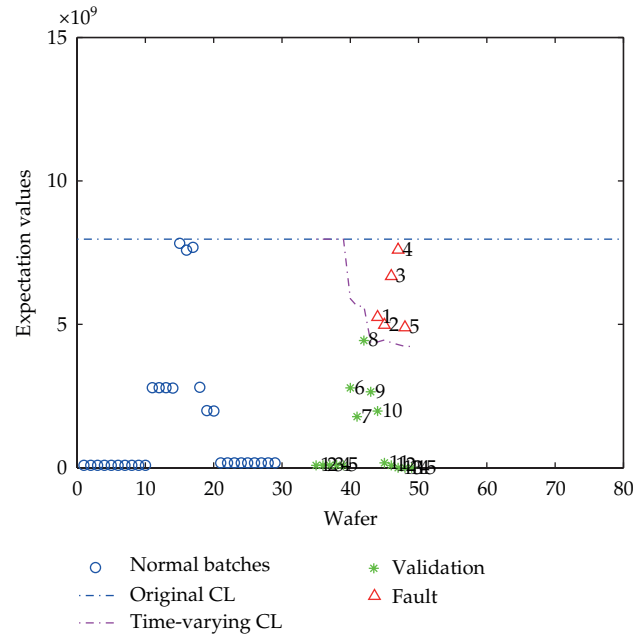


Figure 15: Fault detection for etch process.

resources for detection. Note that Algorithm 2.1 still can be used to undate the raw data set to produce the time-varying CL.

Following the aforementioned steps, threshold $\varepsilon = 0.9$ is determined in the detection for the etch process, and 29 normal batches can be obtained after Step 2. Moreover, the CL $7.9682e + 009$ is computed as original CL under confidence level 0.99. Known that if the query is normal, then it will be put into the simplified data set during identifying the normal data and fault data. It is worth pointing out the number of data in the database will be not surely increasing unlimitedly at the same time of updating the database due to Step 4. In this case, the final number of samples in the database is 36. The detection result is shown in Figure 15, where all of faults are identified correctly under the time-varying control limit by the method proposed in this paper. However, note that if raw data set is not updated, all of 5 faults can not be detected under the original CL computed. It is certain that some original faults with small data drift due to the effect of temperature and environment can be identified more easily in terms of the method of updating raw data set by online approach than the methods without applying the technology of update data set.

6. Conclusion

This research presented aims at highlighting the following two aspects: on one hand, the raw data set is simplified and updated by JIT approach based on the Mahalanobis distance between samples; on the other hand, combining the KNN rule with SPC method, the time-varying CL can be obtained to solve the nonlinear, multimodal, and data drift and shift problems existed in the practical case study. Numerical examples and an industrial case study show that the method proposed is an adaptive, flexible, and high-performance fault detection technique.

Acknowledgments

The authors would like to acknowledge the National Natural Science Foundation of China under Grant 61174119, 61104093, 61034006, 61174026, the Special Program for Key Basic Research Founded by MOST under Grant 2010CB334705, the National High Technology Research and Development Program of China (863 Program) under Grant 2011AA040101, and the Scientific Research Project of Liaoning Province of China under Grant L2012141, L2011064.

References

- [1] M. Wang and T. Zhan, "Analysis of the fault attack ECDLP over prime field," *Journal of Applied Mathematics*, vol. 2011, Article ID 580749, 11 pages, 2011.
- [2] Z. Gao, T. Breikin, and H. Wang, "Reliable observer-based control against sensor failures for systems with time delays in both state and input," *IEEE Transactions on Systems, Man, and Cybernetics A*, vol. 38, no. 5, pp. 1018–1029, 2008.
- [3] Z. Gao, X. Shi, and S. X. Ding, "Fuzzy state/disturbance observer design for T-S fuzzy systems with application to sensor fault estimation," *IEEE Transactions on Systems, Man, and Cybernetics B*, vol. 38, no. 3, pp. 875–880, 2008.
- [4] Z. Gao, X. Dai, T. Breikin, and H. Wang, "Novel parameter identification by using a high-gain observer with application to a gas turbine engine," *IEEE Transactions on Industrial Informatics*, vol. 4, no. 4, pp. 271–279, 2008.
- [5] X. B. Yang, X. Q. Jin, Z. M. Du, and Y. H. Zhu, "A novel model-based fault detection method for temperature sensor using fractal correlation dimension," *Building and Environment*, vol. 46, no. 4, pp. 970–979, 2011.
- [6] L. F. Mendonça, J. M. C. Sousa, and J. M. G. Sá da Costa, "An architecture for fault detection and isolation based on fuzzy methods," *Expert Systems with Applications*, vol. 36, no. 2, pp. 1092–1104, 2009.
- [7] P. M. Frank, E. Alcorta García, and B. Köppen-Seliger, "Modelling for fault detection and isolation versus modelling for control," *Mathematics and Computers in Simulation*, vol. 53, no. 4–6, pp. 259–271, 2000.
- [8] F. Li, B. R. Upadhyaya, and L. A. Coffey, "Model-based monitoring and fault diagnosis of fossil power plant process units using group method of data handling," *ISA Transactions*, vol. 48, no. 2, pp. 213–219, 2009.
- [9] Q. P. He and J. Wang, "Fault detection using the k-nearest neighbor rule for semiconductor manufacturing processes," *IEEE Transactions on Semiconductor Manufacturing*, vol. 20, no. 4, pp. 345–354, 2007.
- [10] Q. P. He and J. Wang, "Principal component based k-nearest-neighbor rule for semiconductor process fault detection," in *Proceedings of the American Control Conference (ACC '08)*, pp. 1606–1611, Washington, USA, June 2008.
- [11] J. Park, I. H. Kwon, S. S. Kim, and J. G. Baek, "Spline regression based feature extraction for semiconductor process fault detection using support vector machine," *Expert Systems with Applications*, vol. 38, no. 5, pp. 5711–5718, 2011.
- [12] K. Manabu, S. Takeaki, and H. Shinji, "Just-in-time statistical process control for flexible fault management," in *Proceedings of the SICE Annual Conference*, pp. 1482–1485, Taipei, Taiwan, August 2010.
- [13] Z. Gao and H. Wang, "Descriptor observer approaches for multivariable systems with measurement noises and application in fault detection and diagnosis," *Systems and Control Letters*, vol. 55, no. 4, pp. 304–313, 2006.
- [14] Z. Gao, S. X. Ding, and Y. Ma, "Robust fault estimation approach and its application in vehicle lateral dynamic systems," *Optimal Control Applications and Methods*, vol. 28, no. 3, pp. 143–156, 2007.
- [15] C. C. Feng, "Robust fault-tolerant control for systems with extended bounded-sensor-faults," in *Proceedings of the 17th World Congress of the International Federation of Automatic Control*, pp. 2582–2587, Seoul, Korea, 2008.
- [16] D. M. Himmelblau, *Fault Detection and Diagnosis in Chemical and Petrochemical Processes*, Elsevier, Amsterdam, The Netherlands, 1978.

- [17] S. Simani and C. Fantuzzi, "Dynamic system identification and model-based fault diagnosis of an industrial gas turbine prototype," *Mechatronics*, vol. 16, no. 6, pp. 341–363, 2006.
- [18] Z. Ge and Z. Song, "Online monitoring of nonlinear multiple mode processes based on adaptive local model approach," *Control Engineering Practice*, vol. 16, no. 12, pp. 1427–1437, 2008.
- [19] C. Cheng and M. S. Chiu, "Nonlinear process monitoring using JITL-PCA," *Chemometrics and Intelligent Laboratory Systems*, vol. 76, no. 1, pp. 1–13, 2005.
- [20] S. Yoon and J. F. MacGregor, "Statistical and causal model-based approaches to fault detection and isolation," *AIChE Journal*, vol. 46, no. 9, pp. 1813–1824, 2000.
- [21] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri, "A review of process fault detection and diagnosis: part I: quantitative model-based methods," *Computers and Chemical Engineering*, vol. 27, no. 3, pp. 293–311, 2003.
- [22] S. Mahadevan and S. L. Shah, "Fault detection and diagnosis in process data using one-class support vector machines," *Journal of Process Control*, vol. 19, no. 10, pp. 1627–1639, 2009.
- [23] D. Fischer, *Artificial Intelligence Techniques Applied to Fault Detection Systems [Ph.D. Dissertation]*, Electrical and Computer Engineering, McMaster University, Ontario, Canada, 2004.
- [24] D. Dong and T. J. Mcavoy, "Nonlinear principal component analysis—based on principal curves and neural networks," *Computers and Chemical Engineering*, vol. 20, no. 1, pp. 65–78, 1996.
- [25] H. G. Hiden, M. J. Willis, M. T. Tham, and G. A. Montague, "Non-linear principal components analysis using genetic programming," *Computers and Chemical Engineering*, vol. 23, no. 3, pp. 413–425, 1999.
- [26] W. Ku, R. H. Storer, and C. Georgakis, "Disturbance detection and isolation by dynamic principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 30, no. 1, pp. 179–196, 1995.
- [27] G. Bontempi, H. Bersini, and M. Birattari, "The local paradigm for modeling and control: from neuro-fuzzy to lazy learning," *Fuzzy Sets and Systems*, vol. 121, no. 1, pp. 59–72, 2001.
- [28] R. D. Nowak, "Nonlinear system identification," *Circuits, Systems, and Signal Processing*, vol. 21, no. 1, pp. 109–122, 2002.
- [29] X. Gao, G. Wang, and J. H. Ma, "An approach of building simplified multivariate statistical model based on mahalanobis distance," *Information and Control*, vol. 30, no. 7, pp. 676–680, 2001.
- [30] A. Ison and C. J. Spanos, "Robust fault detection and fault classification of semiconductor manufacturing equipment," in *Proceedings of the Fifth International Symposium on Semiconductor Manufacturing*, pp. 1–4, Tokyo, Japan, October 1996.
- [31] W. Li, H. H. Yue, S. Valle-Cervantes, and S. J. Qin, "Recursive PCA for adaptive process monitoring," *Journal of Process Control*, vol. 10, no. 5, pp. 471–486, 2000.
- [32] B. M. Wise, N. B. Gallagher, S. W. Butler, D. D. White, and G. G. Barna, "A comparison of principal component analysis, multiway principal component analysis, trilinear decomposition and parallel factor analysis for fault detection in a semiconductor etch process," *Journal of Chemometrics*, vol. 13, no. 3–4, pp. 379–396, 1999.
- [33] E. R. Inc, "Metal etch data for fault detection evaluation," 1999, <http://software.eigenvector.com/Data/Etch/index.html>.

Research Article

Computation of the Added Masses of an Unconventional Airship

**Naoufel Azouz,¹ Said Chaabani,¹
Jean Lerbet,¹ and Azgal Abichou²**

¹ Laboratoire IBISC, Université d'Evry Val d'Essonne, 40 Rue du Pelvoux, 91025 Evry, France

² Lab of Mathematical Engineering, Polytechnic School, 2078 La Marsa, Tunisia

Correspondence should be addressed to Naoufel Azouz, azouz@iup.univ-evry.fr

Received 5 June 2012; Revised 9 August 2012; Accepted 10 August 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 Naoufel Azouz et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a modelling of an unmanned airship. We are studying a quadrotor flying wing. The modelling of this airship includes an aerodynamic study. A special focus is done on the computation of the added masses. Considering that the velocity potential of the air surrounding the airship obeys the Laplace's equation, the added masses matrix will be determined by means of the velocity potential flow theory. Typically, when the shape of the careen is quite different from that of an ellipsoid, designers in preprocessing prefer to avoid complications arising from mathematical analysis of the velocity potential. They use either complete numerical studies, or geometric approximation methods, although these methods can give relatively large differences compared to experimental measurements performed on the airship at the time of its completion. We tried to develop here as far as possible the mathematical analysis of the velocity potential flow of this unconventional shape using certain assumptions. The shape of the careen is assumed to be an elliptic cone. To retrieve the velocity potential shapes, we use the sphericoconal coordinates. This leads to the Lamé's equations. The whole system of equations governing the interaction air-structure, including the boundary conditions, is solved in an analytical setting.

1. Introduction

The rapid expansion of the capabilities of airships in the last decade and the growing of the range of missions they designed to support lead to the design of complex shapes of the careens.

Exhaustive studies in that topic were presented by [1, 2]. Traditionally, ellipsoidal shapes are used for airships [3–5]. However, and in order to optimize their performances, different original shapes have been tested in the last years. This is due to the advances in aerodynamics, control theory, and embedded electronics. The airship studied here departs



Figure 1: The flying wing Airship MC500.

with the traditional shapes. The MC500 is a flying wing (Figure 1), developed by the French network DIRISOFT.

The MC500 is an experimental prototype for a set of great innovating airships. A precise dynamics model is needed for this kind of unmanned airships including the air-structure interaction. This will enable the elaboration of convenient algorithms of control, stabilization, or navigation of these flying objects.

The aerodynamic forces have a large influence on the dynamic behaviour of the airships. Among the aerodynamic solicitations, the added masses phenomenon is one of the most important. In fact, when hovering or in case of low speed displacement, the lift and drag forces and torques could be neglected.

The added masses phenomenon is well known for airships and similarly for submarines. When an airship moves in an incompressible and infinite inviscid fluid, the kinetic energy of the fluid produces an effect equivalent to an important increase of the mass and the inertia moments of the body. This contribution may be of the same magnitude as the terms of mass or inertia of the airship. Apart from ellipsoidal shapes [6] or thin rectangular plates [7] where the theory is well established for many decades, the analysis of this problem for unconventional shapes in preprocessing is usually done by approximate methods. We can quote the geometric method, consisting of an extension of a 2D analysis and requiring the introduction of empirical parameters [8, 9], or the fully numerical methods consisting of modeling both the airship and the surrounding air [10, 11].

As opposed to other treatments of this problem, the derivation proposed here is based on the velocity potential flow theory [6, 7, 12]. We tried to pursue the analytical study to an advanced stage with some assumptions. The shape of the careen of the MC500 is considered as a cone with elliptic section. When we consider that the velocity potential of the air obeys the Laplace's equation, the added masses matrix could be determined by solving this equation, using the spheroconal coordinates. Solving the Laplace's equation for such configuration leads to the Lamé's equations. We have dedicated a section of this paper for the determination of the ellipsoidal harmonic series issued from these equations with the specific boundary conditions. Such series developed for the first time by Lamé [13] were improved particularly by Liouville and Sturm in their famous theory and by Hermite [14] and were applied in different fields. Byerly [15] and later Hobson [16] gave an important comprehensive literature about this topic.

Significant recent works, such as the works of Boersma and Jansen [17] in electromagnetic field, and those of Garmier and Barriot [18] in astrophysics could be quoted. In this work, we are trying to apply this theory for the air-structure interaction.

2. Dynamic Model

2.1. Kinematics

The airship MC500 is assumed to be a rigid flying object.

We use two reference frames. First an earth-fixed frame $R_0 = (O, X_0, Y_0, Z_0)$ assumed to be Galilean. Then a local reference frame $R_m = (G, X_m, Y_m, Z_m)$ fixed to the airship. Its axes are selected as follows:

X_m and Y_m are the transverse axis of the airship, Z_m : the normal axis directed downwards.

To describe the orientation of the airship, we use a parameterization in yaw, pitch, and roll. The configuration of the object is described by means of three rotations defined by three angles of orientation, that is, the yaw ψ , pitch θ , and roll ϕ .

The whole transformation between the local reference frame R_m and the global frame is given by Goldstein [19]:

$$J_1^T(\eta_2) = \begin{pmatrix} c\theta \cdot c\psi & c\theta \cdot s\psi & -s\theta \\ s\phi \cdot s\theta \cdot c\psi - c\phi \cdot s\psi & s\phi \cdot s\theta \cdot s\psi + c\phi \cdot c\psi & s\phi \cdot c\theta \\ c\phi \cdot s\theta \cdot c\psi + s\phi \cdot s\psi & c\phi \cdot s\theta \cdot s\psi - s\phi \cdot c\psi & c\phi \cdot c\theta \end{pmatrix}, \quad (2.1)$$

such as $J_1^T(\eta_2) \cdot J_1(\eta_2) = J_1(\eta_2) \cdot J_1^T(\eta_2) = I_3$. We denote by $c\theta = \cos \theta$; $s\phi = \sin \phi$.

Using the rotation matrix $J_1(\eta_2)$, the expression of the linear speed in the reference frame R_0 is given by

$$\dot{\eta}_1 = J_1(\eta_2) \cdot v_1. \quad (2.2)$$

The angular speed is expressed as follows:

$$\dot{\eta}_2 = J_2(\eta_2) \cdot v_2 \quad (2.3)$$

with

$$J_2(\eta_2) = \begin{pmatrix} 1 & s\phi \tan \theta & c\phi \tan \theta \\ 0 & c\phi & -s\phi \\ 0 & \frac{s\phi}{c\theta} & \frac{c\phi}{c\theta} \end{pmatrix}. \quad (2.4)$$

It is noticed that the parameterization by the Euler angles has a singularity in $\Theta = \pi/2 + k\pi$.

This parameterization is acceptable because it is not possible for an airship to reach this singular orientation of 90 degrees pitching angle.

2.2. Dynamics

As mentioned previously, a specificity of the lighter than air vehicles is illustrated by the added masses phenomenon. When a big and light object moves in the air, the kinetic energy of the particles of air produces an effect equivalent to an important growing of the mass and inertia of the body. As the airship displays a very large volume, its added masses and inertias become significant.

The basis of the analysis of the motion of a rigid body immersed in a perfect fluid is established by Lamb [6]. He proves that the kinetic energy T of the body with the surrounding fluid can be expressed as a quadratic function of the six components of the translation and rotation velocity as follows:

$$T = \frac{1}{2} \mathbf{v}^T \underbrace{(M_b + M_a)}_M \mathbf{v} = T_b + T_a, \quad (2.5)$$

where M_a is the added mass matrix due to the motion of the surrounding air and M_b is the mass matrix of the body.

For an airship fully immersed in the air, the 6×6 added mass matrix M_a is assumed to be positive and definite. As the added kinetic energy T_a , it will be discussed in the next section. The whole mass matrix M is assumed symmetric block-diagonal matrix:

$$M = \begin{pmatrix} M_{TT} & 0 \\ 0 & M_{RR} \end{pmatrix}. \quad (2.6)$$

For the computation of the whole dynamics model, we choose to use the Kirchoff's equation [20]:

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial T}{\partial \mathbf{v}_1} \right) + \mathbf{v}_2 \wedge \frac{\partial T}{\partial \mathbf{v}_1} &= \boldsymbol{\tau}_1, \\ \frac{d}{dt} \left(\frac{\partial T}{\partial \mathbf{v}_2} \right) + \mathbf{v}_2 \wedge \frac{\partial T}{\partial \mathbf{v}_2} + \mathbf{v}_1 \wedge \frac{\partial T}{\partial \mathbf{v}_1} &= \boldsymbol{\tau}_2, \end{aligned} \quad (2.7)$$

where \wedge is the vectorial product, $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$ are, respectively, the external forces and torques, including the rotors effects, the weight (mg), the buoyancy B , and the aerodynamic lift (F_L) and drag (F_D).

The dynamical system of the airship becomes [21–23]

$$\begin{pmatrix} M_{TT} & 0 \\ 0 & M_{RR} \end{pmatrix} \begin{pmatrix} \dot{\mathbf{v}}_1 \\ \dot{\mathbf{v}}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\tau}_1 - \mathbf{v}_2 \wedge (M_{TT} \mathbf{v}_1) \\ \boldsymbol{\tau}_2 - \mathbf{v}_2 \wedge (M_{RR} \mathbf{v}_2) - \mathbf{v}_1 \wedge (M_{TT} \mathbf{v}_1) \end{pmatrix}. \quad (2.8)$$

Description of the Rotors

The MC500 has four electric engines driving rotors. Each rotor has two parallel contrarotating propellers to avoid any aerodynamic torque (Figure 2). The rotor can swivel in two directions. A rotation of angle β_i around the Y axis ($-180^\circ \leq \beta_i \leq 180^\circ$) and a rotation of angle γ_i around

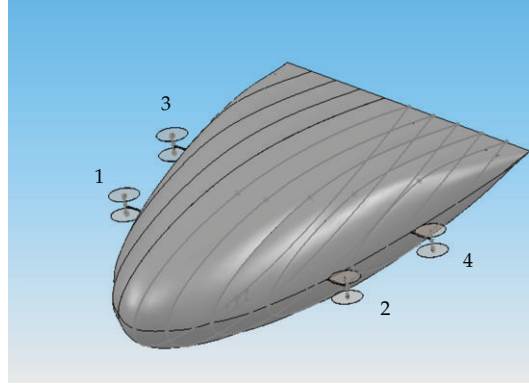


Figure 2: Position of the rotors.

an axis Z_{iR} normal to Y and initially coinciding with the Z axis ($-30^\circ \leq \gamma_i \leq 30^\circ$). A fictive axis X_{iR} completes the rotor frame.

Let us denote P_i by the position of the rotor i . We can then define a rotation matrix J_3^i between the frame (P_i, X_{iR}, Y, Z_{iR}) and the local frame R_m such as

$$J_3^i = \begin{pmatrix} c\gamma_i c\beta_i & -s\gamma_i c\beta_i & s\beta_i \\ s\gamma_i & c\gamma_i & 0 \\ -c\gamma_i s\beta_i & s\gamma_i s\beta_i & c\beta_i \end{pmatrix}. \quad (2.9)$$

If we suppose that the intensity of the thrust force of the rotor i is $\|\mathbf{F}_i\|$, this force could be introduced in the second member of the dynamic equation as

$$\mathbf{F}_i = J_3^i \|\mathbf{F}_i\| \cdot e_{X_m}, \quad (2.10)$$

where e_{X_m} is a unitary vector along the X_m axis.

The torque produced by this rotor in the centre of inertia G is $\mathbf{F}_i \wedge P_i G$.

Weight and Buoyancy

An important characteristic of the airships is the buoyancy B_u . This force represents a natural static lift, corresponding roughly to 1 Kg for each m^3 of helium involved in the careen. We suppose here that this force is applied in the centre of buoyancy B different from the centre of inertia G :

$$B_u = \rho_{\text{air}} \cdot V \cdot g, \quad (2.11)$$

where V is the volume of the careen, ρ_{air} is the density of the air, and g the gravity.

Let us note F_{WB} and M_{WB} the force and the moment due to the weight and buoyancy. We have

$$F_{WB} = (mg - B_u) \cdot J_1^T e_Z, \quad M_{WB} = B_u \cdot \left(J_1^T \cdot e_Z \wedge BG \right). \quad (2.12)$$

Aerodynamic Forces F_A

Such as other flying objects, the airships are subjected to aerodynamic forces. The resultant of these forces could be divided into two component forces, one parallel to the direction of the relative wind and opposite to the motion, called Drag, and the other perpendicular to the relative wind, called Lift. The MC500 is designed with an original shape oriented to a best optimization of the ratio lift upon drag forces.

However, and as first study, we try to evaluate the behaviour of the airship in the case of low velocity or while hovering. In these cases, the effect of these forces could be neglected.

The global dynamic system could be expressed in a compact form as follows [21, 24]:

$$M \cdot \dot{v} = \tau + Q_G \quad (2.13)$$

with

$$\tau = \begin{pmatrix} \tau_1 \\ \tau_2 \end{pmatrix}, \quad (2.14)$$

such as

$$\begin{aligned} \tau_1 &= \begin{pmatrix} \sum_{i=1}^4 \|F_i\| \cdot c\gamma_i \cdot c\beta_i - (mg - B_u) \cdot s\theta \\ \sum_{i=1}^4 \|F_i\| s\gamma_i + (mg - B_u) \cdot s\phi \cdot c\theta \\ - \sum_{i=1}^4 \|F_i\| c\gamma_i \cdot s\beta_i + (mg - B_u) c\phi \cdot c\theta \end{pmatrix}, \\ \tau_2 &= - \begin{pmatrix} c \sum_{i=1}^4 \|F_i\| \cdot s\gamma_i + b_1 (\|F_1\| c\gamma_1 \cdot s\beta_1 - \|F_2\| c\gamma_2 \cdot s\beta_2) \\ + b_3 (\|F_3\| c\gamma_3 \cdot s\beta_3 - \|F_4\| c\gamma_4 \cdot s\beta_4) + B_u z_B \cdot s\phi \cdot c\theta \\ - c \sum_{i=1}^4 \|F_i\| c\gamma_i \cdot c\beta_i \\ + a (\|F_4\| c\gamma_4 \cdot s\beta_4 + \|F_3\| c\gamma_3 \cdot s\beta_3 - \|F_1\| c\gamma_1 \cdot s\beta_1 - \|F_2\| c\gamma_2 \cdot s\beta_2) + B_u z_B \cdot s\theta \\ b_1 (\|F_1\| c\gamma_1 \cdot c\beta_1 - \|F_2\| c\gamma_2 \cdot c\beta_2) + b_3 (\|F_3\| c\gamma_3 \cdot c\beta_3 - \|F_4\| c\gamma_4 \cdot c\beta_4) \\ + a (\|F_4\| s\gamma_4 + \|F_3\| s\gamma_3 - \|F_1\| s\gamma_1 - \|F_2\| s\gamma_2) \end{pmatrix}, \\ Q_G &= \begin{pmatrix} -v_2 \wedge (M_{TT} v_1) \\ -v_2 \wedge (M_{RR} v_2) - v_1 \wedge (M_{TT} v_1) \end{pmatrix}. \end{aligned} \quad (2.15)$$

3. Computation of the Added Masses

3.1. Spheroconal Coordinates

Much of the current airships developed and studied in the literature are considered as ellipsoidal. This particular shape has a wide popularity in this field due to the existence of

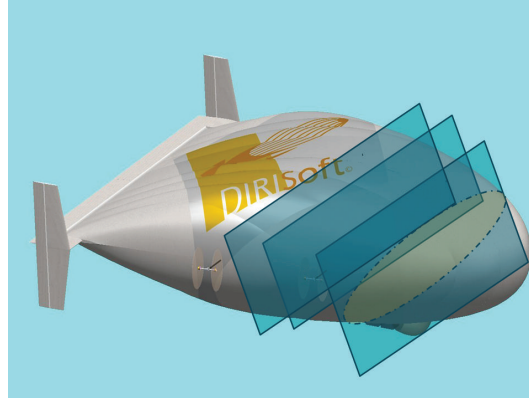


Figure 3: Transverse sections of the MC500.

a large and reliable knowledge and experimentations for both ellipsoidal airships and their alter ego submarines. However for our flying wing airship, we should use a more convenient approximation for the aerodynamic study.

The MC500 airship is a collection of aerodynamic profiles, with symmetry around the x - z axis. Each transverse section of the careen parallel to the plan y - z gives roughly an ellipse. This motivated us to model as first assumption the airship as an elliptic cone (Figure 3).

To take into account the interaction of the airship with the surrounding fluid, a model of the flow is needed.

A variation of this study may be performed by calculating the fluid pressure around the airship through the Bernoulli equation [25].

Here, we use the potential flow theory with the following assumptions.

- (a) The air can be considered as an ideal fluid with irrotational flow and uniform density ρ_f , that is, an incompressible fluid.
- (b) A velocity potential Φ_f exists and satisfies the Laplace's equation throughout the fluid domain:

$$\nabla^2 \Phi_f = 0, \quad (3.1)$$

and satisfies the nonlinear free surface condition, body boundary condition, and initial conditions.

Finally we suppose that the velocity of the air is null far from the airship:

$$\Phi_{f\infty} \longrightarrow 0. \quad (3.2)$$

- (c) The velocity of the fluid obeys the equation:

$$V_f = \nabla \Phi_f. \quad (3.3)$$

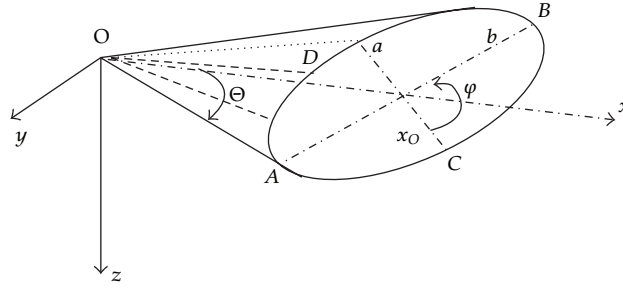


Figure 4: Representation of the spheroconal coordinates in the elliptic cone.

In his study [6], Lamb proves that the kinetic energy T_a of the fluid surrounding a moving rigid body can be expressed as a quadratic function of the six components of the translation and rotation velocity $\mathbf{v} = (u, v, w, p, q, r)^T$ and can also be expressed as function of the velocity potential of this fluid by the following relation:

$$T_a = \frac{1}{2} \mathbf{v}^T M_a \mathbf{v} = -\frac{1}{2} \rho_f \iint_{\partial S} \Phi_f \frac{\partial \Phi_f}{\partial n_O} dS. \quad (3.4)$$

n_O is an outward normal vector. Φ_f can be splitted on

$$\Phi_f = \mathbf{v}^T \Phi = u\Phi_1 + v\Phi_2 + w\Phi_3 + p\Phi_4 + q\Phi_5 + r\Phi_6, \quad (3.5)$$

where Φ appears as a collection of six spatial potential Φ_i functions of x, y, z that verify the Laplace's equation. To extract the terms of the added mass matrix, we should define the spatial potential Φ of the moving fluid. In some cases it could be determined using geometric means [8, 26]. Here we choose to use the potential flow theory. We can then express the components of the added mass matrix M_a in function of the velocity potential flow of the air surrounding the airship Φ such as

$$M_{aij} = -\rho_f \iint_S \Phi_i \frac{\partial \Phi_j}{\partial n_O} dS. \quad (3.6)$$

For solving the Laplace's equation (3.1) and according to the configuration of the careen, we use the spheroconal coordinates [16]. This assumption seems acceptable in that stage.

A first parameterisation by the coordinates (ρ, Θ, φ) is used, for example, by Boersma and Jansen [17], Kraus and Levine [27] (Figure 4), where ρ is the distance of a point to the origin, φ is the azimuthal angle ($0 \leq \varphi \leq 2\pi$), and Θ is the longitudinal angle. $\Theta = \Theta_0$ represents the external surface of the elliptic cone.

This parameterisation has the advantage to be physically significant. However, in our case, it leads to intractable calculations. We prefer to use the parameterisation (ρ, μ, ζ) given

by Hobson [16]. The equation of the surface of an elliptic cone is given by G. A. Korn and T. M. Korn [28] such as

$$\frac{x^2}{\mu^2} + \frac{y^2}{\mu^2 - a^2} + \frac{z^2}{\mu^2 - b^2} = 0. \quad (3.7)$$

a , b , and μ_0 are parameters that define the elliptic cone.

With

$$b^2 > \zeta^2 > a^2 > \mu^2 > \mu_0^2, \quad (3.8)$$

ρ is always the distance of a point to the origin. We define then the Cartesian coordinates (x, y, z) as

$$\begin{aligned} x &= \frac{\rho\mu\zeta}{ab}, \\ y &= \frac{\rho}{a} \sqrt{\frac{(\mu^2 - a^2)(\zeta^2 - a^2)}{a^2 - b^2}}, \\ z &= \frac{\rho}{b} \sqrt{\frac{(\mu^2 - b^2)(\zeta^2 - b^2)}{b^2 - a^2}}. \end{aligned} \quad (3.9)$$

By analogy with the azimuthal angle in the previous parameterisation, we can introduce an angle φ defined by

$$\begin{aligned} \cos^2 \varphi &= \frac{b^2 - \zeta^2}{b^2 - a^2}, \\ \sin^2 \varphi &= \frac{\zeta^2 - a^2}{b^2 - a^2}, \end{aligned} \quad \text{such that } \zeta = \sqrt{a^2 \cos^2 \varphi + b^2 \sin^2 \varphi}, \quad (3.10)$$

with $(0 \leq \varphi \leq 2\pi)$.

The surface of the elliptic cone is now defined by $\mu = \mu_0$. By varying this parameter, one can see its influence on the shape of the cone in Figure 5.

To express the components of the fluid velocity V_f in the conical elliptic reference frame, we define three unitary vectors $\vec{e}_\rho, \vec{e}_\varphi, \vec{e}_\mu$. With

$$\vec{e}_\rho = \frac{1}{h_\rho} \begin{bmatrix} \frac{\partial x}{\partial \rho} \\ \frac{\partial y}{\partial \rho} \\ \frac{\partial z}{\partial \rho} \end{bmatrix}, \quad \vec{e}_\varphi = \frac{1}{h_\varphi} \begin{bmatrix} \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial \varphi} \end{bmatrix}, \quad \vec{e}_\mu = \frac{1}{h_\mu} \begin{bmatrix} \frac{\partial x}{\partial \mu} \\ \frac{\partial y}{\partial \mu} \\ \frac{\partial z}{\partial \mu} \end{bmatrix}, \quad (3.11)$$

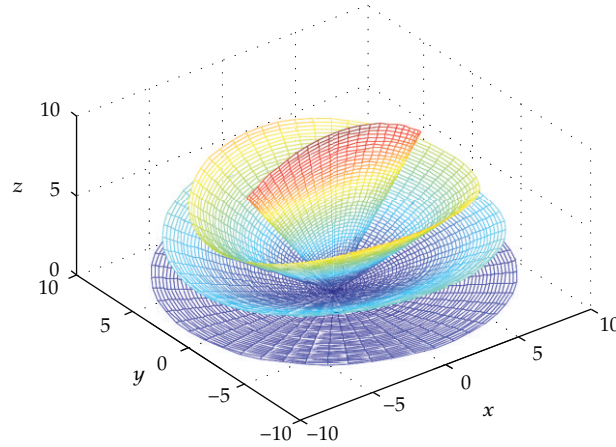


Figure 5: Influence of the parameter μ on the shape of the cone.

h_i are the scale factors. One can easily verify that these three angles form an orthonormal basis, and the velocity of the fluid could be expressed as

$$\vec{V}_f = V_\rho \cdot \vec{e}_\rho + V_\varphi \cdot \vec{e}_\varphi + V_\mu \cdot \vec{e}_\mu. \quad (3.12)$$

The surface of the elliptic cone is defined by $\mu = \mu_0$; \vec{e}_μ appears as the normal vector to the surface in a given point.

3.2. Laplace's Equation

With the spheroconal coordinates, the Laplace's equation (3.1) becomes [16]

$$(\mu^2 - \zeta^2) \frac{\partial^2 \Phi}{\partial \alpha^2} + (\rho^2 - \zeta^2) \frac{\partial^2 \Phi}{\partial \beta^2} + (\rho^2 - \mu^2) \frac{\partial^2 \Phi}{\partial \gamma^2} = 0 \quad (3.13)$$

with

$$\begin{aligned} \alpha &= b \int_b^\rho \frac{dt}{\sqrt{(t^2 - a^2)(t^2 - b^2)}}, & \beta &= b \int_a^\mu \frac{dt}{\sqrt{(b^2 - t^2)(t^2 - a^2)}}, \\ \gamma &= b \int_0^\zeta \frac{dt}{\sqrt{(a^2 - t^2)(b^2 - t^2)}}. \end{aligned} \quad (3.14)$$

To solve the Laplace's equation (3.13), we use the traditional separation of variables:

$$\Phi(\rho, \mu, \zeta) = R(\rho) \cdot Z(\mu) \cdot Y(\zeta). \quad (3.15)$$

Replacing (3.16) in (3.12) gives

$$\frac{\mu^2 - \zeta^2}{R} \frac{d^2 R}{d\alpha^2} + \frac{\rho^2 - \zeta^2}{Z} \frac{d^2 Z}{d\beta^2} + \frac{\rho^2 - \mu^2}{Y} \frac{d^2 Y}{d\gamma^2} = 0. \quad (3.16)$$

This equation can be splitted on three separated differential equations:

$$\begin{aligned} \frac{d^2 R}{d\alpha^2} - \left[n(n+1)\rho^2 - (a^2 + b^2)s \right] R &= 0, \\ \frac{d^2 Z}{d\beta^2} + \left[n(n+1)\mu^2 - (a^2 + b^2)s \right] Z &= 0, \\ \frac{d^2 Y}{d\gamma^2} - \left[n(n+1)\zeta^2 - (a^2 + b^2)s \right] Y &= 0. \end{aligned} \quad (3.17)$$

And by replacing α , β , and γ by their expressions in (3.14), one can obtain

$$\begin{aligned} (\rho^2 - a^2)(\rho^2 - b^2) \frac{d^2 R}{d\rho^2} + r(2\rho^2 - a^2 - b^2) \frac{dR}{d\rho} - \left[n(n+1)\rho^2 - (a^2 + b^2)s \right] R &= 0, \\ (\mu^2 - a^2)(\mu^2 - b^2) \frac{d^2 Z}{d\mu^2} + \mu(2\mu^2 - a^2 - b^2) \frac{dZ}{d\mu} - \left[n(n+1)\mu^2 - (a^2 + b^2)s \right] Z &= 0, \\ (\zeta^2 - a^2)(\zeta^2 - b^2) \frac{d^2 Y}{d\zeta^2} + \zeta(2\zeta^2 - a^2 - b^2) \frac{dY}{d\zeta} - \left[n(n+1)\zeta^2 - (a^2 + b^2)s \right] Y &= 0. \end{aligned} \quad (3.18)$$

The solutions of these equations are ellipsoidal harmonics or Lamé's functions. The Lamé's equations are usually written in a general form as

$$(x^2 - a^2)(x^2 - b^2) \frac{d^2 E_n}{dr^2} + x(2x^2 - a^2 - b^2) \frac{dE_n}{dx} - \left[n(n+1)x^2 - (a^2 + b^2)s \right] E_n = 0. \quad (3.19)$$

For a given n integer, we can find $z = (a^2 + b^2) \cdot s$ such that the particular solution of (3.19) can be written as [18]

$$E_n^z(x) = \varphi_n(x) P_n^z(x). \quad (3.20)$$

Those solutions are the Lamé's functions of first kind with

$$P_n^z(x) = \sum_{j=0}^m a_j x^{2j} \quad \text{is Lamé's polynomial,} \quad (3.21)$$

Table 1: Characteristics of lamé's functions of first kind.

Function	Principal product	m	Number	Value of the index i
$K_n^{z_i}(x)$	$\varphi_n(x) = x^{n-2k}$	k	$k+1$	$i = 0, \dots, k$
$L_n^{z_i}(x)$	$\varphi_n(x) = x^{1-n+2k} \sqrt{ x^2 - a^2 }$	$n-k-1$	$n-k$	$i = k+1, \dots, n$
$M_n^{z_i}(x)$	$\varphi_n(x) = x^{1-n+2k} \sqrt{ x^2 - b^2 }$	$n-k-1$	$n-k$	$i = n+1, \dots, 2n-k$
$N_n^{z_i}(x)$	$\varphi_n(x) = x^{n-2k} \sqrt{ (x^2 - a^2)(x^2 - b^2) }$	$k-1$	k	$i = 2n-k+1, \dots, 2n$
Total = $2n+1$.				

where m depends on the integer k , as shown in Table 1:

$$k = \begin{cases} \frac{n}{2}, & \text{if } n \text{ is even,} \\ \frac{n-1}{2}, & \text{if } n \text{ is odd.} \end{cases} \quad (3.22)$$

$\varphi_n(x)$ is called the principal product.

There are four different Lamé's functions that differ from their characteristics. We present them in Table 1.

z_i is an eigenvalue. We remark that z_i corresponds to a value of the parameter z in (3.19), which determines as many equations as there are values of z_i .

For a given value of n , the unknowns are the coefficients a_j (3.23) and the parameters z_i . We consider as a first step the functions $K_n^{z_i}(x)$; the other functions could be defined similarly.

If we introduce the expressions (3.20) and (3.21) in (3.19), we obtain the following recurrence's relation:

$$\underbrace{2(k-j+1)(2n-2k+2j-1)}_{\lambda_j} a_{j-1} + \underbrace{\left[(a^2 + b^2)(n-2k+2j)(n-2k+2j) - z \right]}_{\varpi_j} a_j - \underbrace{a^2 b^2 (n-2k+2j+2)(n-2k+2j+1)}_{\sigma_j} a_{j+1} = 0, \quad (3.23)$$

or

$$\lambda_j a_{j-1} + (\varpi_j - z) a_j + \sigma_j a_{j+1} = 0 \quad \text{for } j = 0, \dots, k. \quad (3.24)$$

Knowing that $a_{k+1} = 0$, the iteration will stop at rank k . And if we introduce the conditions:

$$a_{-1} = a_{k+1} = 0, \quad (3.25)$$

the whole recurrence's relations (3.28) can be written in a matrix form such as

$$\underbrace{\begin{pmatrix} \varpi_0 & \sigma_0 & 0 & 0 & 0 & \cdots & 0 \\ \lambda_1 & \varpi_1 & \sigma_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \varpi_2 & \sigma_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \varpi_3 & \sigma_3 & & \vdots \\ \vdots & \vdots & & & & \ddots & 0 \\ 0 & 0 & \cdots & 0 & \lambda_{k-1} & \varpi_{k-1} & \sigma_{k-1} \\ 0 & 0 & \cdots & 0 & 0 & \lambda_k & \varpi_k \end{pmatrix}}_{\Omega_K} \cdot \underbrace{\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ \vdots \\ \vdots \\ a_{k-1} \\ a_k \end{pmatrix}}_{\Lambda} = z \cdot \underbrace{\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ \vdots \\ \vdots \\ a_{k-1} \\ a_k \end{pmatrix}}_{\Lambda}, \quad (3.26)$$

or

$$\Omega_K \Lambda = z \cdot \Lambda. \quad (3.27)$$

Ω_K is a square matrix of dimension $(k + 1)$. The vector Λ is an eigenvector associated to the eigenvalue z .

There are basically $(2k + 1)$ eigenvalues and eigenvectors. In fact we show that it is possible to find a diagonal matrix D and a symmetric matrix S_K such as

$$S_K = D \cdot \Omega_K \cdot D^{-1}, \quad (3.28)$$

with

$$D = \begin{pmatrix} c_0 & 0 & 0 & \cdots & 0 \\ 0 & c_1 & 0 & \cdots & \vdots \\ 0 & 0 & c_2 & \cdots & \vdots \\ & & & \ddots & 0 \\ 0 & 0 & \cdots & 0 & c_k \end{pmatrix}, \quad \text{where} \begin{cases} c_0 = 1, \\ c_j = \sqrt{\frac{\sigma_{j-1}}{\lambda_j}} \cdot c_{j-1} \quad \text{for } j = 1, \dots, k. \end{cases} \quad (3.29)$$

The matrix S_k is diagonalizable such as $S_k = R^T \cdot \Delta \cdot R$, where Δ is diagonal and R is an orthogonal matrix. Accordingly, Ω_K is diagonalizable and admits $2k + 1$ separate eigenvalues z_i associated to $k + 1$ eigenvectors Λ_i , so $k + 1$ functions $K_n^{z_i}(x)$ for $i = 0, \dots, k$.

Ω_K has the same eigenvalues than S_k . Knowing that S_k is symmetric, its eigenvalues z_i and eigenvectors Λ_s are obtained conventionally using the algorithm QR. Then the eigenvectors of Ω_K are given by $\Lambda = D^{-1} \Lambda_s$.

The computation of the three other functions $L_n^{z_i}(x)$, $M_n^{z_i}(x)$, and $N_n^{z_i}(x)$ is in the same manner.

Each of the functions $K_n^{z_i}(x)$, $L_n^{z_i}(x)$, $M_n^{z_i}(x)$, and $N_n^{z_i}(x)$ admits exactly i zeros in the interval $]a, b[$.

For $i \geq k/2$, the value of z_i corresponding to $K_n^{z_i}(x)$ is equal to z_{i-1} corresponding to $N_n^{z_{i-1}}(x)$, and the values of z_i corresponding to $L_n^{z_i}(x)$ and $M_n^{z_i}(x)$ are identical.

However, the computation shows a significant numerical instability if the value of n increases beyond 10. This is due to the computation of the coefficients a_j in the expression (3.21).

A technique was proposed by Dobner and Ritter [29] to stabilize such computation. They proposed to use another expression of the polynomial P_n such as

$$P_n^z(x) = \sum_{j=0}^m a_j \left(1 - \frac{x^2}{a^2}\right). \quad (3.30)$$

This variation gives a more stable computation.

The product $E(r) \cdot E(\mu) \cdot E(\nu)$ satisfies the potential equation (3.19) within the elliptic cone. However for the external space other kinds of solutions are needed which vanish at infinity.

There is another kind of function $F_n^z(x)$ solution of (3.19) which vanishes at infinity. These functions are called Lamé's functions of second kind. They are defined by Hobson [16]:

$$F_n(x) = (2n+1)E_n(x) \cdot \int_x^\infty \frac{dt}{(E_n(x))^2 \sqrt{t^2 - a^2} \sqrt{t^2 - b^2}}. \quad (3.31)$$

For each value of $E_n^{z_i}(x)$ we have a function $F_n^{z_i}(x)$. We have then $(2k+1)$ Lamé's functions of second kind.

Recalling that surface of the elliptic cone is given by the relation $\mu = \mu_0$, the spatial velocity potential flow Φ , solution of (3.13), is then [16]

$$\Phi = \sum_{n=0}^{\infty} \sum_{i=1}^{2n+1} A_{n,i} \frac{F_n^{z_i}(\mu)}{F_n^{z_i}(\mu_0)} E_n^{z_i}(\rho) \cdot E_n^{z_i}(\zeta), \quad (3.32)$$

where $A_{n,i}$ are constants to be defined using the boundary conditions.

3.3. Boundary Condition

In addition to the relations (3.1)–(3.3) the velocity potential flow Φ verifies a kinematical boundary condition on the surface of the elliptic cone such as

$$\frac{\partial \Phi_f}{\partial n_o} = \nabla \Phi_f \cdot n_o = v_1^T \cdot n_o. \quad (3.33)$$

n_o is an outward normal vector to the surface:

$$n_o = \left(\frac{\partial x}{\partial \mu}, \frac{\partial y}{\partial \mu}, \frac{\partial z}{\partial \mu} \right)^T. \quad (3.34)$$

The relation (3.33) can then be written as

$$u \frac{\partial x}{\partial \mu} + v \frac{\partial y}{\partial \mu} + w \frac{\partial z}{\partial \mu} = \rho \cdot g(\zeta) = \frac{\partial \Phi_f}{\partial \mu} \quad (3.35)$$

with

$$\begin{aligned} \frac{\partial x}{\partial \mu} &= \frac{\rho \zeta}{ab}, \\ \frac{\partial y}{\partial \mu} &= \frac{\rho \mu}{a} \sqrt{\frac{(\zeta^2 - a^2)}{(a^2 - b^2)(\mu^2 - a^2)}}, \\ \frac{\partial z}{\partial \mu} &= \frac{\rho \mu}{b} \sqrt{\frac{(\zeta^2 - b^2)}{(b^2 - a^2)(\mu^2 - b^2)}}, \\ g(\zeta) &= \frac{u \cdot \zeta}{ab} + \frac{v \cdot \mu_0}{a} \sqrt{\frac{(\zeta^2 - a^2)}{(a^2 - b^2)(\mu_0^2 - a^2)}} + \frac{w \cdot \mu_0}{b} \sqrt{\frac{(\zeta^2 - b^2)}{(b^2 - a^2)(\mu_0^2 - b^2)}}. \end{aligned} \quad (3.36)$$

Let us denote $k = a/b$. And suppose that $0 < \alpha < K$ and $0 < \gamma < 4K$ with K is the complete elliptic integral $F(k, \pi/2)$.

For a function of α , β , and γ on the boundary surface $\mu = \mu_0$ we will have

$$\int f(\alpha, \beta, \gamma) dS = \frac{1}{b^2} \int_0^K d\alpha \int_0^{4K} f(\alpha, \beta_0, \gamma) (\rho^2 - \zeta^2) \sqrt{(\rho^2 - \mu^2)(\zeta^2 - \mu^2)} d\gamma. \quad (3.37)$$

We can then solve the problem for the potential Φ in a given point in the space with the boundary conditions defined on the surface of the airship $\mu = \mu_0$.

The relations (3.32) and (3.35) give

$$\left. \frac{\partial \Phi}{\partial \mu} \right|_{\mu=\mu_0} = \sum_{n=0}^{\infty} \sum_{i=1}^{2n+1} A_{n,i} E_n^{z_i}(\rho) E_n^{z_i}(\zeta) \frac{F_n^{z_i}(\mu_0)}{F_n^{z_i}(\mu_0)} = \rho \cdot g(\zeta). \quad (3.38)$$

Using (3.37) and (3.38) we obtain $A_{n,i}$ and then Φ and by the way M_{aij} .

We can now deduce the different components of the mass matrix M .

Finally, the vector of gyroscopic forces Q_G can then be expressed in a developed form as

$$Q_G = \begin{pmatrix} M_{22}vr - M_{33}qw \\ M_{33}pw - M_{11}ur \\ M_{11}uq - M_{22}vp \\ -M_{46}pq + (M_{55} - M_{66})qr \\ M_{46}p^2 + (M_{66} - M_{44})pr - M_{46}r^2 \\ (M_{44} - M_{55})pq + M_{46}qr \end{pmatrix}. \quad (3.39)$$

This leads to the developed dynamic model:

$$\begin{aligned} M_{11} \cdot \dot{u} &= \sum_{i=1}^4 \|F_i\| c\gamma_i \cdot c\beta_i - (mg - B_u)s\theta - M_{33}q \cdot w + M_{22}r \cdot v, \\ M_{22} \cdot \dot{v} &= \sum_{i=1}^4 \|F_i\| s\gamma_i + (mg - B_u)s\phi \cdot c\theta + M_{33}p \cdot w - M_{11}r \cdot u, \end{aligned} \quad (3.40a)$$

$$\begin{aligned} M_{33} \cdot \dot{w} &= - \sum_{i=1}^4 \|F_i\| c\gamma_i \cdot s\beta_i + (mg - B_u)c\phi \cdot c\theta + M_{11}u \cdot q - M_{22}v \cdot p, \\ \dot{p} &= \frac{1}{(M_{44}M_{66} - M_{46}^2)} \\ &\times \left\{ -M_{66}c \sum_{i=1}^4 \|F_i\| s\gamma_i + (M_{46} - M_{66})b_1(\|F_1\| c\gamma_1 \cdot s\beta_1 - \|F_2\| c\gamma_2 \cdot s\beta_2) \right. \\ &\quad + (M_{46} - M_{66})b_3(\|F_3\| c\gamma_3 \cdot s\beta_3 - \|F_4\| c\gamma_4 \cdot s\beta_4) \\ &\quad + M_{46}a(\|F_4\| s\gamma_4 + \|F_3\| s\gamma_3 - \|F_1\| s\gamma_1 - \|F_2\| s\gamma_2) - M_{66}B_u z_G s\phi \cdot c\theta \\ &\quad \left. - M_{46}(M_{44} - M_{55} + M_{66})pq + (M_{55}M_{66} - M_{46}^2 - M_{66}^2)qr \right\}, \\ M_{55} \cdot \dot{q} &= -c \sum_{i=1}^4 \|F_i\| c\gamma_i \cdot c\beta_i + a(\|F_4\| c\gamma_4 \cdot s\beta_4 + \|F_3\| c\gamma_3 \cdot s\beta_3 - \|F_1\| c\gamma_1 \cdot s\beta_1 - \|F_2\| c\gamma_2 \cdot s\beta_2) \\ &\quad - B_u z_G \cdot s\theta + M_{46}p^2 + (M_{66} - M_{44})pr + M_{46}r^2, \\ \dot{r} &= \frac{1}{(M_{44}M_{66} - M_{46}^2)} \\ &\times \left\{ M_{46} \cdot c \sum_{i=1}^4 \|F_i\| s\gamma_i + (M_{46} - M_{44})b_1(\|F_1\| c\gamma_1 \cdot s\beta_1 - \|F_2\| c\gamma_2 \cdot s\beta_2) \right. \\ &\quad + (M_{46} - M_{44})b_3(\|F_3\| c\gamma_3 \cdot s\beta_3 - \|F_4\| c\gamma_4 \cdot s\beta_4) \\ &\quad - M_{44}a(\|F_4\| s\gamma_4 + \|F_3\| s\gamma_3 - \|F_1\| s\gamma_1 - \|F_2\| s\gamma_2) + M_{46}B_u z_G s\phi \cdot c\theta \\ &\quad \left. + (M_{44}^2 + M_{46}^2 - M_{44}M_{55})pq + M_{46}(M_{44} - M_{55} + M_{66})qr \right\}. \end{aligned} \quad (3.40b)$$

4. Results and Discussion

In this section we present the computation results of the added mass matrix of the MC500. Since any comparison with a similarly shaped airship is impossible at present, we conducted the calculation by the geometric method that we present here briefly in order to compare the results of our method.

Geometric method is well known and discussed extensively elsewhere [6, 8, 12, 20]. In a 2D analysis the planar coefficients m_{ij} are established for the standard shapes (rectangle,

Table 2: Comparison between the two methods of computation.

Added masses terms	M_{11} (kg)	M_{22} (kg)	M_{33} (kg)	M_{44} (kg·m ²)	M_{55} (kg·m ²)	M_{66} (kg·m ²)	M_{46} (kg·m ²)
Velocity potential method	583	620	687	9413	10456	18700	160
Geometric method	648	633	708	9973	11920	19341	168

circle, ellipse, etc.). Following the exhaustive study of Brenner [8], we model our flying wing as a truncated cone (T) with elliptic section.

The airship is divided into a dozen cross-sections to optimize the inclusion of changes in transverse dimensions in the 3D calculation.

The computing of the terms of the added masses matrix can be seen, for example, as

$$M_{a11} = \int_{(T)} m_{11}(y, z) dx, \quad (4.1)$$

where m_{11} is a 2D added mass coefficient for the forward motion.

According to the large difference of size between the diagonal and off-diagonal terms, we will neglect these last terms, keeping only the term M_{a46} .

4.1. Results

We present here some characteristics of the geometry of the airship: $z_G = 0.5$ m; $a = 2.5$ m; $b = 6.5$ m; $c = 2$ m; $b_1 = 5.4$ m; $b_3 = 6.5$ m; volume $V = 500$ m³. Numerical results are presented in Table 2.

4.2. Discussion

The first results described here show that thanks to the application of the velocity potential flow theory to this unconventional airship it is possible to obtain reasonable values of the added masses matrix. To our knowledge, this is the first attempt to compute the added masses using this technique for an elliptic cone-shaped airship.

Some differences can be observed between the two methods to certain terms of the added masses matrix. Experimental studies on the prototype nearing completion will confirm the accuracy of our method.

Although some geometric assumptions are made, it nonetheless demonstrates the capability of this method to compute interesting values of the added masses matrix.

Validation of our technique will allow a good estimation of the added masses matrix in preprocessing phase for this kind of airship. The development of an airship is usually done by iterative techniques (model, stability, propulsion, sizing of the control, rudders, etc.); obtaining a fairly accurate aerodynamic model early design allows better refinement of the final model the airship.

5. Conclusion

In this paper a dynamic model of an unconventional airship is presented. The original shape of the careen induces a necessary reformulation of the dynamic and aerodynamic study of these flying objects. A special focus is put on the computation of the added masses. According with the original shape, an assumption of elliptic cone was made to define the careen. The ellipsoidal harmonics and the Lamé's equations are revisited for the analysis of the velocity potential flow, according to the constraints of fluid-structure interaction. The first results seem promising.

Nomenclature

$\eta_1 = [x_0, y_0, z_0]^T$: Vector position of the origin expressed in the fixed reference frame R_0
$\eta_2 = [\phi, \theta, \psi]^T$: Vector orientation of the pointer R_m in regards to R and given by the Euler angles
$\eta = [\eta_1, \eta_2]^T$: Vector attitude compared to R_0
$\dot{\eta}$: Velocity vector compared to R_0 expressed in R_0
$v_1 = [u, v, w]^T$: Velocity vector expressed in R_m
$v_2 = [p, q, r]^T$: Vector of angular velocities expressed in R_m
$v = [v_1, v_2]^T$: The 6×1 velocity vector
m	: The mass of the airship
I_3	: The identity matrix 3×3 .

References

- [1] L. Liao and I. Pasternak, "A review of airship structural research and development," *Progress in Aerospace Sciences*, vol. 45, no. 4-5, pp. 83–96, 2009.
- [2] Y. Li, M. Nahon, and I. Sharf, "Airship dynamics modeling: a literature review," *Progress in Aerospace Sciences*, vol. 47, no. 3, pp. 217–239, 2011.
- [3] Y. Li and M. Nahon, "Modeling and simulation of airship dynamics," *Journal of Guidance, Control, and Dynamics*, vol. 30, no. 6, pp. 1691–1700, 2007.
- [4] H. Jex and P. Gelhausen, "Control response measurements of the Skyship 500 Airship," in *Proceedings of the 6th AIAA Conference Lighter than Air Technology*, pp. 130–141, New York, NY, USA, 1985.
- [5] S. Bennaceur, N. Azouz, and D. Boukraa, "An efficient modelling of flexible airships: lagrangian approach," in *Proceedings of the ESDA 2006 on ASME International Conference*, Torino, Italy, 2006.
- [6] H. Lamb, *On the Motion of Solids Through a Liquid. HydrodynAmics*, Dover, New York, NY, USA, 6th edition, 1945.
- [7] W. Meyerho, "Added masses of thin rectangular plates calculated from potential theory," *Journal of Ship Research*, vol. 14, no. 2, pp. 100–111, 1970.
- [8] C. H. Brenner, "A review of added mass and fluid inertial forces," Report of the Naval Civil Engineering Laboratory CR 82. 10, 1982.
- [9] M. Munk, "Some tables of the factor of apparent additional mass," Tech. Rep. NACA-TN-197, 1924.
- [10] N. Bessert and O. Frederich, "Nonlinear airship aeroelasticity," *Journal of Fluids and Structures*, vol. 21, no. 8, pp. 731–742, 2005.
- [11] K. El Omari, E. Schall, B. Koobus, and A. Dervieux, "Inviscid flow calculation around flexible airship," *Mathematical Symposium Garcia De GalDeano*, vol. 31, pp. 535–544, 2004.
- [12] A. Korotkin, *Added Masses of Ship Structures*, Springer, New York, NY, USA, 2009.
- [13] G. Lamé, "Sur les surfaces isothermes dans les corps homogènes en équilibre de température," *Journal de Mathématiques Pures et Appliquées*, vol. 2, pp. 147–188, 1837.

- [14] C. Hermite, "Sur quelques conséquences arithmétiques des formules de la théorie des fonctions elliptiques," *Acta Mathematica*, vol. 8, no. 1, 1886.
- [15] W. Byerly, 1893 *An Elementary Treatise on Fourier's Series and on Spherical, Cylindrical and ellipsoidal harmonics*, Ed Dover, New York, NY, USA.
- [16] E. W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics*, Chelsea Publishing Company, New York, NY, USA, 1955.
- [17] J. Boersma and J. K. M. Jansen, *Electromagnetic Field Singularities at the Tip of an Elliptic Cone*, vol. 90, Eindhoven University of Technology, Eindhoven, The Netherlands, 1990.
- [18] R. Garmier and J.-P. Barriot, "Ellipsoidal harmonic expansions of the gravitational potential: theory and application," *Celestial Mechanics & Dynamical Astronomy. An International Journal of Space Dynamics*, vol. 79, no. 4, pp. 235–275, 2001.
- [19] H. Goldstein, *Classical Mechanics*, Saffko & Poole, 3rd edition, 2001.
- [20] T. Fossen, *Guidance and Control of Ocean Vehicles*, Wiley press, 1996.
- [21] S. Bennaceur, A. Abichou, and N. Azouz, "Modelling and control of flexible airship," in *Proceedings of the 1st Mediterranean Conference on Intelligent Systems and Automation (CISA '08)*, pp. 397–407, July 2008.
- [22] Z. Gao and S. X. Ding, "Fault estimation and fault-tolerant control for descriptor systems via proportional, multiple-integral and derivative observer design," *IET Control Theory & Applications*, vol. 1, no. 5, pp. 1208–1218, 2007.
- [23] Z. Gao, H. Wang, and T. Chai, "A robust fault detection filtering for stochastic distribution systems via descriptor estimator and parametric gain design," *IET Control Theory & Applications*, vol. 1, no. 5, pp. 1286–1293, 2007.
- [24] Z. Gao, "PD observer parametrization design for descriptor systems," *Journal of the Franklin Institute*, vol. 342, no. 5, pp. 551–564, 2005.
- [25] F. Axisa and J. Antunes, *Modelling of Mechanical Systems: Fluid-Structure Interaction*, Butterworth-Heinemann, 2006.
- [26] F. Imlay, "The complete expressions for added mass of a rigid body moving in an ideal fluid," Tech. Rep. 1528, Hydrodynamics laboratory, 1961.
- [27] L. Kraus and L. M. Levine, "Diffraction by an elliptic cone," *Communications on Pure and Applied Mathematics*, vol. 14, pp. 49–68, 1961.
- [28] G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers*, McGraw-Hill Book, New York, NY, USA, 1968.
- [29] H.-J. Dobner and S. Ritter, "Verified computation of Lamé functions with high accuracy," *Computing*, vol. 60, no. 1, pp. 81–89, 1998.

Research Article

Modules Identification in Gene Positive Networks of Hepatocellular Carcinoma Using Pearson Agglomerative Method and Pearson Cohesion Coupling Modularity

Jinyu Hu¹ and Zhiwei Gao²

¹ School of Electrical Engineering and Automation, Tianjin University, Tianjin 300072, China

² School of Computing, Engineering and Information Sciences, Northumbria University, Newcastle Upon Tyne NE1 8ST, UK

Correspondence should be addressed to Zhiwei Gao, zhiwei.gao@northumbria.ac.uk

Received 5 June 2012; Accepted 26 June 2012

Academic Editor: Dexing Kong

Copyright © 2012 J. Hu and Z. Gao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this study, a gene positive network is proposed based on a weighted undirected graph, where the weight represents the positive correlation of the genes. A Pearson agglomerative clustering algorithm is employed to build a clustering tree, where dotted lines cut the tree from bottom to top leading to a number of subsets of the modules. In order to achieve better module partitions, the Pearson correlation coefficient modularity is addressed to seek optimal module decomposition by selecting an optimal threshold value. For the liver cancer gene network under study, we obtain a strong threshold value at 0.67302, and a very strong correlation threshold at 0.80086. On the basis of these threshold values, fourteen strong modules and thirteen very strong modules are obtained respectively. A certain degree of correspondence between the two types of modules is addressed as well. Finally, the biological significance of the two types of modules is analyzed and explained, which shows that these modules are closely related to the proliferation and metastasis of liver cancer. This discovery of the new modules may provide new clues and ideas for liver cancer treatment.

1. Introduction

Hepatocellular carcinoma (HCC) is one of the most common malignant tumors in the world. Most of the liver cancer patients are in advanced states when they are firstly clinically diagnosed, which leads to poor treatment and high mortality. It is known that the nature of liver cancer is abnormal expression of genes caused by a variety of reasons. There are a lot of modules in a cell, and these modules work together to implement a function of the cell. The functional modules are composed of genes which are similar to each other

in physiological or functional aspects. When the gene functional modules receive impacts, they may lead to disease [1]. Microarray data are the results of many gene expressions, which consist in the information of the gene function modules [2]. There is a very important biological significance to identify gene functional modules in terms of a large-scale gene expression profiling. Cancer gene therapy has become a new treatment method following surgical resection, radiotherapy and chemotherapy, and interventional therapy. For instance, the recently discovered adeno-associated virus AAV3 may be useful for attacking human liver cancers.

The gene network is a complex dynamic system. Therefore, the process of finding gene modules is actually a process of discovering community structure from a complex network. The correlation between genes may be strong or weak, leading to a variant of collections of genes. Clearly, a strong community of a collection of related genes is what we are looking for. Currently, there are a number of community discovery methods of complex networks such as GN-splitting algorithm [3] and NEWSMAN cohesion algorithm [4], both of them use graphs without weights. In order to reflect the size of the gene intensity, it motivates us to use graphs with weights. It is worthy to mention that Pearson values can be used to measure the correlation between genes. As Pearson values may be positive or negative, the absolute of Pearson was used to express the weights in [5] so that the intensity of correlation was obtained. Unfortunately, the positive correlation (mutual promotion) and negative correlation (mutual inhibition) were not considered in the method [5]. The genes in functional modules are a collaboration which may have mutually reinforcing relationships. Therefore it motivates us to use the Pearson values greater than 0 as the weights to reflect the concerned positive correlations of genes.

In this study, we are looking for the genes with similar function; therefore we will use an undirected weighted graph to describe gene network relationship. To the best of our knowledge, the present functions of modules are mainly for gene network graphs without weights, for example, Q function [6], which are invalid for our weighed graphs. Motivated by this, we design a PCC modularity algorithm to measure the performance of the modular decomposition.

According to experiments, if the module decomposition is optimized without considering the size of the threshold, it may lead to the obtained decompositions make no practical sense. In order to overcome this drawback, we propose a modified algorithm, that is, interval PCC modularity (IPM). For instance, in order to obtain a set of very strong modules, we preset the threshold range at the interval $[0.8, 1]$, and we can find a maximum modularity in the interval.

2. Construction of the GPN Network

2.1. Definition of Network and Storage

A connection matrix C is used to store gene community networks (GCN), whose element is c_{ij} , defined by

$$c_{ij} = \begin{cases} p_{ij} & i > j \\ 0 & i \leq j, \end{cases} \quad (2.1)$$

where p_{ij} is the Pearson correlation coefficient of the nodes i and j , and $p_{ij} \in [-1, 1]$.

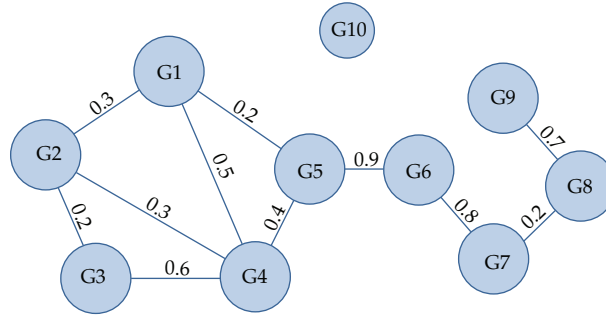


Figure 1: A GPN network: GPN10.

Remark 2.1. (1) $p_{ij} = 0$ indicates there is not linear relationship between two genes.

(2) $p_{ij} > 0$ implies a positive correlation between genes. Particularly when $p_{ij} = 1$, it indicates a completely positive correlation between genes.

(3) $p_{ij} < 0$ means a negative correlation between genes, and $p_{ij} = -1$ represents a completely negative correlation between genes.

Therefore from (2.1), one has $c_{ij} \in [-1, 1]$. When $c_{ij} = 0$ there is no edge between nodes i and j .

The connection matrix is a lower triangular matrix, which stores an undirected weighted graph. It is noticed that no closed loops and no two-directional sides exist in the graph. This matrix is named as Pearson connection matrix (PCM).

2.2. The Type of the Network

The GCN networks can be divided into three kinds.

- (1) Gene positive network (GPN): extract a network from the GCN where the values of all the edges in the network are greater than 0.
- (2) Gene negative network (GNN): remove the edges with the weights greater than 0 to form a network.
- (3) Gene absolute network (GAN): the weights of the edges are taken as the absolute values of the GCN network.

Genes in a module should reinforce mutually, which means the Pearson value of genes should be greater than 0. As a result, the GPN network will be used in this study, which is defined as follows:

$$c_{ij} = \begin{cases} p_{ij} & i > j, p_{ij} \in (0, 1] \\ 0 & i \leq j. \end{cases} \quad (2.2)$$

A GPN network of 10 genes is depicted by Figure 1.

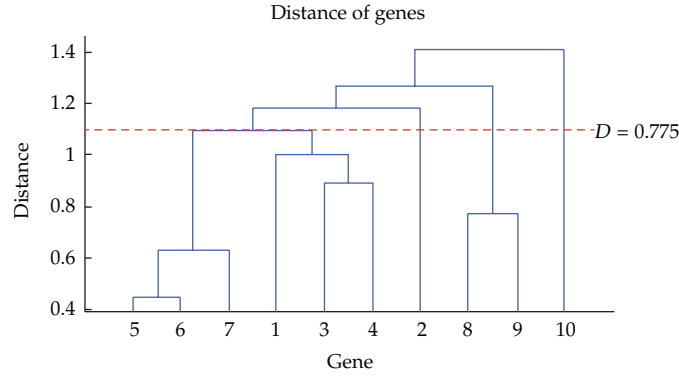


Figure 2: The PAM clustering tree of GNP10.

3. Pearson Agglomerative Method (PAM)

3.1. The Basic Idea of the PAM Algorithm

Hierarchical clustering is a conventional method to find a social network community structure, which can be classified into two types [7]: agglomerative method and divisive method. The main principle of the PAM proposed here is to first calculate the functional similarity between nodes using the Pearson method. Then to add new edges to a raw network composed of n nodes and 0 edge starting from the node pairs with the most similarity. This process is repeated and may end in any node. This procedure from an empty graph to form a resulting graph can be described by Figure 2. In this figure, x -axis is the node and y -axis is the distance between nodes. We call this tree structure as PAM Clustering tree.

The distance of the genes is defined by [8]

$$d_{ij} = \sqrt{1 - c_{ij}}, \quad c_{ij} \in [0, 1]. \quad (3.1)$$

The larger the distance is, the relationship between genes become farther and vice versa. In our GPN, the weights are positives. Therefore the definition (3.1) can be used to measure disturbances of genes in this study.

In Figure 2, when the dashed line moves up from the bottom of the tree gradually, the variant nodes can be integrated into a greater community. The whole network becomes one community as the dashed line move up to the top. The tree structure corresponds to different community structure when it is cutoff from any location using dashed line.

If the red dotted line is placed at 0.775, we can obtain five modules as follows: {G5, G6, G7}, {G1, G3, G4}, {G2}, {G8, G9}, and {G10}.

3.2. Algorithm Implementation

In terms of the definition of the distance matrix $D = [d_{ij}]$, the PAM algorithm can be addressed in the following steps.

- (1) Initialize the network with n communities, where every node is a unique community.

- (2) Calculate the minimum distance using Dijkstra algorithm [9], and then combine the minimum distance node pairs.
- (3) Repeat step 2 until the whole network is merged to one community. The maximum execution times are $n - 1$.

We call this tree structure as “PAM clustering tree”.

4. PCC Modularity Assessments

Similar to the GN algorithm [3] and other decomposition methods, the PAM algorithm cannot determine what kind of decomposition is optimal. It is of significance to value the performance of the decompositions. In [6], the modularity Q function was used to measure the quality of network partitioning. Unfortunately, for communities with big differences, the Q function is not ideal to assess the performance of module partitioning [9]. For the weighted networks, we introduce a standard function to measure the quality of network partitioning, namely, the PCC module function.

Here, the network is assumed to be divided into N modules: C_1, C_2, \dots, C_N .

4.1. Cohesion (Coh)

Cohesion (Coh) is the measure of the relevance of the internal nodes in a module. For a module with n nodes, the maximum edges are $(n(n - 1))/2$ and the maximum weight of the edge is 1. We use the ownership of the internal value which is divided by $(n(n - 1))/2$ to represent the cohesion, described by

$$\text{Coh}(C_i) = \begin{cases} \frac{I(C_i)}{n_i(n_i - 1)/2} & n_i > 1 \\ 1, & n_i = 1, \end{cases} \quad (4.1)$$

where C_i is the i th module; n_i is the number of nodes of the module; $I(C_i)$ is the sum of the ownership values in the module, expressed as $I(C_i) = \sum_{k=1}^{m_i} w_k$, where m_i is the number of edges in the module and w_k is the weight for the k th edge. The weights are not greater than 1, so $I(C_i) \leq m_i$. If the module has only one node, its cohesion is defined as 1.

4.2. Coupling (Cou)

The coupling (Cou) is a measure of the degree of association between modules, defined by

$$\text{Cou}(C_i) = \begin{cases} \frac{O(C_i)}{I(C_i) + O(C_i)}, & n_i > 1 \\ 1, & n_i = 1, \end{cases} \quad (4.2)$$

where $O(C_i)$ is the sum of the weights of external edges connected to the module, expressed as $O(C_i) = \sum_{r=1}^{s_i} w_r$, in which s_i is the number of external edges connected to the module, and w_r is the weight of the r th edge. If the module has only one node, its coupling is defined

as 1. When $1/2 < \text{Cou}(C_i) < 1$, it is equivalent to $I(C_i) < O(C_i)$, which indicates the internal strength of the module is less than the external strength, and the division of this module is generally unreasonable.

Equation (4.2) reflects the dependence of a module to the other. The lower is the coupling, the higher is the independence of the modules.

4.3. PCC Modularity

It is evident that the model partition needs high cohesion and low coupling. The formula $\text{Coh}(C_i)[1 - \text{Cou}(C_i)]$ can be used to describe the tightness of connection within the community C_i and reflect the independence of the community. Thus, the new modularity is defined as

$$\text{PCC}(C_i) = \text{Coh}(C_i)[1 - \text{Cou}(C_i)] \quad (4.3)$$

The PCC can be understood as “Pearson”, “Cohesion”, and “Coupling”. Substitution (4.1) and (4.2) into (4.3) yields:

$$\text{PCC}(C_i) = \begin{cases} \frac{2I(C_i)^2}{n_i(n_i - 1)[I(C_i) + O(C_i)]}, & n_i > 1 \\ 0, & n_i = 1. \end{cases} \quad (4.4)$$

If a module has only one node, it is called outlier module, whose PCC is 0.

The average of all the modules of PCC is used to measure the division of the entire network, which is defined as

$$\text{PCC}(C_1, C_2, \dots, C_N) = \frac{\sum_{i=1}^N \text{PCC}(C_i)}{N}, \quad (4.5)$$

where N denotes the number of modules. When the number of nonisolated point modules is R , the number of outlier modules is $N - R$. Since the PCC value of outlier module is 0, (4.5) can be rewritten as

$$\text{PCC}(C_1, C_2, \dots, C_N) = \frac{\sum_{i=1}^R \text{PCC}(C_i)}{N}. \quad (4.6)$$

In other words, the network is divided into N modules. Since the PCC value of each module is not greater than 1, one thus has $\text{PCC}(C_1, C_2, \dots, C_N) \leq R/N$.

5. The Relationship of Threshold and Modularity

5.1. One-to-One Map of Threshold to Modularity

From Figure 2, when the dotted line is set as $D = 0.775$, the network can be divided into five modules. In terms of (3.1), the threshold can be calculated as $T = \sqrt{1 - D^2} = 0.41$. Removing the edges whose PCC values are less than 0.41, Figure 1 can be transformed into Figure 3.

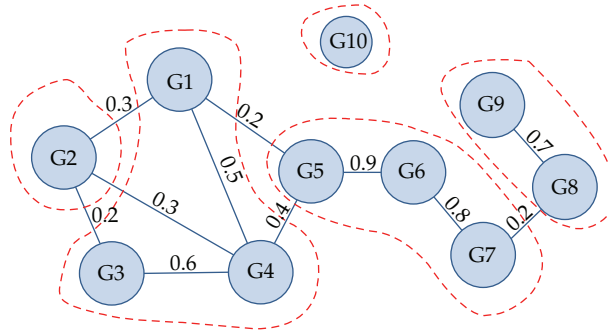


Figure 3: The decomposition of the network GPN10 by using $T = 0.41$.

Table 1: The relationships among T , PCC, and N .

T	0	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
N	2	3	4	5	6	7	8	9	10
PCC	0.07	0.24	0.18	0.22	0.19	0.13	0.05	0.04	0

As a result, when the threshold is $T = 0.41$, the network GPN10 in Figure 1 is divided into five modules: $\{G5, G6, G7\}$, $\{G1, G3, G4\}$, $\{G2\}$, $\{G8, G9\}$, and $\{G10\}$. This partition is the same as Figure 2.

Rather than building PAM-tree to divide the modules, we can decompose the network by using threshold value. A modular decomposition corresponds to a module function; therefore a threshold only has a corresponding module function.

For instance, by setting different thresholds, we can obtain the resulting decompositions (see Table 1), and each module corresponds to the modularity of a PCC. Under the same modularity premise, in order to ensure internal correlation of each module stronger, we choose a larger threshold.

Based on the decomposition of the network GPN10, Table 1 reflects the relationships among " T ", " PCC ", and " N ", where " T " is threshold, " PCC " is PCC modularity, and " N " is the number of modules.

As each threshold corresponds to one decomposition, each threshold corresponds to one modularity as well. From Figure 4, when the threshold $T = 0.2$, the decomposition of the network GPN10 is optimal. In this case, the network GPN10 is broken down into three modules: $\{G1, G2, G3, G4, G5, G6, G7\}$, $\{G8, G9\}$, and $\{G10\}$.

5.2. The Definition of Interval PCC Modularity (IPM)

The absolute value of correlation coefficient is greater and the correlation is stronger. The correlation coefficient is close to 1 or -1 , the correlation is very strong. The correlation coefficient is close to 0, the correlation is weak.

Generally, we judge the intensity of two variables by the range of correlation coefficients (see Table 2).

According to Table 2, we define five different ranges of modularity.

- (1) Modularity of very weak correlation: $PCC(T), T \in [0.0, 0.2)$.
- (2) Modularity of weak correlation: $PCC(T), T \in [0.2, 0.4)$.

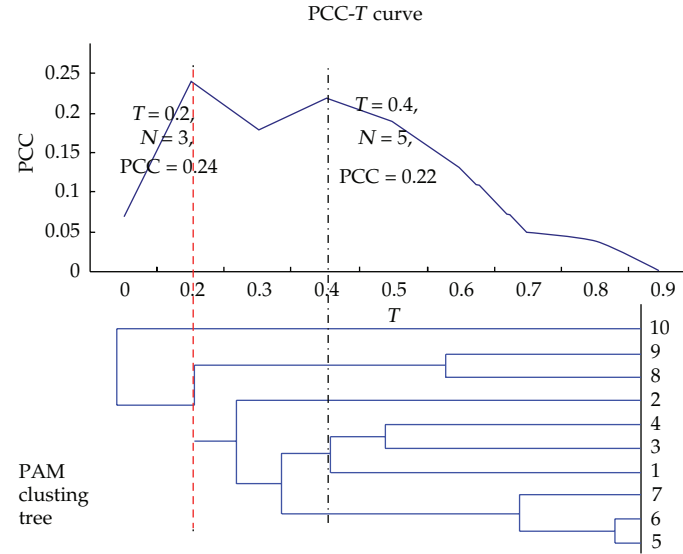


Figure 4: The relationship of thresholds and modularity under the different decomposition of GPN10.

Table 2: The intensity of Pearson correlation coefficient.

Correlation coefficient	0.8–1.0	Very strong correlation
	0.6–0.8	Strong correlation
	0.4–0.6	Moderate correlation
	0.2–0.4	Weak correlation
	0.0–0.2	Very weak or no correlation

(3) Modularity of moderate correlation: $PCC(T)$, $T \in [0.4, 0.6)$.

(4) Modularity of strong correlation: $PCC(T)$, $T \in [0.6, 0.8)$.

(5) Modularity of very strong correlation: $PCC(T)$, $T \in [0.8, 1.0]$.

Generally, we find a strong correlation or strong related modules by using (4.1) and (4.2).

6. Results

6.1. Obtain the HCC Gene Modules

The liver cancer microarray data is taken from Chen et al. [10], which is available at <http://genome-www.stanford.edu/hcc/supplement.shtml>. The 1648 genes are differentially expression in HCC and nontumor liver in 156 liver tissues (74 nontumor liver and 82 HCC). We only study the gene expression of HCC. The Missing values are replaced by the average of the gene expression data under corresponding data column or sequence.

We build the GPN network of 1648 HCC genes. Next we, respectively, test the PCC value in the threshold interval $[0.8, 1]$ and $[0.6, 0.8)$. According to the maximum of PCC value, GPN network is, respectively, divided into the HCC very strong correlation modules and the HCC strong correlation modules.

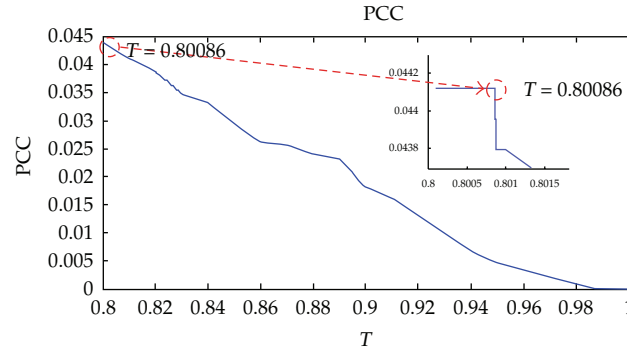


Figure 5: The movements of PCC modularity at the threshold interval $[0.8, 1]$.

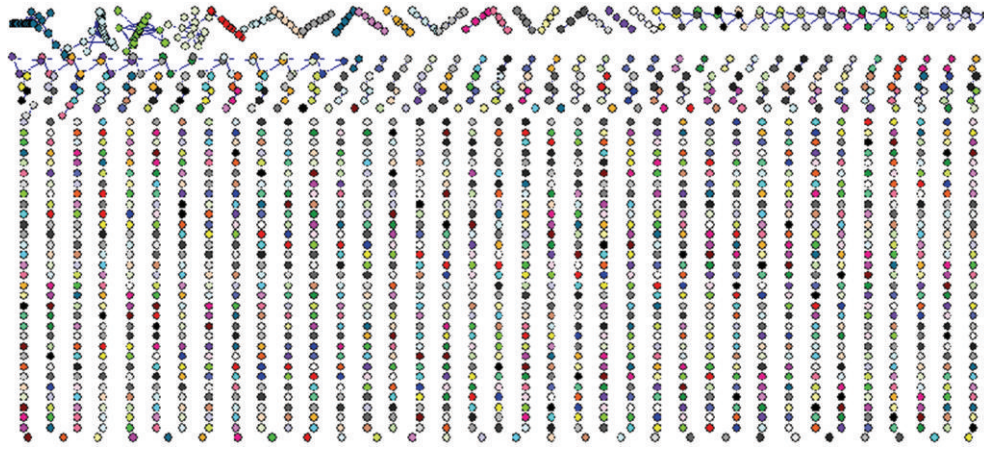


Figure 6: The decomposition diagram of HCC GPN network when the threshold $T = 0.80086$.

6.1.1. Very Strong Correlation Modules (VSCM) of HCC

In the threshold interval $[0.8, 1]$, the PCC curve is given by Figure 5, which shows that the PCC is downtrend. Within the threshold interval $[0.8, 1]$, as the threshold is greater, the modular decomposition is getting worse. When the threshold is between 0.8 and 0.80086, the modularity PCC values are equal. In order to make the module correlation coefficient greater, we choose the threshold $T = 0.80086$. In this case, the modularity $PCC = 0.0441$.

In Figure 6, when the threshold $T = 0.80086$, the network is broken down into 1360 modules, including 150 nonisolated point module. According to formula (4.6), the $PCC < 150/1360 = 0.1103$.

In Figure 7, there are 13 modules and 121 genes in total, where each module is not less than five nodes. The modules numbered and arranged according to gene-related strengths from strong to weak. In order to distinguish the Very Strong Correlation Modules (VSCM) from the strong correlation modules (SCM), we mark VSCM and SCM with "S" and "W" respectively, which means "strong" and "weak".

In Table 3, "NO" is the abbreviation for "No Gene information" and "Trans" means "Transcribed locus". The number of genes is given in the bracket, for example, SERPINA5(5)

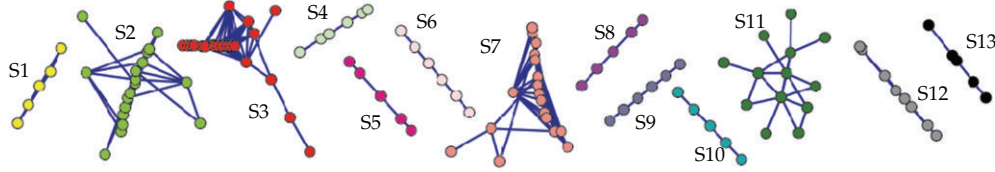


Figure 7: Thirteen gene modules of HCC very strong correlation modules with no less than five nodes.

Table 3: Thirteen gene modules of HCC very strong correlation modules (VSCM).

Module	Number	The detailed genes of each module
S1	5	SERPINA5(5), LOC100507281
S2	18	WNT4, SLU7, CSF2RA, IGKC(3), NCF1, EPB72, IGHG3, Trans, IGL(5), TNFSF10, NAPA, CSF2
S3	21	NO(11), Trans(7), C2orf55, SLC35E1, TRIOBP
S4	6	Trans, ZFP92(2), TAGLN(2), AEBP1
S5	5	C1R(2), C1S(2), FGA
S6	6	CKAP2, AQP4, HAMP, Transcribed locus(3)
S7	18	LRRC8C, EDNRA, BIRC5, MT1B(2), Trans, AGXT, MT1H, MT1G, MT1F(2), MT1E, MT1L, NO, LARP4(3), DLG4
S8	5	NO, Trans, RS10(2), CDNA
S9	6	GRN, C19orf6, RAD23A, ZNF451, RER1, ABCF1
S10	5	TUBA2, TUBA1(2), TUBA3, Trans
S11	13	PLK, TROAP, Trans, CENPM, MYBL2, PTTG1, NUSAP1, CDC20, FOXM1, UBE2C, CDC2, KIAA0101, IFIT1
S12	8	RPS20(2), EIF3S6(2), NO(2), RPL30, Trans
S13	5	SPARC(2), THY1(2), COL4A2

means that there are five SERPINA5 genes. Green denotes the genes in the module with low expression and red indicates the genes with high expression.

6.1.2. Strong Correlation Modules (SCM) of HCC

In the threshold interval $(0.6, 0.8]$, we obtain strong modularity PCC curve as found in Figure 8.

When $T = 0.67302$, the optimal PCC is 0.0687. In this case, the HCC GPN network is divided into $N = 955$ modules (see Figure 9) and the number of the nonisolated modules is $R = 164$. By the formula (4.6), we can get that the $PCC < 164/955 = 0.172$.

We selected the modules with more than five nodes and arranged according to the order of strength from strong to weak order as follows (see Figure 10).

In Figure 10, there are 14 modules and 505 nodes in total. The modules with fewer nodes are W1, W6, W8, W12, W13, and W14. The modules containing a large number of nodes are W11, W4, W3, W2, W5, and W7.

In Table 4, there are 14 strong modules in total, involving 504 genes. The number of gene duplication is marked in brackets. Red means the gene is highly expressed in HCC and green indicates genes in low expression. The genes in W3 have both high expression and low expression, therefore it is not colored.

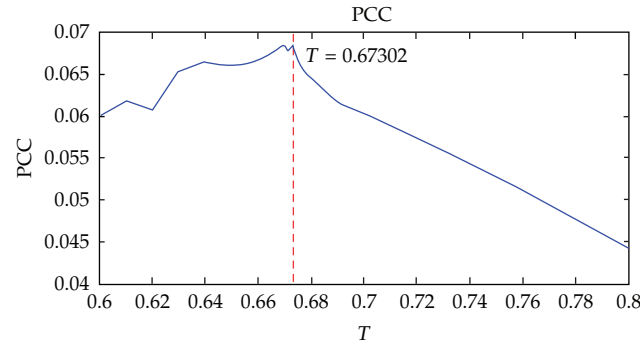


Figure 8: The PCC modularity trend on the threshold interval $[0.6, 0.8)$.

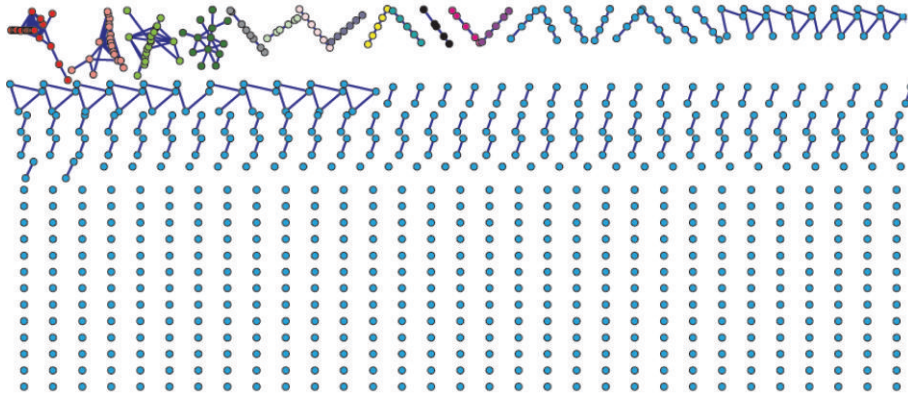


Figure 9: When the threshold $T = 0.67302$, the network is broken down into 955 modules in which there are 164 nonisolated point modules.

Generally, HCC strong correlation modules (SCMs) may include the genes of HCC with very strong correlation modules (VSCMs). In Table 4, blue bolds mark genes which appeared in the VSCMs. This kind of inclusion relations are shown in Table 5.

Table 5 shows the inclusion relationship of the VSCM and the SCMs. The table can also be described as $W5 \geq S3$; $W7 \geq S7$; $W2 \geq B2$; $W11 \geq S11 + S12 + S10 + S8$; $W10 \geq S6$; $W13 \geq S9$; $W1 = S1$; $W3 \geq S4 + S13$; $W4 \geq S5$.

6.2. Biological Explanation of the HCC Gene Modules

The biological explanation of the HCC gene modules, if not otherwise specified, all refers to the Stanford gene database: <http://smd.stanford.edu/cgi-bin/source/sourceSearch>. As there are too many genes, we only provide the biological explanation for the important genes in the each module. According to the major functions of the genes in every group, we name each module.

Table 4: Fourteen gene modules of HCC strong correlation modules (SCM).

Module	Number	The detailed genes of each module
W1	5	SERPINA5(5), LOC100507281
W2	29	WNT4, SLU7, CSF2RA, IGKC(3), NCF1, EPB72, IGHG3, Trans, IGL(5), TNFSF10, NAPA, CSF2, CD69, TF, NO, PSMF1, EDR2, HNT, KLRK1, SYT6, ID4, HCLS1, CD53
W3	40	Trans, ZFP92(2), TAGLN(2), AEBP1, SPARC(2), THY1(2), COL4A2, ID3, COL6A1, FGF12B, TMEM204, MYO10, CSNK2B, PDGFRA, SVEP1, SVEP1, SRPX, CRISPLD2, RBMS3, PYGM, MFAP4, COL6A2, PODN, LAMA2, NGFR, NRG2, CYR61, SLC15A2, SCYA2, TSPYL1, ID4, CRHBP, THY1, NOTCH3, COL15A1, LOXL2
W4	123	CPS1, ZNF248, HRSP12, PCBD1, ALDOB(2), ENC1, APOC3(2), PAH, CD302, POR, Trans(7), SERPING1(2), IVD, APOH, SCYA14, PBP, SORD, EVX1, UGP2(2), C21ORF4, GALE, HSD17B6, CYP2A7, MST1(2), APOA1, C1R(2), C1S(2), FGA, C1RL, PROML1, LRRN3, LANCL1, ACOX1, CYP2C, BDH1, PIPOX, MPDZ, HSD11B1, RGN, PCK1, CHD9, ACAA2(2), FACL2, PON3(2), D4S234E, AZGP, RNAC(2), ADH6(2), ADH4(2), ADH2, APOC4, SLC27A5, MMSDH(3), PCK2, CPB2(2), CPN2, DEPDC7, CYP4V2, LY9, GRHPR, AMDHD1(2), ACADSB, ST3GAL6, SPRYD4, CYB5, ADI1, NO(4), QDPR, PLG, CYP27A1, GYS2, CTH, SHMT, ARHB(2), OGDHL, ACY1, APCS, PXMP2, EDNRB, C14orf45(2), SCP2(2), DHTKD1, KNG, ALAS1, MARC2, SULT2A1, CYP2J2, CTSO, SOD1, MYO1B(2), SYBU, PVRL3, PDK, KIAA0317
W5	25	NO(13), Trans(9), C2orf55, SLC35E1, TRIOBP, L3MBTL4
W6	6	PTMS, SDHAP1, RBP5(2), IKBKAP, HAAO
W7	22	LRRC8C, EDNRA, BIRC5, MT1B(2), Trans(2), AGXT, MT1H, MT1G(2), MT1F(2), MT1E, MT1L, NO, LARP4(3), DLG4, CDK5, TFG
W8	5	AFF4, NUFIP2, LEAP2(3)
W9	9	SERPINA3, FGB(4), Trans, FGA, FGG, CFI
W10	8	CKAP2, AQP4, HAMP, Trans(4), LOC257396
W11	212	EIF4B, RB1CC1(3), MTF2, MAL2, MAL2, CDNA(2), HMG17, CSNK2B, CUTA, ASAP1(2), RCC2, LMNB2, MAPK13, HJURP, SMC4, CMTM1, NO(6), SEMG2, 14ORF4, RPS10(3), Trans(6), RPS16, KLK3, HBG1, RPLP0, RPS5, RPS19, CPNE1, ETV1, TUBB(2), WNK1, RTN3, C1orf43, PAX8, FAM83H, TUBG2, TUBG1, TSEN54, UBE2M, TRIM28, SNRNP, HGS, STARD3, GPS1, CLPTM1, ARF3, ASNA1, TAF2E, USP5, SHC1, VARS2, ASF1B, PKMYT1, SERPINB3, E2F1, NLRP2, H2AFX, MLF1IP, ILF3(2), C1orf9, NAP1L1, SCNM1, LAPTM4B(2), TOP2A, HN1, TUBA2, TUBA1(2), TUBA3, BUB1, HSU, CKS1, CBX1, SLC1A4, KPNA2, EIF4A2, TMEM106C, EHMT2, SF3B4, SCAMP3, FLAD1, TCFL1, UBAP2L, PRCC, UBE2Q1, HTCD37, SNX27(3), PYGO2, FAM189B, NCSTN, RPRD2, USP21, MCM4, SNHG10, GMNN(2), PLK, TROAP, CENPM, MYBL2, CENPW, TPX2, ZNF261, ZWINT, LAP18, PTTG1, NUSAP1, CDC20, FOXM1, UBE2C, CDC2, MAD2L1, KIAA0101, CDKN3(2), RRM2(2), IFIT1(2), FAM72B(2), CEP55(3), KNSL5(2), NUDT1(2), TRIP13, MCM5(2), NRM, CDK4, KIAA1522, RDBP, PEA15, NPM1, UBR5, MRPL42(2), XPOT(2), MZT1, ACLY, PAPP, ILF2(2), TCEB1, ASPH, ATP6V1H, YWHAZ(3), ZNF706, RPS20(2), EIF3S6(2), RPL30, RAD21, BIG1, MTDH, POLR2K, ARMC1(3), COPS5, CANX, KIAA0196, PTK2, TCEA1, NSMCE2, ZHX1, UQCRB, NBS1, FAM49B, DEK, UBA2, TIMP1, PSPH, LAMB1, SRXN1, PIR, TACC3, MCM3, DR1, CDC7, MCM6, RASSF3, POLA, YKT6
W12	7	TIGD5, MAF1, PUF60, CYC1, SHARPIN, GPAA1, Trans
W13	6	GRN, C19orf6, RAD23A, ZNF451, RER1, ABCF1
W14	7	HIST1H2AC, HIST1H2BK, HIST1H2BC, HIST2H2BE(2), CPS1, Trans

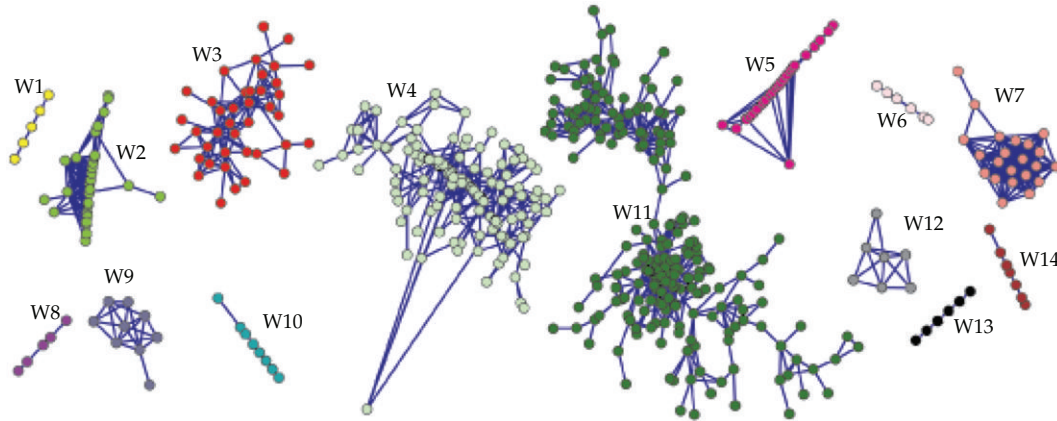


Figure 10: Fourteen gene modules of HCC strong correlation modules with no less than five nodes.

Table 5: The inclusion relationship of the VSCM and the SCMs.

SCM	W5	W7	W2	W11				W10	W13	W1	W3		W4
VSCM	S3	S7	S2	S11	S12	S10	S8	S6	S9	S1	S4	S13	S5

6.2.1. Comparison with Other Results

In this section, we will show the comparison of our results with the experimental results of Chen et al. [10] and Yan et al. [1].

The module S2 (or W2) agrees with the module D in Chen et al. [10] and the module C in Yan et al. [1], which is in relation to B lymphocytes. The disorder of B-cell immune function has a lot to do with liver cancer.

The module S4 agrees with the module E in Chen et al. [10], which is stroma cell module. The function of the module S13 relates to the endothelial cell, and the module G in Chen et al. [10] has the similar capability.

The module W3 contains the genes both from the module S4 and S13. Since S4 genes means stroma cells, while the S13 genes are all located in the stroma, the module W4 is a generalized stroma cell module. The W4 functions as the module D given by Yan et al. [1]. In the HCC, the genes of S4 are in lower level expression, while the expression level of S13 is higher. The module W3 not only contains the highly expressed genes, but also includes the gene with low expression. From this viewpoint, the model W3 is different from any other modules.

The module S5 (or W4) functions as the module K by Yan et al. [1], whose function is about complement. Bacterial infection of the liver cells may be related to genetic disorders of the complement component module.

The module S10 is consistent with the module J by Yan et al. [1], in which the main genes are about the tubulins $\alpha 1$, $\alpha 2$, and $\alpha 3$. Tubulin abnormalities have impacts on the occurrence and development of liver cancer.

The genes of module S11 all appear in the module A given by Chen et al. [10]. The function of the module is also identical to the feature of the module A from Yan et al. [1], which is related to the cell cycle and proliferation of cancer cells.

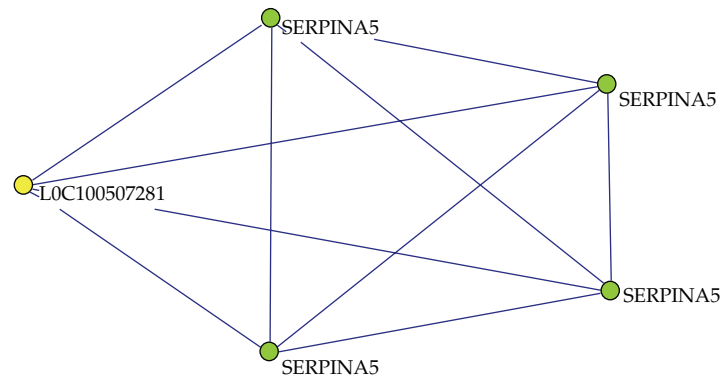


Figure 11: The gene network of module S1.

The genes in the module S12 are about ribosomal proteins, which functions as the same as in the module G in the reference by Yan et al. [1]. Ribosomal protein dysfunction may be related with liver cancer.

The function of module W11 is consistent with the module A given by Chen et al. [10], whose function is about cell proliferation. It is noted that the proliferation module A from Chen et al. [10] does not include S8, S10, and S12. In fact, tumor cell proliferation not only relates to cell cycle, but also relates to ribosomes, and tubulin. Because ribosomal synthesis protein and microtubule protein involves in mitosis, both of the two process are essential in cell proliferation.

The module W14 of histone is consistent with the module I from Yan et al. [1]. The occurrence of liver cancer may be associated with abnormal expression of the histone.

We have found that twelve gene modules such as S2, S4, S13, S5, S10, S11, S12, W2, W3, W4, W11, and W14 have specific functions and are largely in line with the gene clusters found by Yan et al. [1] or Chen et al. [10]. It has proven that the PAM algorithm and the PCC modularity can effectively discover gene function modules.

6.2.2. The Specific Modules of VSCM

In this subsection, we will focus on the modules only existing in this study. The specific modules of VSCM are S1, S3, S6, S7, S8, and S9, in which only S8, S9 are highly expressed.

(1) Hemostasis Module S1

In Figure 11, the key gene of the module S1 is SERPINA5, which plays the hemostatic role in the blood plasma. Additionally, SERPINA5 is able to inhibit the migration of HCC cells. The low expression of SERPINA5 genes probably promotes the occurrence and development of HCC.

(2) Transport Module S3

In Figure 12, NO means no gene information and trans represents Transcribed locus. TROPBP and SLC35E1 are the key genes in S3. TRIOB is closely related to HCC [11]. Moreover,

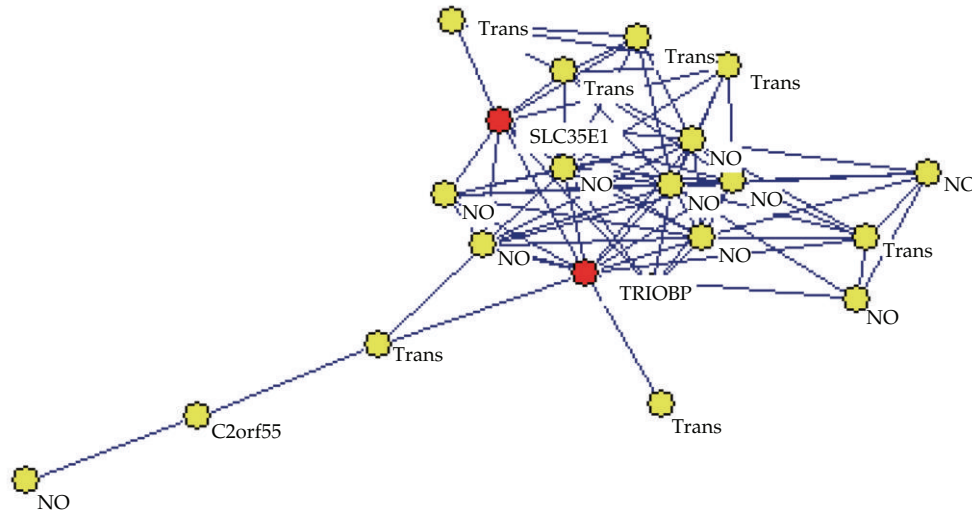


Figure 12: The gene network of module S3.

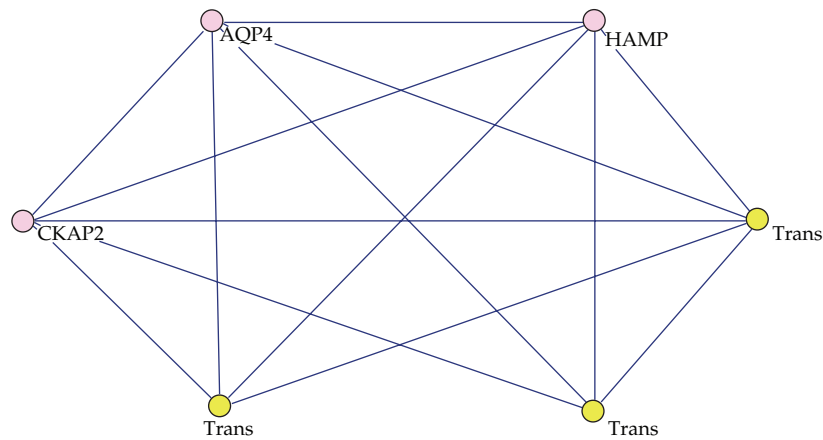


Figure 13: The gene network of module S16.

the transport function of the solute carrier SLC35E1 is related to cancer [12]. Therefore the module should be responsible for transport in hepatoma cells.

(3) Iron Regulation Module S6

The key genes in Figure 13 are CHAP2, AQP4, and HAMP. CKAP2 is a cytoskeleton-associated protein involved in mitotic progression. AQP4 encodes a member of the aquaporin family of intrinsic membrane proteins. Decrease in aquaporin expression [13] may lead to the increase of the resistance to apoptosis in hepatocellular carcinoma. The product encoded by this gene HAMP is involved in the maintenance of iron homeostasis. This module relates to the regulation of the iron.

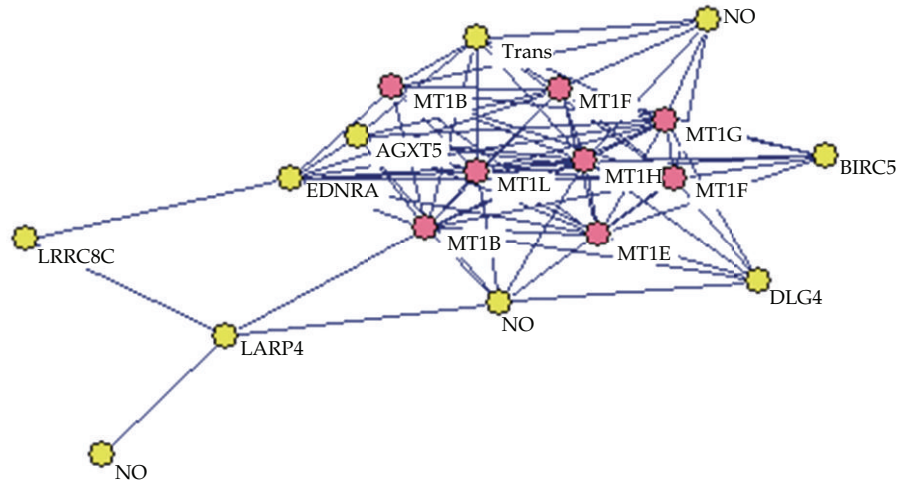


Figure 14: The gene network of module S7.

(4) Metalloproteins Module S7 (see Figure 14)

The genes in this module are most from the metalloproteins (MTs) family, such as MT1B and MT1G. MTs play a key role in the transport of essential heavy metals, detoxification of toxic metals, and protection of cells against oxidation stress. The increasing of the MTs levels in cancer cells are probably related to their increased proliferation and protection against apoptosis [14].

(5) Antiterminator Module S8

From Figure 15, RPS10 (ribosomal protein S10) are the most important genes in this module. RPS10 (as known as NusE) reflects another function [15] which is different from the ribosome module S12. Ribosomal S10 relates to liver cancer [16], and RPS10 play an antitermination role in the transcription process [17]. Therefore, ribosomal S10 may be the antitermination factor for liver cancer.

(6) Immortal Module S9 (see Figure 16)

This module is the most important module of all modules, where any two genes are strongly correlated. For each gene function, we can conclude as follows.

- (1) GRN are a potent growth factor, which can promote the excessive proliferation of tumor cells.
- (2) C19orf6 (also known as membralin) relates NMDAR1 receptor activity, which promotes tumor to differentiation and invasion and metastasis.
- (3) RAD23A involves in negative regulation of HIV-1 replication, and VPR prevents cell division. Therefore RAD23A with high expression is to promote tumor cell division.

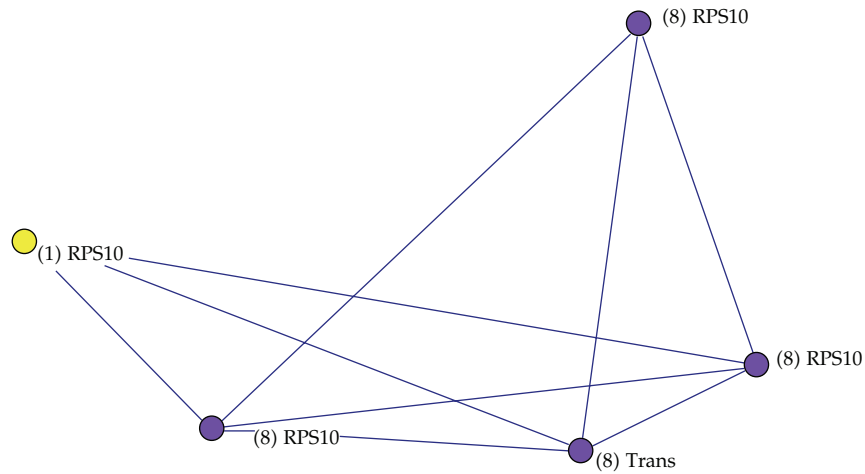


Figure 15: The gene network of module S8.

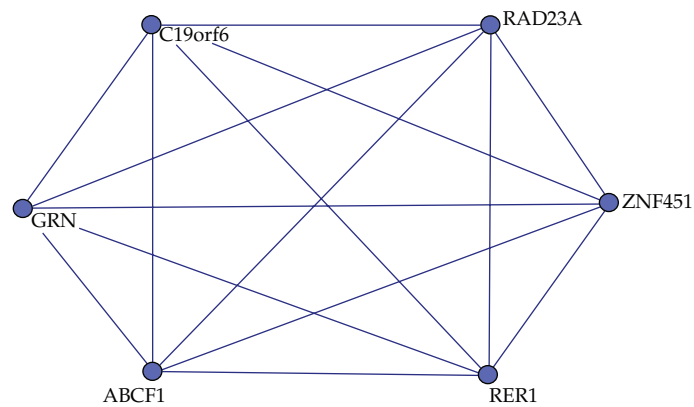


Figure 16: The gene network of module S9.

- (4) ZNF451 may negatively regulate the steroid hormone receptor coactivator of transcription factor (Src), where Src protein plays an important role in the proliferation of hepatoma cells during apoptosis.
- (5) The protein encoded by RER1 is a multipass membrane protein, which facilitates gamma-secretase complex assembly.
- (6) The protein encoded by ABCF1 is a member of the superfamily of ATP-binding cassette (ABC) transporters. This protein may be regulated by tumor necrosis factor-alpha and play a role in enhancement of protein synthesis and the inflammation process. The gene overexpression in HCC will reduce the efficiency of drug treatment.

Summing up the narrative, the module may be the secret of liver cancer cells “immortal”.

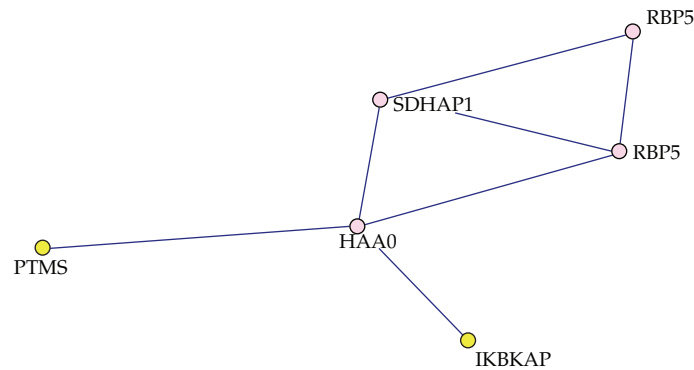


Figure 17: The gene network of module W6.

6.2.3. The Specific Modules of SCM

The specific modules of SCM are W6, W8, W9, and W12, in which only W12 is highly expressed.

(1) Antigrowth Inhibition Module W6

From Figure 17, one can see that the important genes of this module are HAAO, SDHAP1, and RBP5. Specifically, the quinolinic acid, which is the resulting product of HAAO, inhibits the growth of hepatoma cells. SDHAP1 is a marker enzyme of mitochondrial, which provides electron to respiratory chain. Retinoic acid that produced by RBP5 after the second oxidation can inhibit the growth of hepatoma cells.

As a result, the low expression of the genes in this module is to cut off the aerobic capacity of the respiratory chain of electronic sources, making the oxidation products inhibit the growth of cancer cells to be synthesized. Therefore, this module is named as antigrowth inhibitory module.

(2) Antimicrobial Peptides Module W8

In Figure 18, LEAP2 (liver expressed antimicrobial peptide 2) is the most important genes of the module W8, which has antibacterial activity.

(3) Fibrinogen Module W9 (see Figure 19)

There are nine fibrinogen (FIB) genes in module W9, such as FGA, FGB, and FGG. FIB is a glycoprotein synthesized by the liver and plays an important physiological role in the coagulation process. It is worthy to point out that the FIB increases in early stage, but decreases in advanced liver cancer [18]. This module is a low expression; therefore the data should be from advanced liver cancer.

(4) Antiapoptotic Module W12

From Figure 20, the most important three genes in the module W12 are SHARPIN, CYC1 and PUF60. SHARPIN interferes with TNF-induced cell death [19], CYC1 access to electrons

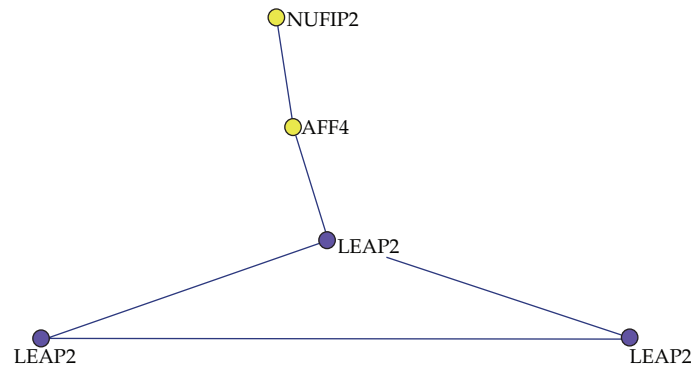


Figure 18: The gene network of module W8.

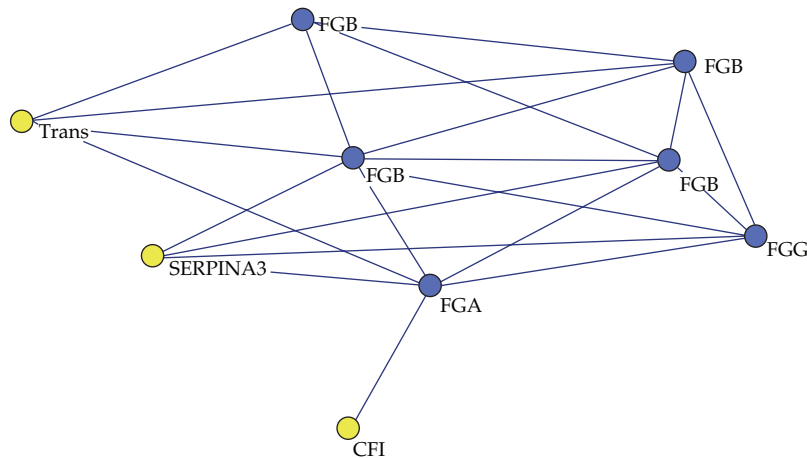


Figure 19: The gene network of module W9.

for respiratory chain, and PUF60 may increase a greater degree of apoptosis resistance of cancer cells. This module provides electrons of respiratory chain to smoothly synthesis anti-apoptotic protein; therefore this module plays the role of antiapoptotic.

7. Conclusions

By using the Pearson agglomerative method (PAM) and Pearson correlation coefficient (PCC) modularity, we have investigated the modules decompositions and the decompositions valuations for liver cancer genes. By using the data from Chen et al. [10], and the proposed methods in this study, we have obtained 13 very strong correlation modules and 14 strong correlation modules. In addition to some common modules, we have found a number of new functional modules.

Coagulation modules are the hemostatic module S1 and fibrin module W9. It is noted that the fibrinogen will be a huge increase in the early liver cancer, but in advanced liver cancer, the fiber protein content would be down to a level slightly lower than normal. Fibrinogen can be used as one of the detection of early stage liver cancer.

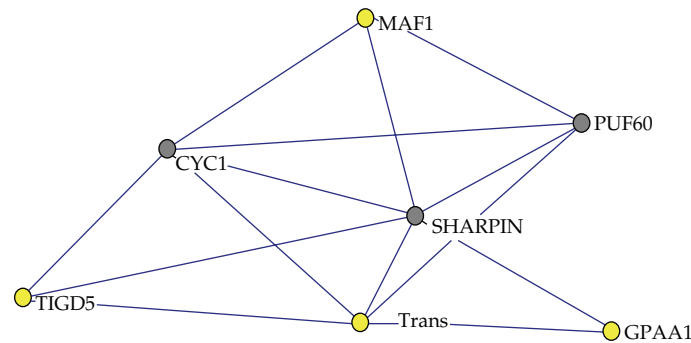


Figure 20: The gene network of module W12.

With respect to the reasons of the cancer cells unlimited reproduction, we have found antiterminate module S8, die module S9, antigrowth inhibition module W6, and antiapoptotic module W12. In which only W6 is in a low expression, and the others are highly expressed. One of the most modules is the immortal module S9, which may be the command center of unlimited reproduction of the entire tumor.

There are two modules associated with the metals. Iron regulation module S6 and metalloproteins module S7. These two modules are in low expression, but their functions are very different. (1) Module S6 is to increase the iron content, making more iron ions combined with more oxygen, and provide a steady stream of energy for the proliferation and metastasis of liver cancer. (2) Metalloproteins in module S7 relates to tumor differentiation, and the content of which is low in liver cancer.

References

- [1] P. Yan, R. Li, J. Chen, Y. Li et al., "Appliction of random matrix theory to microay data for discovering functional gene modules of hepatocellular carcinoma," *Acta Biophysica Sinica*, vol. 25, no. 3, pp. 192–202, 2009.
- [2] C. Ling, Y. Lu, J. K. Kalsi et al., "Human hepatocyte growth factor receptor is a cellular coreceptor for adeno-associated virus serotype 3," *Human Gene Therapy*, vol. 21, no. 12, pp. 1741–1747, 2010.
- [3] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 12, pp. 7821–7826, 2002.
- [4] M. Newsman, "Fast algorithm for detecting community structure in networks," *Physical Review E*, vol. 69, no. 6, Article ID 066133, 2004.
- [5] X. G. Ruan and J. L. Wang, "Method for finding tumor functional modules using DNA microarray data," *Journal of Beijing University of Technology*, vol. 33, no. 4, pp. 366–371, 2007.
- [6] E. Newman and M. Grivan, "Finding and evaluating community structure in networks," *Physical Review E*, vol. 69, no. 2, Article ID 026113, 2004.
- [7] J. Scott, *Social Network Analysis: A Handbook*, London, UK, 2002.
- [8] X. Wang, *The Application of Multivariate Analysis*, Tsinghua University Press, 3rd edition, 2009.
- [9] L. Wang, G. Dai, and H. Zhao, "Research on modularity for evaluating community structure," *Computer Engineering*, vol. 36, no. 14, pp. 227–229, 2010.
- [10] X. Chen, S. T. Cheung, S. So et al., "Gene expression patterns in human liver cancers," *Molecular Biology of the Cell*, vol. 13, no. 6, pp. 1929–1939, 2002.
- [11] S. I. Kitajiri, T. Sakamoto, I. A. Belyantseva et al., "Actin-bundling protein TRIOBP forms resilient rootlets of hair cell stereocilia essential for hearing," *Cell*, vol. 141, no. 5, pp. 786–798, 2010.
- [12] C. Rimkus, J. Friederichs, A. L. Boulesteix et al., "Microarray-based prediction of tumor response to neoadjuvant radiochemotherapy of patients with locally advanced rectal cancer," *Clinical Gastroenterology and Hepatology*, vol. 6, no. 1, pp. 53–61, 2008.

- [13] A. Warth, T. Muley, M. Meister et al., "Loss of aquaporin-4 expression and putative function in non-small cell lung cancer," *BMC Cancer*, vol. 11, article 161, 2011.
- [14] S. Krizkova, I. Fabrik, V. Adam et al., "Metallothionein-a promising tool for cancer diagnostics," *Bratislavske Lekarske Listy*, vol. 110, pp. 429–447, 2009.
- [15] R. B. Bhavsar, L. N. Makley, and P. A. Tsonis, "The other lives of ribosomal proteins," *Human Genomics*, vol. 4, no. 5, pp. 327–344, 2010.
- [16] M. Shuda, N. Kondoh, K. Tanaka et al., "Enhanced expression of translation factor mRNAs in hepatocellular carcinoma," *Anticancer Research*, vol. 20, no. 4, pp. 2489–2494, 2000.
- [17] X. Luo, H. H. Hsiao, M. Bubunenko et al., "Structural and functional analysis of the E.coli NusB-S10 transcription antitermination complex," *Molecular Cell*, vol. 32, no. 6, pp. 791–802, 2008.
- [18] X. Guo, M. Chen, L. Ding et al., "Application of cox model in coagulation function in patients with primary liver cancer," *Hepato-Gastroenterology*, vol. 58, no. 106, pp. 326–330, 2011.
- [19] B. Gerlach, S. M. Cordier, A. C. Schmukle et al., "Linear ubiquitination prevents inflammation and regulates immune signalling," *Nature*, vol. 471, no. 7340, pp. 591–596, 2011.

Research Article

Multi-State Dependent Impulsive Control for Pest Management

Huidong Cheng, Fang Wang, and Tongqian Zhang

College of Science, Shandong University of Science and Technology, Qingdao 266510, China

Correspondence should be addressed to Huidong Cheng, chd900517@sdust.edu.cn

Received 17 April 2012; Accepted 6 June 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 Huidong Cheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

According to the integrated pest management strategies, we propose a model for pest control which adopts different control methods at different thresholds. By using differential equation geometry theory and the method of successor functions, we prove the existence of order one periodic solution of such system, and further, the attractiveness of the order one periodic solution by sequence convergence rules and qualitative analysis. Numerical simulations are carried out to illustrate the feasibility of our main results. Our results show that our method used in this paper is more efficient and easier than the existing ones for proving the existence of order one periodic solution.

1. Introduction

It is of great value to study pest management method applied in agricultural production; entomologists and the whole society have been paying close attention to how to control pests effectively and to save manpower and material resources. In agricultural production, pesticides-spraying (chemical control) and release of natural enemies (biological control) are the ways commonly used for pest control. But if we implement chemical control as soon as pests appear, many problems are caused: the first is environmental pollution; the second is increase of costs including human and material resources and time; the third is killing natural enemies, such as parasitic wasp; the last is pests' resistance to pesticides, which brings great negative effects instead of working as well as had been expected [1–3]. The second way, which controls pests with the help of the increasing natural enemies, can avoid problems caused by chemical control and gets more and more attention. So many scholars have been studying and discussing it [4–8]. Considering the effectiveness of the chemical control and nonpollution and limitations of the biological one, people have proposed the method of integrated pest management (IPM), which is a pest management system integrating all

appropriate ways and technologies to control economic injury level (EIL) caused by pest populations in view of population dynamics and its relevant environment. In the process of practical application, people usually implement the following two schemes for the integrated pest management: one is to implement control at a fixed time to eradicate pests [9, 10]; the other is to implement measures only when the amount of pests reaches a critical level, which is to make the amount less than certain economic impairment level, not to wipe out pests [11–13]. Salazar conducted an experiment of broad bean being damaged by bean sprouts worm in 1976 and found “crops’ compensation to damage of pests”, that is, yields of crops which had been damaged a little by pests in the early growth are actually higher than those without damage. In other words, we do not want to wipe out pests but to control them to a certain economic injury level (EIL). So, the second is used most in the process of agricultural industry. Tang and Cheke [14] first proposed the “Volterra” model in the form of a state-dependent impulsive model:

$$\begin{aligned} x'(t) &= x(t)(a - by(t)), & x \neq ET, \\ y'(t) &= y(t)(-d + cx(t)), \\ \Delta x(t) &= -\alpha x(t), & x = ET, \\ \Delta y(t) &= q, \end{aligned} \tag{1.1}$$

and they applied this model to pest management and proved existence and stability of periodic solution of first and second order. Then Tang and Cheke [14] also proposed bait-dependent digestive model with state pulse:

$$\begin{aligned} x'(t) &= x(t)(a - by(t)), \\ y'(t) &= y(t)\left(\frac{\lambda bx(t)}{1 + bhx(t)} - d\right), & x \neq h_{\max}, \\ \Delta x(t) &= -\alpha x(t), & x = h_{\max}, \\ \Delta y(t) &= q, \end{aligned} \tag{1.2}$$

they had the existence of positive periodic solution and stability of orbit. Recently Jiang and Lu et al. [15–17] have proposed pest management model with state pulse and phase structure and several predator-prey models with state pulse and had the existence of semi-trivial periodic solution and positive periodic solution and stability of orbit.

It is worth mentioning that the vast majority of research on population dynamics system with state pulse considers single state pulse, which is to say, only when the amount of population reaches the same economic threshold can measures be taken (e.g., chemical control and biological control); but this single state-pulse control does not confirm to reality. In fact, we often need to use different control methods under different states in real life. For example, in the process of pest management, when the amount of pests is small, biological control is implemented; when the amount is large, combination control is applied. Tang et al. [18] have investigated and developed a mathematical model with hybrid

impulsive model:

$$\begin{aligned}
 x'(t) &= rx(t)(1 - \delta x(t)) - bx(t)y(t), & x < ET, \\
 y'(t) &= y(t)(cx(t) - a), & t = \lambda_m, \\
 x(t^+) &= (1 - p_1)x(t), & x(t) = ET, \\
 y(t^+) &= (1 - p_2)y(t), & \\
 y(\lambda_m^+) &= (1 + p_3)y(\lambda_m) + q, & t = \lambda_m.
 \end{aligned} \tag{1.3}$$

Motivated by Tang, on the basis of the above analysis, we set up the following predator-prey system with different control methods in different thresholds:

$$\begin{aligned}
 x'(t) &= x(t)(a - by(t)), \\
 y'(t) &= y(t) \left(\frac{\lambda bx(t)}{1 + b\lambda x(t)} - d \right), & x \neq h_1, h_2 \text{ or } x = h_1, y > y^*, \\
 \Delta x(t) &= 0, \\
 \Delta y(t) &= \delta, & x = h_1, y \leq y^*, \\
 \Delta x(t) &= -\alpha x(t), \\
 \Delta y(t) &= -\beta y(t) + q, & x = h_2,
 \end{aligned} \tag{1.4}$$

where $x(t)$ and $y(t)$ represent, respectively, the prey and the predator population densities at time t ; a, b, λ, h_1, h_2 and d are all positive constants and $h_1 < h_2$; $y^* = a/b$. $\alpha, \beta \in (0, 1)$ represent the fraction of pest and predator, respectively, which die due to the pesticide when the amount of prey reaches economic threshold h_2 and q is the release amount of predator. $\lambda bx(t)/(1 + b\lambda x(t))$ is the per capita functional response of the predator. When the amount of the prey reaches the threshold h_1 at time t_{h_1} , controlling measures are taken (releasing natural enemies) and the amount of predator abruptly turns to $y(t_{h_1}) + \delta$. When the amount of the prey reaches the threshold h_2 at time t_{h_2} , spraying pesticide, and releasing natural enemies and the amount of prey and predator abruptly turn to $(1 - \alpha)x(t_{h_2})$ and $(1 - \beta)y(t_{h_2}) + q$, respectively. Refer to [17] Liu et al. for details.

2. Preliminaries

First, we give some basic definitions and lemmas.

Definition 2.1. A triple (X, Π, R^+) is said to be a semidynamical system if X is a metric space, R^+ is the set of all nonnegative real, and $\Pi(P, t) : X \times R^+ \rightarrow X$ is a continuous map such that:

- (i) $\Pi(P, 0) = P$ for all $P \in X$;
- (ii) $\Pi(P, t)$ is continuous for t and s ;
- (iii) $\Pi(\Pi(P, t)) = \Pi(P, t + s)$ for all $P \in X$ and $t, s \in R^+$. Sometimes a semi-dynamical system (X, Π, R^+) is denoted by (X, Π) .

Definition 2.2. Assuming that

- (i) (X, Π) is a semi-dynamical system;
- (ii) M is a nonempty subset of X ;
- (iii) function $I : M \rightarrow X$ is continuous and for any $P \in M$, there exists a $\varepsilon > 0$ such that for any $0 < |t| < \varepsilon$, $\Pi(P, t) \notin M$.

Then, (X, Π, M, I) is called an impulsive semi-dynamical system.

For any P , the function $\Pi_P : \mathbb{R}^+ \rightarrow X$ defined as $\Pi_P(t) = \Pi(P, t)$ is continuous, and we call $\Pi_P(t)$ the trajectory passing through point P . The set $C^+(P) = \{\Pi(P, t) / 0 \leq t < +\infty\}$ is called positive semitrajectory of point P . The set $C^-(P) = \{\Pi(P, t) / -\infty < t \leq 0\}$ is called the negative semi-trajectory of point P .

Definition 2.3. One considers state-dependent impulsive differential equations:

$$\begin{aligned} x'(t) &= P(x, y), & (x, y) \notin M(x, y), \\ y'(t) &= Q(x, y), \\ \Delta x(t) &= \alpha(x, y), & (x, y) \in M(x, y), \\ \Delta y(t) &= \beta(x, y), \end{aligned} \tag{2.1}$$

where $M(x, y)$ and $N(x, y)$ represent the straight line or curve line on the plane, $M(x, y)$ is called impulsive set. The function I is continuous mapping, $I(M) = N$, I is called the impulse function. $N(x, y)$ is called the phase set. We define “dynamic system” constituted by the definition of solution of state impulsive differential equation (2.1) as “semicontinuous dynamic systems”, which is denoted as (Ω, f, I, M) .

Definition 2.4. Suppose that the impulse set M and the phase set N are both lines, as shown in Figure 1. Define the coordinate in the phase set N as follows: denote the point of intersection Q between N and x -axis as O , then the coordinate of any point A in N is defined as the distance between A and Q and is denoted by y_A . Let C denote the point of intersection between the trajectory starting from A and the impulse set M , and let B denote the phase point of C after impulse with coordinate y_B . Then, we define B as the successor point of A , and then the successor function of point A is that $f(A) = y_B - y_A$.

Definition 2.5. A trajectory $\tilde{\Pi}(P_0, t)$ is called order one periodic solution with period T if there exists a point $P_0 \in N$ and $T > 0$ such that $P = \Pi(P_0, t) \in M$ and $P^+ = I(P) = P_0$.

We get these lemmas from the continuity of composite function and the property of continuous function.

Lemma 2.6. Successor function defined in Definition 2.1 is continuous.

Lemma 2.7. In system (1.4), if there exist $A \in N$, $B \in N$ satisfying successor function $f(A)f(B) < 0$, then there must exist a point P ($P \in N$) satisfying $f(P) = 0$ the function between the point of A and the point of B , thus there is an order one periodic solution in system (1.4).

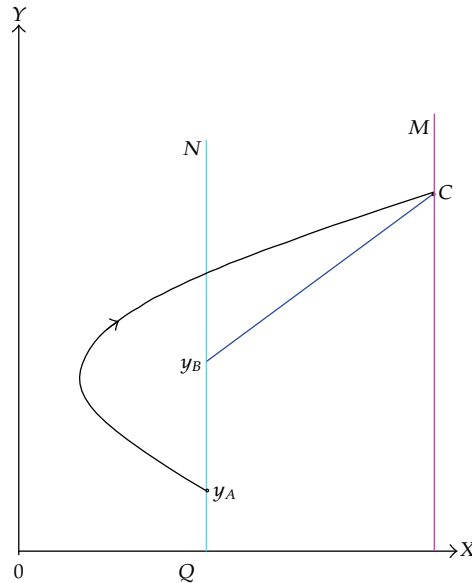


Figure 1: Successor function defined.

Next, we consider the model (1.4) without impulse effects:

$$\begin{aligned} x'(t) &= x(t)(a - by(t)), \\ y'(t) &= y(t)\left(\frac{\lambda bx(t)}{1 + bhx(t)} - d\right). \end{aligned} \quad (2.2)$$

It is well known that the system (2.2) possesses

- (I) two steady states $O(0,0)$ -saddle point, and $R(d/b(\lambda - dh), a/b) = R(x^*, y^*)$ ($\lambda > dh$)-stable centre;
- (II) a unique closed trajectory through any point in the first quadrant contained inside the point R .

In this paper, we assume that the condition $\lambda > dh$ holds. By the biological background of system (1.4), we only consider $D = \{(x, y) : x \geq 0, y \geq 0\}$. Vector graph of system (2.2) can be seen in Figure 2.

This paper is organized as follows. In the next section, we present some basic definitions and an important lemmas as preliminaries. In Section 3, we prove existence for an order one periodic solution of system (1.4). The sufficient conditions for the attractiveness of order one periodic solutions of system (1.4) are obtained in Section 4. At last, we state conclusion, and the main results are carried out to illustrate the feasibility by numerical simulations.

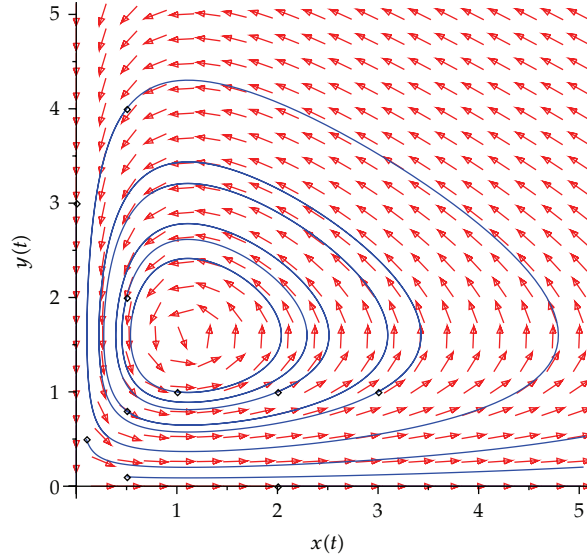


Figure 2: Illustration of vector graph of system (2.2).

3. Existence of the Periodic Solution

In this section, we will investigate the existence of an order one periodic solution of system (1.4) by using the successor function defined in this paper and qualitative analysis. For this goal, we denote that $M_1 = \{(x, y) / x = h_1, 0 \leq y \leq a/b\}$, and that $M_2 = \{(x, y) \mid x = h_2, y \geq 0\}$. Phase set of set M is that $N_1 = I(M_1) = \{(x, y) \mid x = h_1, a/b < y \leq (a/b) + \delta\}$ and that $N_2 = I(M_2) = \{(x, y) \mid x = (1 - \alpha)h_2, y \geq q\}$. Isoclinic line is denoted, respectively, by lines: $L_1 = \{(x, y) \mid y = a/b, x \geq 0\}$ and $L_2 = \{(x, y) \mid x = d/b(\lambda - dh), y \geq 0\}$.

For the convenience, if $P \in \Omega - M$, $F(P)$ is defined as the first point of intersection of $C^+(P)$ and M , that is, there exists a $t_1 \in R_+$ such that $F(P) = \Pi(P, t_1) \in M$, and for $0 < t < t_1$, $\Pi(P, t) \notin M$; if $B \in N$, $R(B)$ is defined as the first point of intersection of $C^-(P)$ and N , that is, there exists a $t_2 \in R_+$ such that $R(B) = \Pi(B, -t_2) \in N$, and for $-t < t < 0$, $\Pi(B, t) \notin N$. For any point P , we denote y_P as its ordinate. If the point $P(h, y_P) \in M$, then pulse occurs at the point P , the impulsive function transfers the point P into $P^+ \in N$. Without loss of generality, unless otherwise specified we assume the initial point of the trajectory lies in phase set N .

Due to the practical significance, in this paper we assume the set always lies in the left side of stable centre R , that is, $h_1 < d/b(\lambda - dh)$ and $(1 - \alpha)h_2 < d/b(\lambda - dh)$.

In the light of the different position of the set N_1 and the set N_2 , we consider the following three cases.

Case 1 ($0 < h_1 < d/b(\lambda - dh)$). In this case, set M_1 and N_1 are both in the left side of stable centre R (as shown in Figure 3). Take a point $B_1(h_1, (a/b) + \varepsilon) \in N_1$ above A , where $\varepsilon > 0$ is small enough, then there must exist a trajectory passing through B_1 which intersects with M_1 at point $P_1(h_1, y_{p_1})$, we have $y_{p_1} < a/b$. Since $p_1 \in M_1$, pulse occurs at the point P_1 , the impulsive function transfers the point P_1 into $P_1^+(h_1, y_{p_1} + \delta)$ and P_1^+ must lie above B_1 , therefore inequation $(a/b) + \varepsilon < y_{p_1} + \delta$ holds, thus the successor function of B_1 is $f(B_1) = y_{p_1} + \delta - ((a/b) + \varepsilon) > 0$.

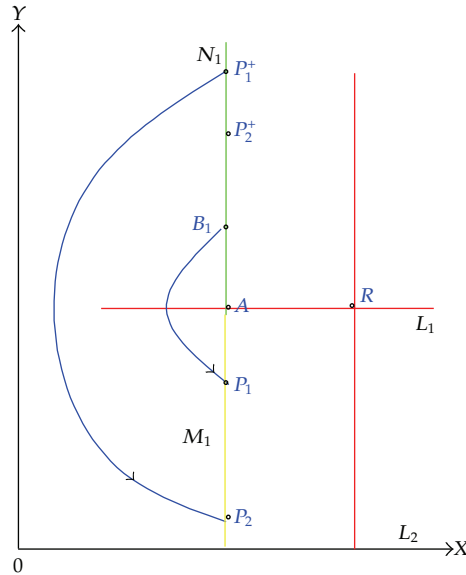


Figure 3: $0 < h_1 < (d/b(1 - dh))(a/b) < y_{p_2} + \delta < y_{p_1} + \delta$.

On the other hand, the trajectory with the initial point P_1^+ intersects M_1 at point $P_2(h_1, y_{p_2})$, in view of vector field and disjointness of any two trajectories, we know $y_{p_2} < y_{p_1} < a/b$. Supposing the point P_2 is subject to impulsive effects to point $P_2^+(h_1, y_{p_2}^+)$, where $y_{p_2}^+ = y_{p_2} + \delta$, the position of P_2^+ has the following two cases.

Subcase 1.1 ($a/b < y_{p_2} + \delta < y_{p_1} + \delta$). In this case, the point P_2^+ lies above the point A and below P_1^+ , then we have $f(P_1^+) = y_{p_2} + \delta - (y_{p_1} + \delta) < 0$.

By Lemma 2.7, there exists an order one periodic solution of system (1.4), whose initial point is between B_1 and P_1^+ in set N_1 .

Subcase 1.2 ($a/b \geq y_{p_2} + \delta$ (as shown in Figure 4)). The point P_2^+ lies below the point A , that is, $P_2^+ \in M_1$, then pulse occurs at the point P_2^+ , the impulsive function transfers the point P_2^+ into $P_2^{++}(h_1, y_{p_2} + 2\delta)$.

If $a/b < y_{p_2} + 2\delta < y_{p_1} + \delta$, like the analysis of Subcase 1.1, there exists an order one periodic solution of system (1.4).

If $a/b > y_{p_2} + 2\delta$, that is, $P_2^{++} \in M_1$, then we repeat the above process until there exists $k \in \mathbb{Z}_+$ such that P_2^{++} jumps to $P_2^{i+}((h_1, y_{p_2} + (k+2)\delta))$ after k times' impulsive effects which satisfies $a/b < y_{p_2} + (k+2)\delta < y_{p_1} + \delta$. Like the analysis of Subcase 1.1, there exists an order one periodic solution of system (1.4).

Now, we can summarize the above results as the following theorem.

Theorem 3.1. *If $\lambda > dh, 0 < h_1 < d/b(\lambda - dh)$, then there exists an order one periodic solution of the system (1.4).*

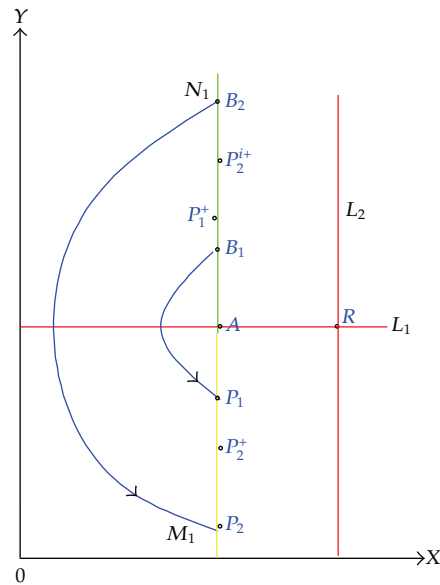


Figure 4: $0 < h_1 < (d/b(1 - dh))(a/b) > y_{P_2} + \delta$.

Remark 3.2. It shows from the proved process of Theorem 3.1 that the number of natural enemies should be selected appropriately, which aims to reduce releasing impulsive times to save manpower and resources.

Case 2 ($h_2 < d/b(\lambda - dh)$). In this case, set M_2 and N_2 are both in the left side of stable center R , in the light of the different position of the set N_2 , we consider the following two cases.

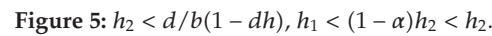
Subcase 2.1 ($0 < h_1 < (1 - \alpha)h_2 < h_2 < d/b(\lambda - dh)$). In this case, the set N_2 is in the right side of M_1 (as shown in Figure 5). The trajectory passing through point A which tangents to N_2 at point A intersects with M_2 at point $P_0(h_2, y_{P_0})$. Since the point $P_0 \in M_2$, then impulse occurs at point P_0 . Supposing the point P_0 is subject to impulsive effects to point $P_0^+((1 - \alpha)h_2, y_{P_0}^+)$, where $y_{P_0}^+ = (1 - \beta)y_{P_0} + q$, the position of P_0^+ has the following three cases:

(1) $((1 - \beta)y_{P_0} + q > a/b)$. Take a point $B_1((1 - \alpha)h_2, \varepsilon + a/b) \in N_2$ above A , where $\varepsilon > 0$ is small enough. Then there must exist a trajectory passing through the point B_1 which intersects with the set M_2 at point $P_1(h_2, y_{P_1})$. In view of continuous dependence of the solution on initial value and time, we know $y_{P_1} < y_{P_0}$ and the point P_1 is close to P_0 enough, so we have the point P_1^+ is close to P_0^+ enough and $y_{P_1^+} < y_{P_0^+}$, then we obtain $f(B_1) = y_{P_1^+} - y_{B_1} > 0$.

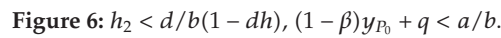
On the other hand, the trajectory passing through point B tangents to N_1 at point B . Set $F(S) = P_2(h_2, y_{P_2}) \in M_2$. Denote the coordinates of impulsive point $P_2^+((1 - \alpha)h_2, y_{P_2}^+)$ corresponding to the point $P_2(h_2, y_{P_2})$.

If $y_S \geq y_{P_0^+}$ then $y_{P_2^+} < y_{P_0^+} < y_S$. So we obtain $f(S) = y_{P_2^+} - y_S < 0$. There exists an order one periodic solution of system (1.4), whose initial point is between the point B_1 and the point S in set N_2 (Figure 5).

If $y_S < y_{P_0^+}$ and $y_{P_2^+} \leq y_S$, we have $f(S) = y_{P_2^+} - y_S \leq 0$, we conclude that there exists an order one periodic solution of system (1.4).



If $y_{P_0^+} > y_A$, $y_S < y_{P_0^+}$ and $y_S < y_{P_2^+}$, there is no an order one periodic solutions of the system (1.4). The trajectory of system (1.4) with any initiating point on the N_2 will ultimately stay in $\Omega_1 = \{(x, y) / 0 \leq x \leq h_1, y \geq 0\}$ after one impulsive effect.



If $(1 - \beta)y_{P_1} + q < a/b$ (Figure 7), the point P_1^+ lies below the point B . Like the analysis of Subcase 2.1(2), we can prove there exists an order one periodic solution to the system (1.4) in this case.

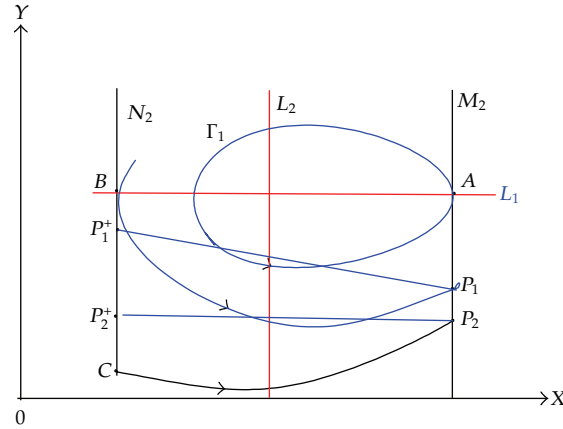


Figure 7: $h_1 < (1 - \alpha)h_2 < \delta_{\min} < d/(\lambda - hd)(1 - \beta)y_{P_1} + q < a/b$.

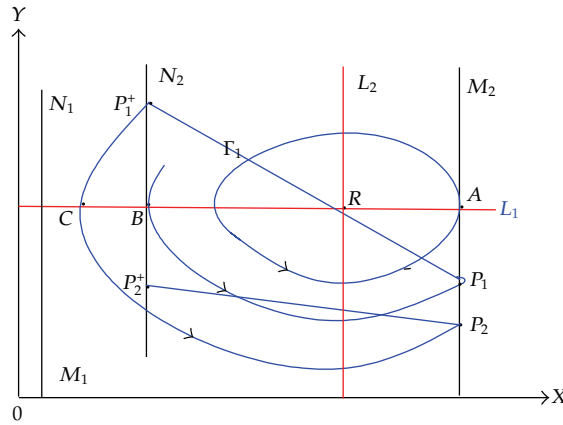


Figure 8: $h_1 < (1 - \alpha)h_2 < \delta_{\min} < d/(\lambda - hd)(1 - \beta)y_{P_1} + q > a/b$.

If $(1 - \beta)y_{P_1} + q > a/b$, the point P_1^+ lies above the point B ; the trajectory from initiating point P_1^+ intersects with the line L_1 at point C . If $h_1 \leq y_C$ (Figure 8), we have $y_{P_1} > y_{P_2}$ and $y_{P_1^+} > y_{P_2^+}$, then the successor function of P_1^+ is that $f(P_1^+) = y_{P_2^+} - y_{P_1^+} < 0$. Then, we know that there exists an order one periodic solution of system (1.4), whose initial point is between the point P_1^+ and B in set N_2 . If $h_1 > y_C$ (Figure 9), there is a trajectory which is tangent to the N_1 at a point D intersects with M_2 at a point $P_3(h_2, y_{P_3}) \in M_2$, P_3 jumps to P_3^+ after the impulsive effects. If $y_{P_3^+} \leq y_{B_1}$, we can easily know that there exists an order one periodic solution of system (1.4). If $y_{P_3^+} > y_{B_1}$, by the qualitative analysis of the system (1.4), we know that trajectory with any initiating point on the N_2 will ultimately stay in Γ_1 after a finite number of impulsive effects.

If $(1 - \beta)y_{P_1} + q = a/b$, the point P_1^+ coincides with the point B , and the successor function of the point B is that $f(B) = 0$; then there exists an order one periodic solution which is just a part of the trajectory passing through the point B .

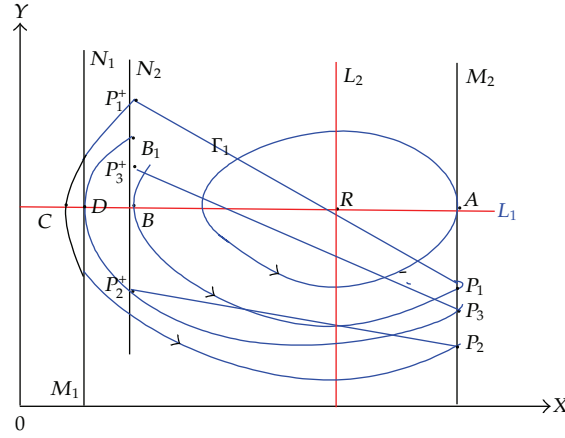


Figure 9: $h_1 < (1 - \alpha)h_2 < \delta_{\min} < d/(\lambda - hd) < h_2(1 - \beta)y_{P_1} + q > a/b, h_1 > y_c$.

Now, we can summarize the above results as the following theorem.

Theorem 3.5. Assuming that $\lambda > dh$ and $h_1 < (1 - \alpha)h_2 < \delta_{\min} < d/b(\lambda - dh) < h_2$.

If $(1 - \beta)y_{P_1} + q \leq a/b$, there exists an order one periodic solution to the system (1.4).

If $(1 - \beta)y_{P_1} + q > a/b$ and $h_1 \leq y_c$, then there exists an order one periodic solution to the system (1.4).

If $(1 - \beta)y_{P_1} + q > a/b$, $h_1 > y_c$ and $y_{P_3^+} \leq y_{B_1}$, then there exists an order one periodic solution to the system (1.4).

(2) ($h_1 < \delta_{\min} < (1 - \alpha)h_2 < d/b(\lambda - dh) < h_2$). In this case, denote the closed trajectory Γ_1 of system (1.4) intersects with N_2 two points $A_1((1 - \alpha)h_2, y_{A_1})$ and $A_2((1 - \alpha)h_2, y_{A_2})$ (as shown in Figure 10). Since $A \in M_2$, impulse occurs at the point A . Suppose point A is subject to impulsive effects to point $P_0^+((1 - \alpha)h_2, y_{P_0^+})$, here $y_{P_0^+} = (1 - \beta)(a/b) + q$.

If $(1 - \beta)(a/b) + q = y_{A_1}$ or $(1 - \beta)(a/b) + q = y_{A_2}$, then P_0^+ coincides with A_1 or P_0^+ coincides with A_2 , and the successor function of A_1 or A_2 is that $f(A_1) = 0$ or $f(A_2) = 0$. So, there exists an order one periodic solution of system (1.4) which is just a part of the trajectory Γ_1 .

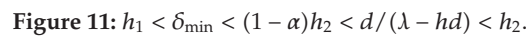
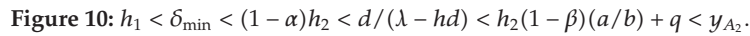
If $(1 - \beta)(a/b) + q < y_{A_2}$, the point P_0^+ lies below the point A_2 . Like the analysis of Subcase 2.1(2), we can prove there exists an order one periodic solution to the system (1.4) in this case.

If $(1 - \beta)(a/b) + q > y_{A_1}$ (as shown in Figure 11), the point P_0^+ is above the point A_1 . Like the analysis of Subcase 3.1(1), we obtain sufficient conditions of existence of order one periodic solution to the system (1.4).

Theorem 3.6. Assuming that $\lambda > dh, h_1 < (1 - \alpha)h_2 < \delta_{\min} < d/b(\lambda - dh) < h_2$.

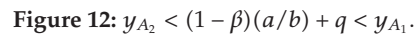
If $(1 - \beta)(a/b) + q \leq y_{A_2}$, there exists an order one periodic solution to the system (1.4).

If $(1 - \beta)(a/b) + q > y_{A_1}$ and $h_1 \leq y_c$, then there exists an order one periodic solution to the system (1.4).



(3) ($y_{A_2} < (1-\beta)(a/b)+q < y_{A_1}$). In this case, we note that the point P_0^+ must lie between the point A_1 and the point A_2 (As shown in Figure 12). Taking a point $B_1 \in M_2$ such that B_1 jumps to A_2 after the impulsive effect, denote $A_2 = B_1^+$. Since $y_{P_0^+} > y_{B_1^+}$, we have $y_A > y_{B_1}$. Let $R(B_1) = B_2^+ \in N_2$, take a point $B_2 \in M_2$ such that B_2 jumps to B_2^+ after the impulsive effects, then we have $y_{B_1^+} > y_{B_2^+}$, $y_{B_1} > y_{B_2}$. This process continues until there exists a $B_K^+ \in N_2$ ($K \in \mathbb{Z}_+$) satisfying $y_{B_K^+} < q$. So we obtain a sequence $\{B_k^+\}_{k=1,2,\dots,K}$ of the set M_2 and a sequence $\{B_k\}_{k=1,2,\dots,K}$ of set N_2 satisfying $R(B_{k-1}) = B_k^+ \in N_2$, $y_{B_{k-1}^+} > y_{B_k^+}$. In the following, we will prove the trajectory of system (1.4) with any initiating point of set N_2 will ultimately stay in Γ_1 .

From the vector field of system (1.4), we know the trajectory of system (1.4) with any initiating point between the point A_1 and A_2 will be free from impulsive effect and ultimately will stay in Γ_1 .



Denote the intersection of the trajectory passing through the point B which tangents to N_1 at point B with the set N_2 at $S((1-\alpha)h_2, y_S)$. With time increasing, the trajectory of system (1.4) from any initiating point on segment $\overline{A_1S}$ intersect with the set N_2 at some point which is below A_2 ; so just like the analysis above we obtain, it will ultimately stay in Γ_1 . So for any point below S , will ultimately stay in region Γ_1 with time increasing.

Now, we can summarize the above results as the following theorem.

Subcase 3.2 ($0 < (1 - \alpha)h_2 < h_1 < d/b(\lambda - dh) < h_2$). In this case, the set N_2 is on the left side of the set N_1 and M_2 in the right side of R . Like the analysis of Subcase 2.2, we can know that any trajectory with initial point $(x_0^+, y_0^+) \in \Omega_0 = \{(x, y) \mid x \geq 0, y \geq 0, x \leq h_2\}$ will stay in the region $\Omega_1 = \{(x, y) \mid x \geq 0, y \geq 0, x \leq h_1\}$ after one impulsive effect or free from impulsive effect.

In this section, under the condition of existence of order one periodic solution to system (1.4) and the initial value of pest population $x(0) \leq h_2$, we discuss its attractiveness. We focus on Case 1 and Case 2; by similar method, we can obtain similar results about Case 3.

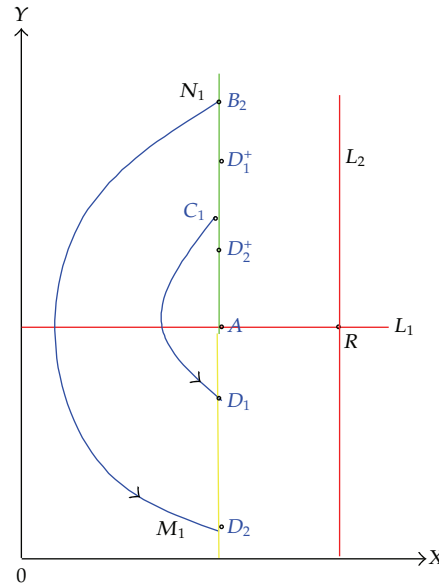


Figure 13: There is a unique order one periodic solution (Theorem 4.1).

Theorem 4.1. Assuming that $\lambda > dh, h_1 < d/(b(\lambda - dh))$ and $\delta \geq a/b$.

If $y_{P_0^+} > y_{P_2^+} > y_{P^+}$ or $y_{P_0^+} < y_{P_2^+} < y_{P^+}$ (Figure 14), then

(I) there exists a unique order one periodic solution of system (1.4),

(II) if $(1 - \alpha)h_2 < h_1$, order one periodic solution of system (1.4) is attractive in the region $\Omega_0 = \{(x, y) \mid x \geq 0, y \geq 0, x \leq h_2\}$.

Proof. By the derivation of Theorem 3.1, we know there exists an order one periodic solution of system (1.4). We assume trajectory $\widehat{PP^+}$ and segment $\widehat{PP^+}$ formulate an order one periodic solution of system (1.4), that is, there exists a $P^+ \in N_2$ such that the successor function of P^+ satisfies $f(P^+) = 0$. First, we will prove the uniqueness of the order one periodic solution.

We take any two points $C_1(h_1, y_{C_1}) \in N_1, C_2(h_1, y_{C_2}) \in N_1$ satisfying $y_{C_2} > y_{C_1} > y_A$, then we obtain two trajectories whose initiate points are C_1 and C_2 intersects the set M_1 two points $D_1(h_1, y_{D_1})$ and $D_2(h_1, y_{D_2})$, respectively, (Figure 13). In view of the vector field of system (1.4) and the disjointness of any two trajectories without impulse, we know $y_{D_1} > y_{D_2}$. Suppose the points D_1 and D_2 are subject to impulsive effect to points $D_1^+(h_1, y_{D_1^+})$ and $D_2^+(h_2, y_{D_2^+})$, respectively, then we have $y_{D_1^+} > y_{D_2^+}$ and $f(C_1) = y_{D_1^+} - y_{C_1}, f(C_2) = y_{D_2^+} - y_{C_2}$, so we get $f(C_1) - f(C_2) < 0$, thus we obtain the successor function $f(x)$ is decreasing monotonously of N_1 , so there is a unique point $P^+ \in N_1$ satisfying $f(P^+) = 0$, and the trajectory $\widehat{P^+PP^+}$ is a unique order one periodic solution of system (1.4).

Next, we prove the attractiveness of the order one periodic solution $\widehat{P^+PP^+}$ in the region $\Omega_0 = \{(x, y) \mid x \geq 0, y \geq 0, x \leq h_2\}$. We focus on the case $y_{P_0^+} > y_{P_2^+} > y_{P^+}$; by similar method, we can obtain similar results about case $y_{P_0^+} < y_{P_2^+} < y_{P^+}$ (Figure 14).

Take any point $P_0^+(h_1, y_{P_0^+}) \in N_1$ above P^+ . Denote the first intersection point of the trajectory from initiating point $P_0^+(h_1, y_{P_0^+})$ with the set M_1 at $P_1(h_1, y_{P_1})$, and the corresponding consecutive points are $P_2(h_1, y_{P_2}), P_3(h_1, y_{P_3}), P_4(h_1, y_{P_4}), \dots$, respectively.

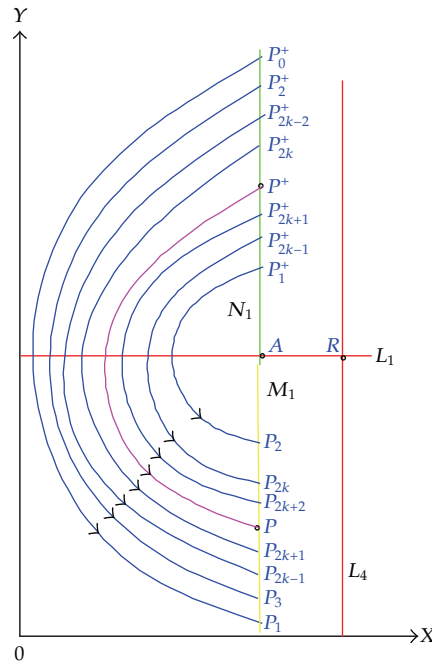


Figure 14: Order one periodic solution is attractive (Theorem 4.1).

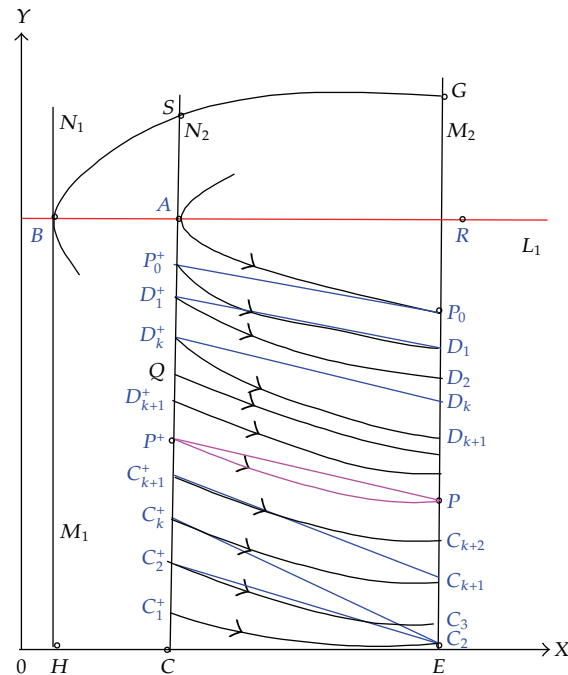
Consequently, under the effect of impulsive function, the corresponding points after pulse are $P_1^+(h_1, y_{P_1^+})$, $P_2^+(h_1, y_{P_2^+})$, $P_3^+(h_1, y_{P_3^+})$, \dots .

Due to conditions $y_{P_0^+} > y_{P_2^+} > y_{P^+}$, $y_{P_k^+} = y_{P_k} + \delta$, $\delta \geq a/b$ and disjointness of any two trajectories, then we get a sequence $\{P_k^+\}_{k=1,2,\dots}$ of the set N_1 satisfying

$$y_{P_1^+} < y_{P_3^+} < \dots < y_{P_{2k-1}^+} < y_{P_{2k+1}^+} < \dots < y_{P^+} < \dots < y_{P_{2k}^+} < y_{P_{2k-2}^+} < \dots < y_{P_2^+} < y_{P_0^+}. \quad (4.1)$$

So the successor function $f(P_{2k-1}^+) = y_{P_{2k}^+} - y_{P_{2k-1}^+} > 0$ and $f(P_{2k}^+) = y_{P_{2k+1}^+} - y_{P_{2k}^+} < 0$ hold. Series $\{y_{P_{2k-1}^+}\}_{k=1,2,\dots}$ increases monotonously and has upper bound, so $\lim_{k \rightarrow \infty} y_{P_{2k-1}^+}$ exists. Next, we will prove $\lim_{k \rightarrow \infty} y_{P_{2k-1}^+} = y_{P^+}$. Set $\lim_{k \rightarrow \infty} P_{2k-1} = C^+$, we will prove $P^+ = C^+$. Otherwise $P^+ \neq C^+$, then there is a trajectory passing through the point C^+ which intersects the set M_1 at point \tilde{C} , then we have $y_{\tilde{C}} > y_P$, $y_{\tilde{C}^+} > y_{P^+}$. Since $f(C^+) \geq 0$ and $P^+ \neq C^+$, according to the uniqueness of the periodic solution, then we have $f(C^+) = y_{\tilde{C}^+} - y_{C^+} > 0$, thus $y_{C^+} < y_{P^+} < y_{\tilde{C}^+}$ hold. Analogously, let trajectory passing through the point C^+ which intersects the set M_1 at point $\tilde{\tilde{C}}$, and the corresponding consecutive points is $\tilde{\tilde{C}}$, then $y_{\tilde{\tilde{C}}} > y_{\tilde{C}} > y_P > y_{\tilde{\tilde{C}}^+} > y_{\tilde{C}^+} > y_{\tilde{\tilde{C}}^+} > y_{P^+} > y_{C^+}$, then we have $f(\tilde{\tilde{C}}^+) = y_{\tilde{\tilde{C}}^+} - y_{\tilde{\tilde{C}}} > 0$, this is, contradict to the fact that C^+ is a limit of sequence $\{P_{2k-1}^+\}_{k=1,2,\dots}$, so we obtain $P^+ = C^+$. So, we obtain $\lim_{k \rightarrow \infty} y_{P_{2k-1}^+} = y_{P^+}$. Similarly, we can prove $\lim_{k \rightarrow \infty} y_{P_{2k}^+} = y_{P^+}$.

From above analysis, we know there exists a unique order one periodic solution in system (1.4), and the trajectory from initiating any point of the N_1 will ultimately tend to be order one periodic solution $\widehat{P^+ P P^+}$.



(II) By the derivation of Theorem 3.3, we know there exists an order one periodic solution of system (1.4) whose initial point is between C_1^+ and P_0^+ in the set N_2 . Assume

trajectory $\widehat{P^+P}$ and segment $\widehat{PP^+}$ formulate the unique order one periodic solution of system (1.4) with initial point $P^+ \in N_2$.

On the one hand, take a point $C_1^+((1-\alpha)h_2, y_{C_1^+}) \in N_2$ satisfying $y_{C_1^+} = \varepsilon < q$ and $y_{C_1^+} < y_{P^+}$. The trajectory passing through the point $C_1^+((1-\alpha)h_2, \varepsilon)$ which intersects with the set M_2 at point $C_2(h_2, y_{C_2})$, that is, $F(C_1^+) = C_2 \in M_2$, then we have $y_{C_2} < y_P$, thus $y_{C_2} < y_{P^+}$, since $y_{C_2^+} = (1-\beta)y_{C_2} + q > \varepsilon$. So, we obtain $f(C_1^+) = y_{C_2^+} - y_{C_1^+} = y_{C_2^+} - \varepsilon > 0$; Set $F(C_2^+) = C_3 \in M_2$, because $y_{C_1^+} < y_{C_2^+} < y_{P^+}$, we know $y_{C_2} < y_{C_3} < y_P$, then we have $y_{C_2^+} < y_{C_3^+} < y_{P^+}$ and $f(C_2^+) = y_{C_3^+} - y_{C_2^+} > 0$. This process is continuing, then we get a sequence $\{C_k^+\}_{k=1,2,\dots}$ of the set N_2 satisfying

$$y_{C_1^+} < y_{C_2^+} < \dots < y_{C_k^+} < \dots < y_{P^+} \quad (4.2)$$

and $f(C_k^+) = y_{C_{k+1}^+} - y_{C_k^+} > 0$. Series $\{y_{C_k^+}\}_{k=1,2,\dots}$ increase monotonously and have upper bound, so $\lim_{k \rightarrow \infty} y_{C_k^+}$ exists. Like the proof of Theorem 4.1, we can prove $\lim_{k \rightarrow \infty} y_{C_k^+} = y_{P^+}$.

On the other hand, set $F(P_0^+) = D_1 \in M_2$, then D_1 jumps to $D_1^+ \in N_2$ under the impulsive effects. Since $y_{P^+} < y_{P_0^+} < y_A$, we have $y_P < y_{D_1} < y_{P_0}$, thus we obtain $y_{P^+} < y_{D_1^+} < y_{P_0^+}$, $f(P_0^+) = y_{D_1^+} - y_{P_0^+} < 0$. Set $F(D_1^+) = D_2 \in M_2$, then D_2 jumps to $D_2^+ \in N_2$ under the impulsive effects. We have $y_{P^+} < y_{D_2^+} < y_{D_1^+}$; this process is continuing, we can obtain a sequence $\{D_k^+\}_{k=1,2,\dots}$ of the set N_2 satisfying

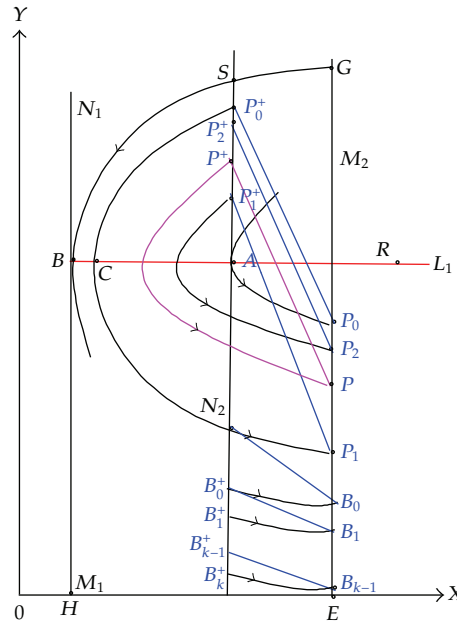
$$y_{P_0^+} > y_{D_1^+} > y_{D_2^+} > \dots > y_{D_k^+} > \dots > y_{P^+} \quad (4.3)$$

and $f(D_k^+) = y_{D_{k+1}^+} - y_{D_k^+} < 0$. Series $\{y_{D_k^+}\}_{k=1,2,\dots}$ decreases monotonously and has lower bound, so $\lim_{k \rightarrow \infty} y_{D_k^+}$ exists. Similarly, we can prove $\lim_{k \rightarrow \infty} y_{D_k^+} = y_{P^+}$.

Any point $Q \in N_2$ below A must be in some interval $[y_{D_{k+1}^+}, y_{D_k^+}]_{k=1,2,\dots}$, $[y_{D_1^+}, y_{P_0^+}]$, $[y_{P_0^+}, y_A]$, $[y_{C_k^+}, y_{C_{k+1}^+}]_{k=1,2,\dots}$. Without loss of generality, we assume the point $Q \in [y_{D_{k+1}^+}, y_{D_k^+}]$. The trajectory with initiating point Q moves between trajectory $\widehat{D_k^+D_{k+1}}$ and $\widehat{D_{k+1}^+D_{k+2}}$ and intersects with M_2 at some point between D_{k+2} and D_{k+1} ; under the impulsive effects, it jumps to the point of N_2 which is between $[y_{D_{k+2}^+}, y_{D_{k+1}^+}]$, then trajectory $\tilde{\Pi}(Q, t)$ continues to move between trajectory $\widehat{D_{k+1}^+D_{k+2}}$ and $\widehat{D_{k+2}^+D_{k+3}}$. This process can be continued unlimitedly. Since $\lim_{k \rightarrow \infty} y_{D_k^+} = y_{P^+}$, the intersection sequence of trajectory $\tilde{\Pi}(Q, t)$, and the set N_2 will ultimately tend to the point P^+ . Similarly, if $Q \in [y_{C_k^+}, y_{C_{k+1}^+}]$, we also can get the intersection sequence of trajectory $\tilde{\Pi}(Q, t)$ and the set N_2 will ultimately tend to point P^+ . Thus, the trajectory initiating any point below A ultimately tend to the unique order one periodic solution $\widehat{P^+PP^+}$.

Denote the intersection of the trajectory passing through the point B which tangents to N_1 at the point B with the set N_2 at a point $S((1-\alpha)h_2, y_S)$. The trajectory from any initiating point on segment \overline{AS} will intersect with the set N_2 at some point below A with time increasing. So like the analysis above, we obtain the trajectory from any initiating point on segment \overline{AS} will ultimately tend to be the unique order one periodic solution $\widehat{P^+PP^+}$.

Since the trajectory with any initiating point of the Ω_2 will certainly intersect with the set N_2 , then from the above analysis, we know the trajectory with any initiating point on segment \overline{AS} will ultimately tend to be order one periodic solution $\widehat{P^+PP^+}$. Therefore, the unique order one periodic solution $\widehat{P^+PP^+}$ is attractive in the region Ω_2 . This completes the proof. \square



So $f(P_{2k-1}^+) = y_{2k}^{P^+} - y_{2k-1}^{P^+} > 0$ and $f(P_{2k}^+) = y_{2k+1}^{P^+} - y_{2k}^{P^+} < 0$ hold. Like the proof of Theorem 4.1, we can prove $\lim_{k \rightarrow \infty} y_{2k-1}^{P^+} = \lim_{k \rightarrow \infty} y_{2k}^{P^+} = y^{P^+}$.

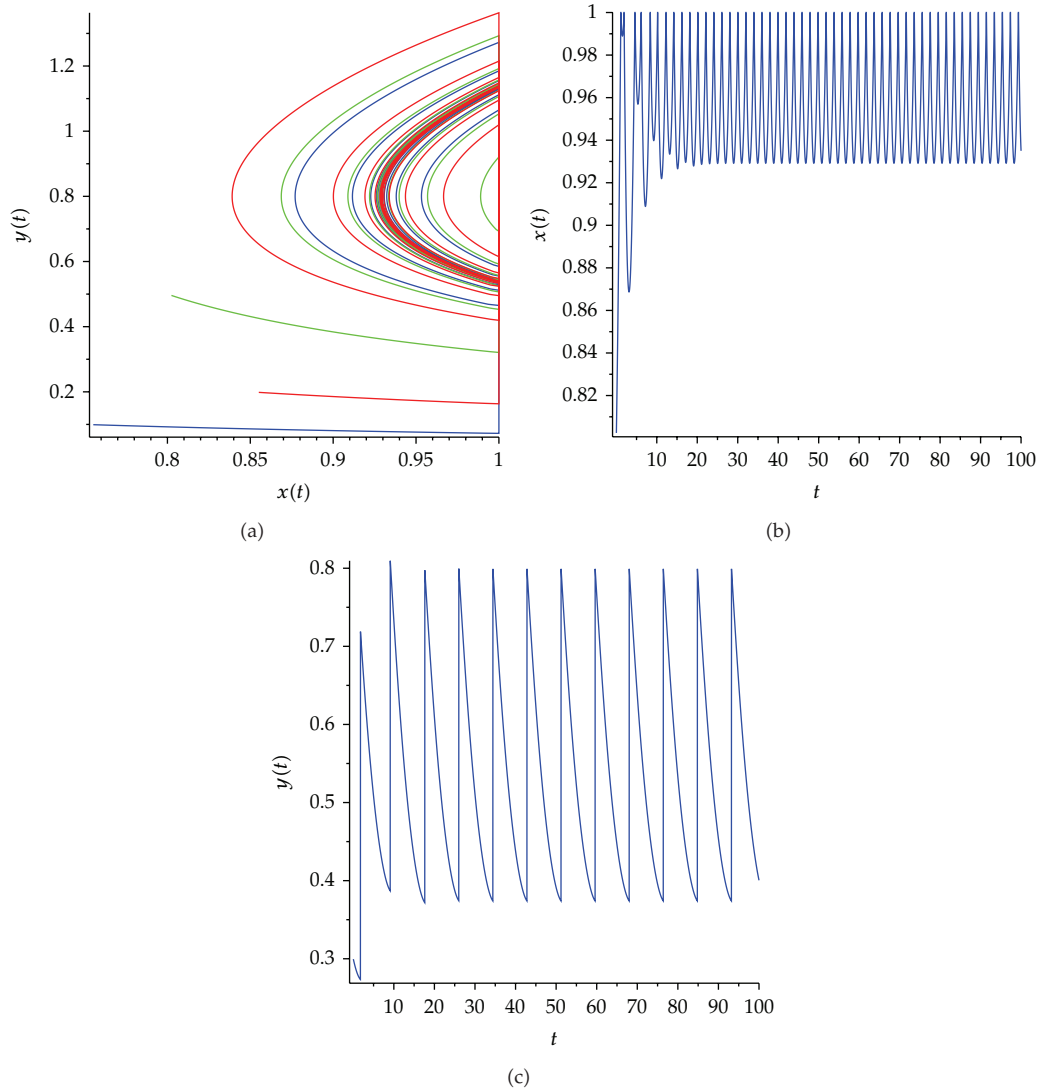


Figure 17: The time series and phase diagram for system (1.4) starting from initial value $(0.85, 0.2)$ (red), $(0.8, 0.5)$ (green), and $(0.75, 0.11)$ (blue), $\delta = 0.6$, $h_1 = 1 < x^*$.

The trajectory from initiating point between B_0^+ and P_0^+ will intersect with impulsive set N_2 with time increasing, under the impulsive effects it arrives at a point of N_2 which is between $[y_{P_{2k-1}^+}, y_{P_{2k+1}^+})$ or $[y_{P_{2k}^+}, y_{P_{2k-2}^+})$. Then like the analysis of Theorem 4.3, we know the trajectory from any initiating point between B_0^+ and P_0^+ will ultimately tend to be order one periodic solution $\widehat{P^+PP^+}$.

Denote the intersection of the trajectory passing through point B which tangents to N_1 at point B with the set N_2 at S . Since the trajectory from initiating any point below S of the set N_2 will certain intersect with set N_2 , next we only need to prove the trajectory with any initiating point below S of the set N_2 will ultimately tend to be order one periodic solution $\widehat{P^+PP^+}$.

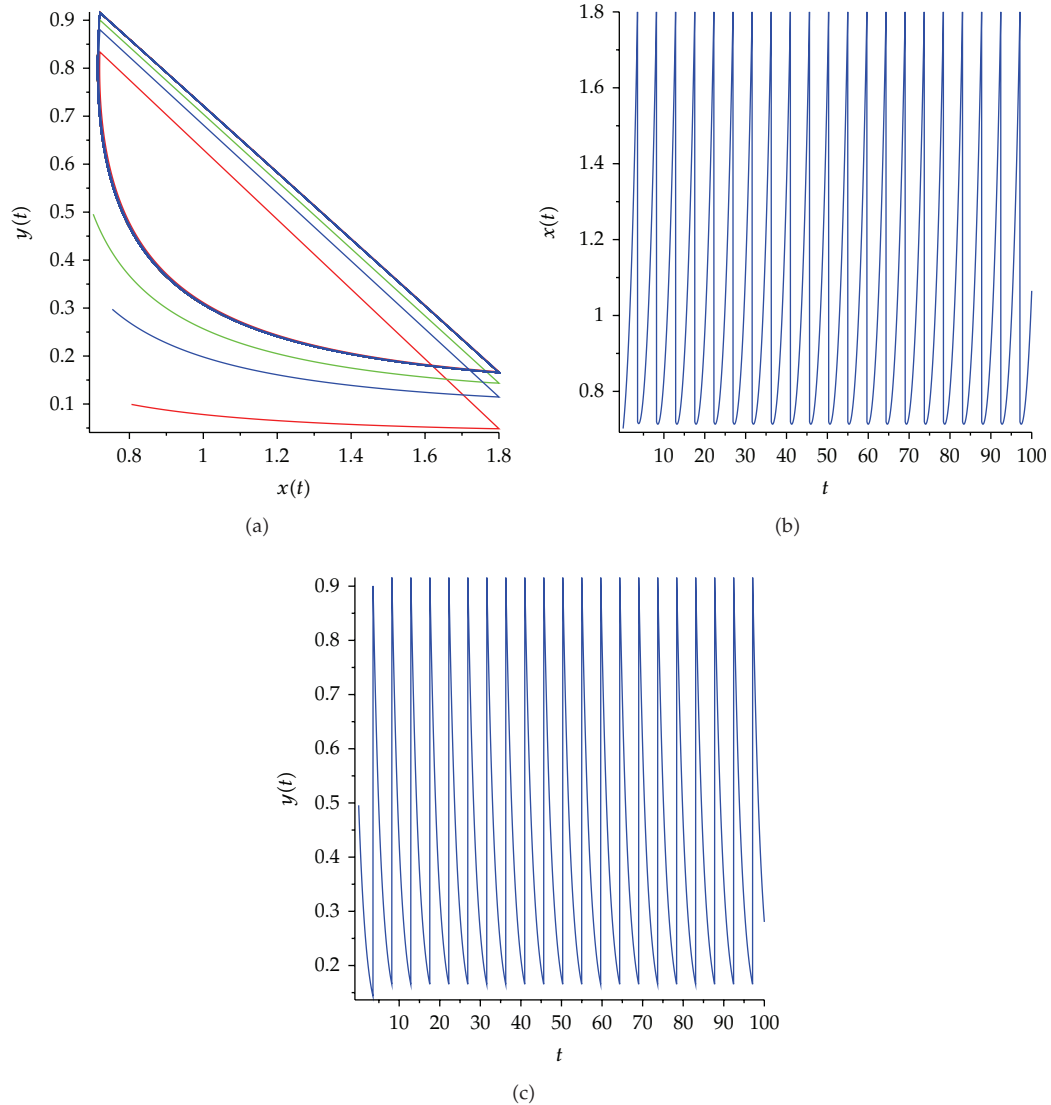


Figure 18: The time series and phase diagram for system (1.4) starting from initial value (0.8,0.1) (red), (0.7,0.5) (green), and (0.75,0.3) (blue) $\alpha = 0.6, \beta = 0.3, q = 0.8, h_2 = 1.8, h_1 < h_2 < x^*$.

Assume a point B_0 of set M_2 jumps to B_0^+ under the impulsive effect. Set $R(B_0) = B_1^+ \in N_2$. Assume point B_1 of set N_2 jumps to B_1^+ under the impulsive effect. Set $R(B_1) = B_2^+ \in N_2$. This process is continuing until there exists a $B_{K_0}^+ \in N$ ($K_0^+ \in N_2$) satisfying $y_{B_{K_0}^+} < q$. So we obtain a sequence $\{B_k\}_{k=0,1,2,\dots,K_0}$ of set M_2 and a sequence $\{B_k^+\}_{k=0,1,2,\dots,K_0}$ of set N_2 satisfying $R(B_{k-1}) = B_k^+, y_{B_k^+} < y_{B_{k-1}^+}$. For any point of set N_2 below B_0^+ , it must lie between B_{k+1}^+ and B_k^+ here $k = 1, 2, \dots, K_0$. After $K_0 + 1$ times' impulsive effects, the trajectory with this initiating point will arrive at some point of the set N_2 which must be between B_0^+ and P_0^+ , and then will ultimately tend to order one periodic solution $\widehat{P^+PP^+}$. There is no order one periodic solution with the initial point below B_0^+ .

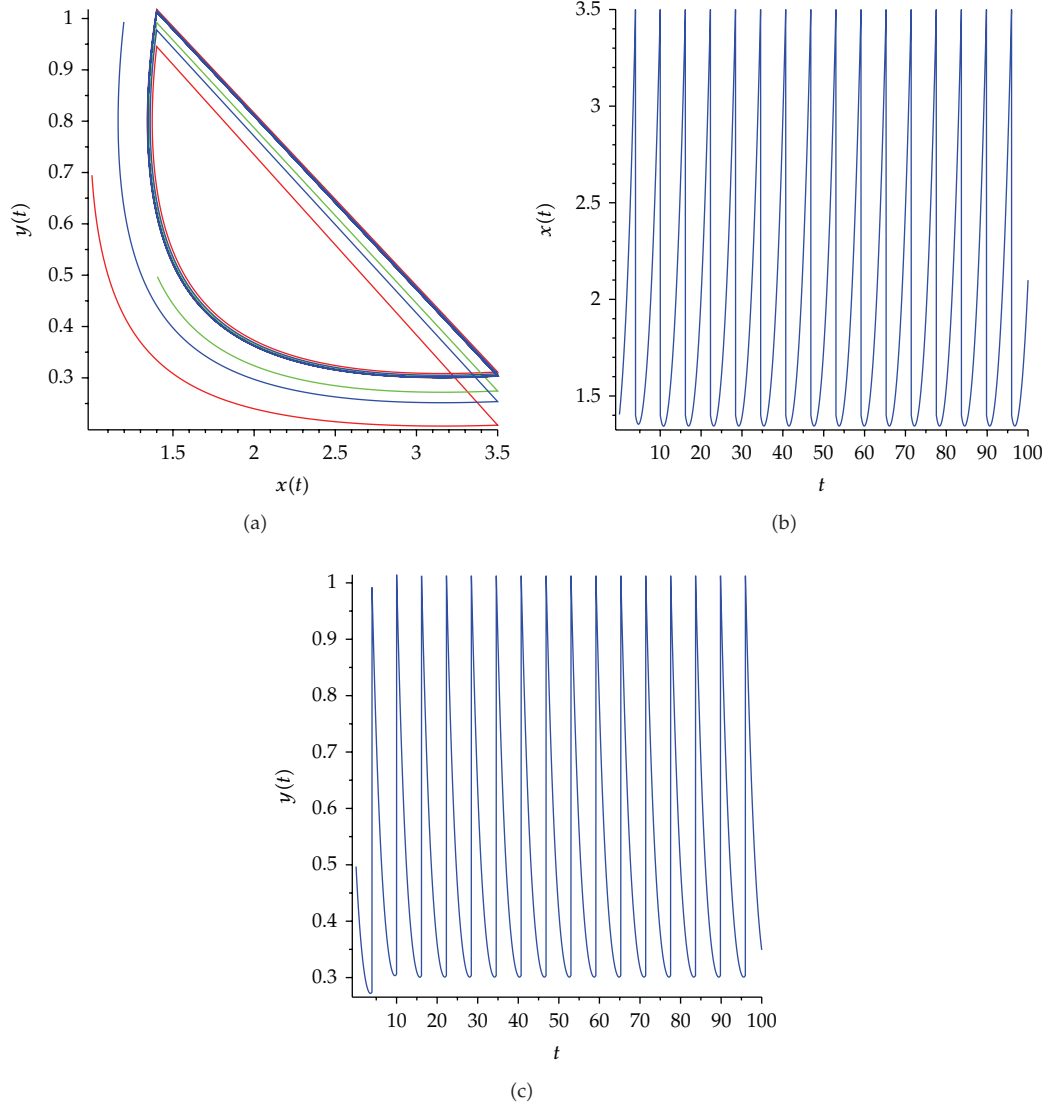


Figure 19: The time series and phase diagram for system (1.4) starting from initial value (1, 0.7) (red), (1.4, 0.5) (green), and (1.2, 1) (blue) $\alpha = 0.6$, $\beta = 0.3$, $q = 0.8$, $h_2 = 3.5$, $h_1 < x^* < h_2$.

The trajectory with any initiating point in segment \overline{AS} will intersect with the set N_2 at some point below B_0^+ with time increasing. Like the analysis above, we obtain the trajectory initiating any point on segment \overline{AS} will ultimately tend to be the unique order one periodic solution $\widehat{P^+PP^+}$.

From above analysis, we know there exists a unique order one periodic solution in system (1.4), and the trajectory from any initiating point below S will ultimately tend to be order one periodic solution $\widehat{P^+PP^+}$. Therefore, order one periodic solution $\widehat{P^+PP^+}$ is attractive in the region Ω_2 . This completes the proof. \square

5. Conclusion

In this paper, a state-dependent impulsive dynamical model concerning different control methods at different thresholds is proposed, we find a new method to study existence and attractiveness of order one periodic solution of such system. We define semicontinuous dynamical system and successor function, demonstrate the sufficient conditions that system (1.4) exists order one periodic solution with differential geometry theory and successor function. Besides, we successfully prove the attractiveness of the order one periodic solution by sequence convergence rules and qualitative analysis. The method can be also extended to mechanical dynamical systems with impacts, for example [19, 20].

These results show that the state-dependent impulsive effects contribute significantly to the richness of the dynamics of the model. The conditions of existence of order one periodic solution in this paper have more extensively applicable scope than the conditions given in [14]. Our results show that, in theory, a pest can be controlled such that its population size is no larger than its ET by applying effects impulsively once, twice, or at most, a finite number of times, or according to a periodic regime. The methods of the theorems are proved to be new in this paper, and these methods are more efficient and easier to operate than the existing research methods which have been applied the models with impulsive state feedback control [16–18, 21], so they are deserved further promotion. In this paper, according to the integrated pest management strategies, we propose a model for pest control which adopts different control methods at different thresholds, the corresponding control is exerted, which leads to the two state impulses in model. Certainly, many biological systems will always be described by three or more state variables, which are the main work in the future.

In order to testify the validity of our results, we consider the following example.

$$\begin{aligned} x'(t) &= x(t)(0.4 - 0.5y(t)), \\ y'(t) &= y(t) \left(\frac{0.25x(t)}{1 + 0.1x(t)} - 0.6 \right), \quad x \neq h_1, h_2 \text{ or } x = h_1, y > y^*, \\ \Delta x(t) &= 0, \\ \Delta y(t) &= \delta, \quad x = h_1, y \leq y^*, \\ \Delta x(t) &= -\alpha x(t), \\ \Delta y(t) &= -\beta y(t) + q, \quad x = h_2, \end{aligned} \tag{5.1}$$

where $\alpha, \beta \in (0, 1)$, $\delta > 0$, $q > 0$, $0 < h_1 < h_2$. Now, we consider the impulsive effects influences on the dynamics of system (5.1).

Example 5.1. Existence and attractiveness of order one periodic solution.

We set $h_1 = 1$, $\alpha = 0.6$, $\beta = 0.8$, $q = 0.8$, $h_2 = 1.8$, initiating points are (0.85, 0.2) (red), (0.8, 0.5) (green), and (0.75, 0.11) (blue), respectively. Figure 17 shows that the conditions of Theorems 3.1 and 4.1 hold, system (5.1) exists order one periodic solution. The trajectory from different initiating must ultimately tend to be the order one periodic solution. Therefore, order one periodic solution is attractive.

Example 5.2. Existence and attractiveness of positive periodic solution.

We set $h_1 = 0.7$, $\alpha = 0.6$, $\beta = 0.8$, $q = 0.8$, $h_2 = 1.8$, $h_1 < (1 - \alpha)h_2 < x^*$, initiating points are (0.8, 0.1) (red), (0.7, 0.5) (green), and (0.75, 0.3) (blue), respectively. Figure 18 shows that

the conditions of Theorems 3.3 and 4.3 hold, there exists order one periodic solution of the system (5.1), and the trajectory from different initiating must ultimately tend to be the order one periodic solution. Therefore, order one periodic solution is attractive.

Example 5.3. Existence and attractive of positive periodic solutions.

We set $h_1 = 0.7$, $\alpha = 0.6$, $\beta = 0.8$, $q = 0.8$, $h_2 = 3.5$, $h_1 < (1 - \alpha)h_2 < x^* < h_2$, initiating points are (1, 0.7) (red), (1.4, 0.5) (green), and (1.2, 1) (blue) as shown in Figure 19. Therefore, the conditions of Theorems 3.6 and 4.5 hold, then system (5.1) exists order one periodic solution, and it is attractive.

Acknowledgment

This Project supported by the National Natural Science Foundation of China (no. 10872118).

References

- [1] B. Liu, Y. J. Zhang, L. S. Chen, and L. H. Sun, "The dynamics of a prey-dependent consumption model concerning integrated pest management," *Acta Mathematica Sinica*, vol. 21, no. 3, pp. 541–554, 2005.
- [2] X. Song, M. Hao, and X. Meng, "A stage-structured predator-prey model with disturbing pulse and time delays," *Applied Mathematical Modelling*, vol. 33, no. 1, pp. 211–223, 2009.
- [3] P. Georgescu, H. Zhang, and L. Chen, "Bifurcation of nontrivial periodic solutions for an impulsively controlled pest management model," *Applied Mathematics and Computation*, vol. 202, no. 2, pp. 675–687, 2008.
- [4] Z. Xiong, Y. Xue, and S. Li, "A food chain system with Holling IV functional responses and impulsive effect," *International Journal of Biomathematics*, vol. 1, no. 3, pp. 361–375, 2008.
- [5] B. Liu, Y. Zhang, and L. Chen, "Dynamic complexities of a Holling I predator-prey model concerning periodic biological and chemical control," *Chaos, Solitons and Fractals*, vol. 22, no. 1, pp. 123–134, 2004.
- [6] G. Jiang, Q. Lu, and L. Peng, "Impulsive ecological control of a stage-structured pest management system," *Mathematical Biosciences and Engineering*, vol. 2, no. 2, pp. 329–344, 2005.
- [7] Z. Gao, "PD observer parametrization design for descriptor systems," *Journal of the Franklin Institute*, vol. 342, no. 5, pp. 551–564, 2005.
- [8] Z. W. Gao and D. W. C. Ho, "On state-space realization of Bezout factorizations in singular systems," *Dynamics of Continuous, Discrete & Impulsive Systems A*, vol. 13, no. 3-4, pp. 387–410, 2006.
- [9] A. D'Onofrio, "Pulse vaccination strategy in the SIR epidemic model: global asymptotic stable eradication in presence of vaccine failures," *Mathematical and Computer Modelling*, vol. 36, no. 4-5, pp. 473–489, 2002.
- [10] X. Z. Meng and L. S. Chen, "A stage-structured SI eco-epidemiological model with time delay and impulsive controlling," *Journal of Systems Science & Complexity*, vol. 21, no. 3, pp. 427–440, 2008.
- [11] L. Nie, J. Peng, Z. Teng, and L. Hu, "Existence and stability of periodic solution of a Lotka-Volterra predator-prey model with state dependent impulsive effects," *Journal of Computational and Applied Mathematics*, vol. 224, no. 2, pp. 544–555, 2009.
- [12] D. Bařnov and P. Simeonov, *Impulsive Differential Equations: Periodic Solutions and Applications*, Longman Scientific & Technical, New York, NY, USA, 1993.
- [13] X. Meng and Z. Li, "The dynamics of plant disease models with continuous and impulsive cultural control strategies," *Journal of Theoretical Biology*, vol. 266, no. 1, pp. 29–40, 2010.
- [14] S. Tang and R. A. Cheke, "State-dependent impulsive models of integrated pest management (IPM) strategies and their dynamic consequences," *Journal of Mathematical Biology*, vol. 50, no. 3, pp. 257–292, 2005.
- [15] S. Tang, Y. Xiao, L. Chen, and R. A. Cheke, "Integrated pest management models and their dynamical behaviour," *Bulletin of Mathematical Biology*, vol. 67, no. 1, pp. 115–135, 2005.
- [16] G. Jiang, Q. Lu, and L. Peng, "Impulsive ecological control of a stage-structured pest management system," *Mathematical Biosciences and Engineering*, vol. 2, no. 2, pp. 329–344, 2005.
- [17] B. Liu, Y. Zhang, and L. Chen, "Dynamic complexities of a Holling I predator-prey model concerning periodic biological and chemical control," *Chaos, Solitons and Fractals*, vol. 22, no. 1, pp. 123–134, 2004.

- [18] S. Tang, G. Tang, and R. A. Cheke, "Optimum timing for integrated pest management: modelling rates of pesticide application and natural enemy releases," *Journal of Theoretical Biology*, vol. 264, pp. 623–638, 2010.
- [19] J. Awrejcewicz, K. Tomczak, and C.-H. Lamarque, "Controlling system with impacts," *International Journal of Bifurcation and Chaos*, vol. 9, no. 3, pp. 547–553, 1999.
- [20] J. Awrejcewicz, M. Fečkan, and P. Olejnik, "On continuous approximation of discontinuous systems," *Nonlinear Analysis*, vol. 62, no. 7, pp. 1317–1331, 2005.
- [21] J. Jiao and L. Chen, "Global attractivity of a stage-structure variable coefficients predator-prey system with time delay and impulsive perturbations on predators," *International Journal of Biomathematics*, vol. 1, no. 2, pp. 197–208, 2008.

Research Article

Nonlinearities in Drug Release Process from Polymeric Microparticles: Long-Time-Scale Behaviour

Elena Simona Bacaita,^{1,2} Costica Bejinariu,³ Borsos Zoltan,⁴ Catalina Peptu,¹ Gabriela Andrei,¹ Marcel Popa,¹ Daniela Magop,⁵ and Maricel Agop^{2,6}

¹ Department of Natural and Synthetic Polymers, Faculty of Chemical Engineering and Environmental Protection, "Gheorghe Asachi" Technical University of Iasi, Prof. Dr. Docent Dimitrie Mangeron Road, No. 73, 700050 Iasi, Romania

² Department of Physics, Faculty of Machine Manufacturing and Industrial Management, "Gheorghe Asachi" Technical University of Iasi, Prof. Dr. Docent Dimitrie Mangeron Road, No. 59A, 700050 Iasi, Romania

³ Department of Materials Engineering and Industrial Security, Faculty of Materials Science and Engineering, "Gheorghe Asachi" Technical University of Iasi, Prof. Dr. Docent Dimitrie Mangeron Road, No. 59A, 700050 Iasi, Romania

⁴ Department of Technology of Information, Mathematics and Physics, Faculty of Letters and Sciences, Petroleum-Gas University of Ploiesti, Bucuresti Boulevard, No. 39, 100680 Ploiesti, Romania

⁵ Physics Department, "Al. I. Cuza" University, Carol I Road, No. 11, 700506 Iasi, Romania

⁶ Lasers, Atoms and Molecules Physics Laboratory, University of Science and Technology, Villeneuve d'Ascq, 59655 Lille, France

Correspondence should be addressed to Costica Bejinariu, c.bejinariu@yahoo.com and Maricel Agop, m.agop@yahoo.com

Received 4 May 2012; Revised 8 July 2012; Accepted 20 July 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 Elena Simona Bacaita et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A theoretical model of the drug release process from polymeric microparticles (a particular type of polymer matrix), through dispersive fractal approximation of motion, is built. As a result, the drug release process takes place through cnoidal oscillations modes of a normalized concentration field. This indicates that, in the case of long-time-scale evolutions, the drug particles assemble in a lattice of nonlinear oscillators occur macroscopically, through variations of drug concentration. The model is validated by experimental results.

1. Introduction

Polymer matrices can be produced in one of the following forms: micro/nanoparticles, micro/nanocapsules, hydrogels, films, and patches. Due to the multitude of biocompatible

polymers in the experimental protocol, drug proper delivery via many administration routes occurs. No matter what their form might be, drug carrier polymer matrices should have the following characteristics: biocompatibility, biodegradability, and controlled release capacity. The last one refers to the relationship between the efficient, nontoxic drug administration and therapeutic window type concentration, that is, minimum concentration is required to produce the wanted effect, but in the case of high levels, a toxic barrier occurs.

Given the importance of the released drug concentration, numerous studies have been performed with the purpose of identifying the mathematical function that describes time dependence. Many papers show how various factors, such as polymer molecular weight [1, 2], polymer chemistry, monomer ratios [3, 4], pH of release media, additives to the release media [5, 6], and particle size [7], affect the release kinetics. At the same time, certain phenomena appearing in the release process have been studied. Of these, we mention (in approximate order of their occurrence) polymer swelling and degradation [8–11], drug dissolution and diffusion [12, 13], and above all, permanent chemical and physical interaction among components (drug, polymer, and release medium). Since all these phenomena are not independent, their analysis becomes complicated; consequently, it will not be possible to treat them separately and cumulate the effects. For example, microparticle morphology changes due to polymer degradation, their surfaces becoming highly porous. This will lead to increased diffusion coefficients and hence certain connected phenomena, such as polymer degradation and drug diffusion [7].

The multitude of phenomena and dependencies occurring in drug release process as well as numerous structural entities (polymer, drug, and release medium) will turn the system into a complex one. Consequently, the complete theoretical analysis becomes more difficult in terms of performing.

Nevertheless, significant amount of work has been accomplished in mathematical modeling, with the purpose of predicting the concentration of the released drug and providing the analysis of fundamental processes that govern release. Higuchi [14] was among the first who produced a drug release model from nonswelling and nondissolving polymer matrices, assuming that such phenomenon is purely controlled through diffusion. A number of other models have also been proposed in order to predict drug release in the case of erosion [9, 11], swelling [8], and dissolution [13] influenced processes. These mathematical patterns have chosen only two phenomena, with the purpose of simplifying mathematical modeling, which, otherwise, proves to be quite difficult.

That is the main reason why it is necessary to use alternative approaches with reduced number of the approximations. One of such possible approaches is the fractal one [15, 16]. Its use is justified by natural and synthetic polymers that have been included in the category of fractional-dimensioned objects whose structures and behaviour can be described by means of fractal geometry [17, 18]. Moreover, it has been observed that the dynamics of drug release systems is a fractal one, because, in spite of complex phenomena and factors, mathematical expressions describing drug release kinetics from a variety of polymer matrices are power type laws (Higuchi [14] for nonswelling and nondissolving polymer, Ritger and Peppas [19] for nonswellable polymer in the form of slabs, spheres, cylinders, or discs, Peppas Sahlin [20] for solute release, Alfrey et al. [21] for diffusion in glassy polymers, etc.) specific for the fractal system evolution [22]. At the same time, it is quite important to emphasize that correlation of experimental data with the above-mentioned laws revealed good correlation in the first part (approximate 60%) of the release kinetics, the correlation coefficient decreasing according to time evolution.

Studies on the release from different types of systems (HPMC matrices [23], inert porous matrices [24], and sponges [25, 26]) have been performed. Such approach analyzes drug release kinetics through Monte Carlo simulation. In this perspective, release systems are considered as three-dimensional lattices with leak sites located at the boundaries of the lattice pattern. Particles are free to move inside the porous network according to the random walk model of the Fickian diffusion (the moving particles act as hard spheres colliding with each other and having no possibility to mutually penetrate).

The first studies started with simplifying approximations. Kosmidis et al. [23] consider that porosity has a constant value. Later on, Villalobos et al. [24] improved the model, assumed that network porosity behaves dynamically, and considered the effects of drug spatial distribution and initial drug concentration. All these approaches proved the validity of Weibull function (a continuous probability distribution function) for the entire release kinetics and consequently eliminate Peppas' temporal limitation of the equation and criticism lacking kinetic basis and physical nature of parameters [27].

Our new approach considers the entire system (drug-loaded polymer matrix in the release environment) as a type of "fluid" totally lacking interaction or neglecting physical interactions among particles. At the same time, the induced complexity is replaced by fractality. This will lead to particles moving on certain trajectories called geodesics within fractal space. This assumption represents the basis of the fractal approximation of motion in scale relativity theory (SRT) [28, 29], leading to a generalized fractal "diffusion" equation that can be analyzed in terms of two approximations (dissipative and dispersive).

The comparison between dissipative approximation (with dominant convective and dissipative processes) and the dispersive one allows theoretical demonstration of Weibull function that best describes the behaviour of drug release systems at short time scales. These phenomena will be the object of subsequent analysis since they are responsible for certain types of behaviour and characterized by high degree of nonlinearity in drug release systems.

This paper is structured as follows: theoretical model (Section 2), experimental results that validate the theoretical model (Section 3), and conclusions (Section 4).

2. Theoretical Model

2.1. Consequences of Nondifferentiability

We suppose that the drug release process takes place on continuous, but nondifferentiable curves (fractal curves). Then, nondifferentiability implies [28–30] the following.

- (i) A continuous and a nondifferentiable curve (or almost nowhere differentiable) is explicitly scale dependent, and its length tends to infinity, when the scale interval tends to zero. In other words, a continuous and nondifferentiable space is fractal, and in the general meaning Mandelbrot used this concept [15];
- (ii) Physical quantities will be expressed through fractal functions, namely, through functions that are dependent both on coordinate field and resolution scale. The invariance of the physical quantities in relation with the resolution scale generates special types of transformations, called resolution-scale transformations. In what follows, we will explain the above statement.

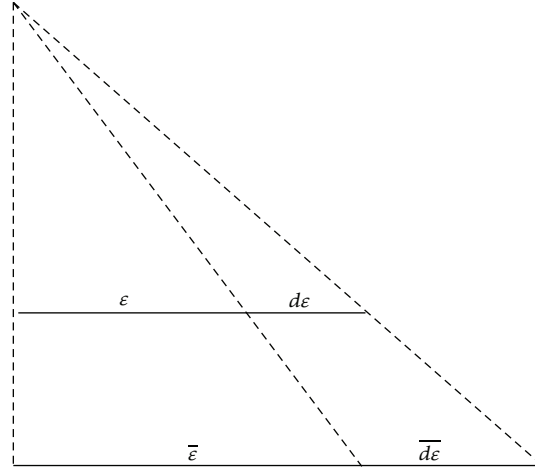


Figure 1: Dilatation-scale invariance.

Let $F(x)$ be a fractal function in the interval $x \in [a, b]$, and let the sequence of values for x be

$$x_a = x_0, \quad x_1 = x_0 + \varepsilon, \dots, \quad x_k = x_0 + k\varepsilon, \quad x_n = x_0 + n\varepsilon = x_b. \quad (2.1)$$

Let us note by $F(x, \varepsilon)$ the broken line that connects the points

$$F(x_0), \dots, F(x_k), \dots, F(x_n). \quad (2.2)$$

We can now say that $F(x, \varepsilon)$ is a ε -scale approximation.

Let us now consider $F(x, \bar{\varepsilon})$ as a $\bar{\varepsilon}$ -scale approximation of the same function. Since $F(x)$ is everywhere almost self-similar, if ε and $\bar{\varepsilon}$ are sufficiently small, both approximations $F(x, \varepsilon)$ and $F(x, \bar{\varepsilon})$ must lead to the same results; in the particular case, a fractal phenomenon is studied through approximation. By comparing the two cases, one notices that scale expansion is related to the increase $d\varepsilon$ of ε , according to an increase $d\bar{\varepsilon}$ of $\bar{\varepsilon}$ (see Figure 1). But, in this case, we have

$$\frac{d\varepsilon}{\varepsilon} = \frac{d\bar{\varepsilon}}{\bar{\varepsilon}} = d\rho, \quad (2.3)$$

a situation in which we can consider the infinitesimal-scale transformation as being

$$\varepsilon' = \varepsilon + d\varepsilon = \varepsilon + \varepsilon d\rho. \quad (2.4)$$

Such transformation in the case of function $F(x, \varepsilon)$ leads to

$$F(x, \varepsilon') = F(x, \varepsilon + \varepsilon d\rho), \quad (2.5)$$

respectively, if we limit ourselves to a first-order approximation:

$$F(x, \varepsilon') = F(x, \varepsilon) + \frac{\partial F(x, \varepsilon)}{\partial \varepsilon}(\varepsilon' - \varepsilon) = F(x, \varepsilon) + \frac{\partial F(x, \varepsilon)}{\partial \varepsilon} \varepsilon d\rho. \quad (2.6)$$

Moreover, let us notice that for an arbitrary but fixed ε_0 , we obtain

$$\frac{\partial \ln(\varepsilon/\varepsilon_0)}{\partial \varepsilon} = \frac{\partial (\ln \varepsilon - \ln \varepsilon_0)}{\partial \varepsilon} = \frac{1}{\varepsilon}, \quad (2.7)$$

a situation in which (2.6) can be written as follows:

$$F(x, \varepsilon') = F(x, \varepsilon) + \frac{\partial F(x, \varepsilon)}{\partial \ln(\varepsilon/\varepsilon_0)} d\rho = \left[1 + \frac{\partial}{\partial \ln(\varepsilon/\varepsilon_0)} d\rho \right] F(x, \varepsilon). \quad (2.8)$$

Therefore, we can introduce the dilatation operator

$$\hat{D} = \frac{\partial}{\partial \ln(\varepsilon/\varepsilon_0)}. \quad (2.9)$$

At the same time, relation (2.9) shows that the intrinsic variable of resolution is not ε , but $\ln(\varepsilon/\varepsilon_0)$.

The fractal function is explicitly dependant on the resolution $(\varepsilon/\varepsilon_0)$; therefore, we have to solve the differential equation

$$\frac{dF}{d \ln(\varepsilon/\varepsilon_0)} = P(F), \quad (2.10)$$

where $P(F)$ is now an unknown function. The simplest explicit suggested form for $P(F)$ is linear dependence [29]

$$P(F) = A + BF, \quad A, B = \text{const.}, \quad (2.11)$$

in which case the differential equation (2.10) takes the form

$$\frac{dF}{d \ln(\varepsilon/\varepsilon_0)} = A + BF. \quad (2.12)$$

Hence, by integration and substituting

$$B = -\tau, \quad (2.13a)$$

$$-\frac{A}{B} = F_0, \quad (2.13b)$$

we obtain

$$F\left(\frac{\varepsilon}{\varepsilon_0}\right) = F_0 \left[1 + \left(\frac{\varepsilon_0}{\varepsilon}\right)^\tau \right]. \quad (2.14)$$

This solution is independent as compared to parameterization on fractal curve.

We can now generalize the previous result by considering that F is dependent on parameterization of the fractal curve. If p characterizes the position on the fractal curve, then, following the same algorithm as above, the solution will be as a sum of two terms, that is, both classical and differentiable (depending only on position) and fractal, nondifferentiable (depending on position and, divergently, on $\varepsilon/\varepsilon_0$)

$$F\left(p, \frac{\varepsilon}{\varepsilon_0}\right) = F_0(p) \left[1 + \xi(p) \left(\frac{\varepsilon_0}{\varepsilon}\right)^{\tau(p)} \right], \quad (2.15)$$

where $\xi(p)$ is a function depending on parameterization of the fractal curve.

The following particular cases are to be considered.

(ii1) In asymptotic small-scale regime $\varepsilon \ll \varepsilon_0$, τ is constant (with no scale dependence) and power-law dependence on resolution is obtained:

$$F\left(p, \frac{\varepsilon}{\varepsilon_0}\right) = T(p) \left(\frac{\varepsilon_0}{\varepsilon}\right)^\tau, \quad (2.16a)$$

$$T(p) = F_0(p)Q(p). \quad (2.16b)$$

At this stage, some power laws should also be considered, namely, those equations describing drug release kinetics from a different type of polymer matrix [14, 19–21]. Consequently, through the appropriate correspondence among quantities from (2.16a) and (2.16b) and those from drug release processes, we will obtain the following:

(a) Higuchi law:

$$\frac{M_t}{M_\infty} = k_H \cdot t^{1/2}, \quad (2.17)$$

where M_t and M_∞ are cumulative amounts of drug release at time t and infinity, respectively, and k_H is a constant characteristic of the system [14];

(b) Peppas law:

$$\frac{M_t}{M_\infty} = k \cdot t^n, \quad (2.18)$$

where k is an experimentally obtained parameter, and n is a real number geometrically related to the system and to drug release mechanism. The n value is used to characterize different release mechanisms, that is, $n = 0.5$ indicates a Fickian diffusion. In their turn,

different from 0.5 n values refer to mass transport according to non-Fickian model [31]. This equation is a generalization of a square-root time law and an approximation for short times of Weibull function.

In these expressions, we recognize the standard form of self-similar fractal behaviour with fractal dimension $D_F = D_T + \tau$, which has already been used for accurately describing many physical and biological systems [15]. The topological dimensions are here $D_T = 1$, since we deal with length, but this can be easily generalized to surfaces ($D_T = 2$) and volumes ($D_T = 3$). Therefore, such result is not a consequence of postulation or deduction, but an aftermath of first principle theoretical analysis.

Considering that the resolution ε is a length, $\varepsilon = \delta X$, the scale-dependent length is given, by definition, by the law

$$X(p, \delta X) = X_0(p) \cdot \left(\frac{\lambda}{\delta X} \right)^{D_F-1}, \quad (2.19)$$

where λ is a scale characteristic length, and the exponent is identified with $\tau = D_F - 1$.

Now, in the above solution, one may use time t as parameter, and if one constantly moves along the curve, one obtains $X_0(t) = at$. Then, a differential version of the above relation will be

$$\delta X = a \delta t \cdot \left(\frac{\lambda}{\delta X} \right)^{D_F-1}, \quad (2.20)$$

so that the following fundamental relation among space and time elements on a fractal curve or function is obtained:

$$\delta X^{D_F} \propto \delta t. \quad (2.21)$$

In other words, they are differential elements of different orders.

(ii2) In the asymptotic big-scale regime $\varepsilon \gg \varepsilon_0$, τ is constant (with no scale dependence), and, in terms of resolution, one obtains an independent law

$$F\left(p, \frac{\varepsilon}{\varepsilon_0}\right) \rightarrow F_0(p). \quad (2.22)$$

Particularly, if $F(p, \varepsilon/\varepsilon_0)$ are the coordinates in given space, we can write

$$X\left(p, \frac{\varepsilon}{\varepsilon_0}\right) = x(p) \left[1 + \xi(p) \left(\frac{\varepsilon_0}{\varepsilon} \right)^\tau \right]. \quad (2.23)$$

In this situation, $\xi(p)$ becomes a highly fluctuating function which can be described by stochastic process, while τ represents (according to previous description) the difference between fractal and topological dimensions. The result is a sum of two terms, a classical, differentiable one (dependent only on the position) and a fractal, nondifferentiable one (dependent both on the position and, divergently, on $\varepsilon/\varepsilon_0$). This represents the importance of the above analysis.

By differentiating these two parts, we obtain

$$dX = dx + d\xi, \quad (2.24)$$

where dx is the classical differential element, and $d\xi$ is a differential fractal one.

(iii) There is infinity of fractal curves (geodesics) relating to any couple of points (or starting from any point) and applied for any scale. The phenomenon can be easily understood at the level of fractal surfaces, which, in their turn, can be described in terms of fractal distribution of conic points of positive and negative infinite curvature. As a consequence, we have replaced velocity on a particular geodesic by fractal velocity field of the whole infinite ensemble of geodesics. This representation is similar to that of fluid mechanics [32] where the motion of the fluid is described in terms of its velocity field $v = (x(t), t)$, density $\rho = (x(t), t)$, and, possibly, its pressure. We will, indeed, recover the fundamental equations of fluid mechanics (Euler and continuity equations), but we will write them in terms of a density of probability (as defined by the set of geodesics) instead of a density of matter and adding an additional term of quantum pressure (the expression of fractal geometry).

(iv) The local differential time invariance is broken, so the time derivative of the fractal field Q can be written as twofold:

$$\frac{d_+ Q}{dt} = \lim_{\Delta t \rightarrow 0_+} \frac{Q(t + \Delta t) - Q(t)}{\Delta t}, \quad (2.25a)$$

$$\frac{d_- Q}{dt} = \lim_{\Delta t \rightarrow 0_-} \frac{Q(t) - Q(t - \Delta t)}{\Delta t}. \quad (2.25b)$$

Both definitions are equivalent in the differentiable case $dt \rightarrow -dt$. In the nondifferentiable situation, these definitions are no longer valid, since limits are not defined anymore. Fractal theory defines physics in relationship with the function behavior during the “zoom” operation on the time resolution δt , here identified with the differential element dt (substitution principle), which is considered an independent variable. The standard field $Q(t)$ is therefore replaced by fractal field $Q(t, dt)$, explicitly dependent on time resolution interval, whose derivative is not defined at the unnoticeable limit $dt \rightarrow 0$. As a consequence, this leads to the two derivatives of the fractal field Q as explicit functions of the two variables t and dt ,

$$\frac{d_+ Q}{dt} = \lim_{\Delta t \rightarrow 0_+} \frac{Q(t + \Delta t, \Delta t) - Q(t, \Delta t)}{\Delta t}, \quad (2.26a)$$

$$\frac{d_- Q}{dt} = \lim_{\Delta t \rightarrow 0_-} \frac{Q(t, \Delta t) - Q(t - \Delta t, \Delta t)}{\Delta t}. \quad (2.26b)$$

Notation “+” corresponds to the forward process, while “−” to the backward one.

(v) Let $P(x^1, x^2)$ be a point of the fractal curve, and let us consider a line which starts from this point. Let M be the first intersection of this line with the fractal curve. By dX_+^i , we denote the components of the vector PM , to the right of the line (d), and by dX_-^i the components of the vector PM' , to the left of the line (d)—see Figure 2.

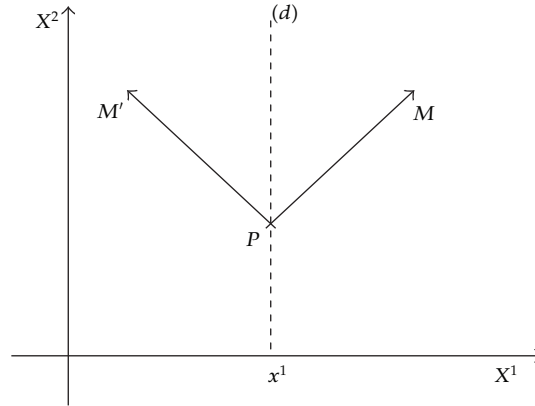


Figure 2: The continuous curves which are not fractals but have certain points where they are not differentiable.

If we consider all the lines (segments) starting from P , we denote the average of these vectors by dx_{\pm}^i , that is,

$$\langle dX_+^i \rangle = dx_+^i, \quad i = 1, 2, \quad (2.27a)$$

$$\langle dX_-^i \rangle = dx_-^i, \quad i = 1, 2. \quad (2.27b)$$

Since, according to (2.24), we can write

$$dX_+^i = dx_+^i + d\xi_+^i, \quad (2.28a)$$

$$dX_-^i = dx_-^i + d\xi_-^i, \quad (2.28b)$$

it results that

$$\langle d\xi_+^i \rangle = 0, \quad (2.29a)$$

$$\langle d\xi_-^i \rangle = 0. \quad (2.29b)$$

(vi) The differential fractal part satisfies, according to (2.21), the fractal equation

$$d_+ \xi^i = \lambda_+^i (dt)^{1/D_F}, \quad (2.30a)$$

$$d_- \xi^i = \lambda_-^i (dt)^{1/D_F}, \quad (2.30b)$$

where λ_+^i and λ_-^i are some constant coefficients, and D_F is a constant fractal dimension. We note that the use of any Kolmogorov or Hausdorff [15, 28, 33–35] definitions can be accepted for fractal dimension, but once a certain definition is admitted, it should be used until the end of analyzed dynamics.

(vii) The local differential time reflection invariance is recovered by combining the two derivatives, d_+/dt and d_-/dt , in the complex operator

$$\frac{\hat{d}}{dt} = \frac{1}{2} \left(\frac{d_+ + d_-}{dt} \right) - \frac{i}{2} \left(\frac{d_+ - d_-}{dt} \right). \quad (2.31)$$

Applying this operator to the “position vector,” a complex speed yields

$$\hat{\mathbf{V}} = \frac{\hat{d}\mathbf{X}}{dt} = \frac{1}{2} \left(\frac{d_+\mathbf{X} + d_-\mathbf{X}}{dt} \right) - \frac{i}{2} \left(\frac{d_+\mathbf{X} - d_-\mathbf{X}}{dt} \right) = \frac{\mathbf{V}_+ + \mathbf{V}_-}{2} - i \frac{\mathbf{V}_+ - \mathbf{V}_-}{2} = \mathbf{V} - i\mathbf{U}, \quad (2.32)$$

with

$$\mathbf{V} = \frac{\mathbf{V}_+ + \mathbf{V}_-}{2}, \quad (2.33a)$$

$$\mathbf{U} = \frac{\mathbf{V}_+ - \mathbf{V}_-}{2}. \quad (2.33b)$$

The real part, \mathbf{V} , of the complex speed $\hat{\mathbf{V}}$ represents the standard classical speed, which does not depend on resolution, while the imaginary part, \mathbf{U} , is a new quantity coming from resolution-dependant fractal.

2.2. Covariant Total Derivative in Drug Release Mechanism

Let us now assume that curves describing drug release (continuous but nondifferentiable) are immersed in a 3-dimensional space, and that \mathbf{X} of components X^i ($i = \overline{1,3}$) is the position vector of a point on the curve. Let us also consider the fractal field $Q(\mathbf{X}, t)$ and expand its total differential up to the third order

$$d_+Q = \frac{\partial Q}{\partial t} dt + \nabla Q \cdot d_+\mathbf{X} + \frac{1}{2} \frac{\partial^2 Q}{\partial X^i \partial X^j} d_+X^i d_+X^j + \frac{1}{6} \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} d_+X^i d_+X^j d_+X^k, \quad (2.34a)$$

$$d_-Q = \frac{\partial Q}{\partial t} dt + \nabla Q \cdot d_-\mathbf{X} + \frac{1}{2} \frac{\partial^2 Q}{\partial X^i \partial X^j} d_-X^i d_-X^j + \frac{1}{6} \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} d_-X^i d_-X^j d_-X^k, \quad (2.34b)$$

where only the first three terms were used in Nottale’s theory (i.e., second-order terms in the motion equation). Relations (2.34a) and (2.34b) are valid in any point both for the spatial

manifold and for the points \mathbf{X} on the fractal curve (selected in relations (2.34a) and (2.34b)). Hence, the forward and backward average values of these relations take the form

$$\begin{aligned} \langle d_{\pm}Q \rangle &= \left\langle \frac{\partial Q}{\partial t} dt \right\rangle + \langle \nabla Q \cdot d_{\pm}X \rangle + \frac{1}{2} \left\langle \frac{\partial^2 Q}{\partial X^i \partial X^j} d_{\pm}X^i d_{\pm}X^j \right\rangle \\ &\quad + \frac{1}{6} \left\langle \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} d_{\pm}X^i d_{\pm}X^j d_{\pm}X^k \right\rangle, \end{aligned} \quad (2.35)$$

$$\begin{aligned} \langle d_{-}Q \rangle &= \left\langle \frac{\partial Q}{\partial t} dt \right\rangle + \langle \nabla Q \cdot d_{-}X \rangle + \frac{1}{2} \left\langle \frac{\partial^2 Q}{\partial X^i \partial X^j} d_{-}X^i d_{-}X^j \right\rangle \\ &\quad + \frac{1}{6} \left\langle \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} d_{-}X^i d_{-}X^j d_{-}X^k \right\rangle. \end{aligned} \quad (2.36)$$

The following aspects should be mentioned: the mean value of function f and its derivatives coincide with themselves, and the differentials $d_{\pm}X^i$ and dt are independent; therefore, the average of their products coincides with the product of averages. Consequently, (2.35) and (2.36) become

$$d_{+}Q = \frac{\partial Q}{\partial t} dt + \nabla Q \langle d_{+}X \rangle + \frac{1}{2} \frac{\partial^2 Q}{\partial X^i \partial X^j} \langle d_{+}X^i d_{+}X^j \rangle + \frac{1}{6} \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} \langle d_{+}X^i d_{+}X^j d_{+}X^k \rangle, \quad (2.37a)$$

$$d_{-}Q = \frac{\partial Q}{\partial t} dt + \nabla Q \langle d_{-}X \rangle + \frac{1}{2} \frac{\partial^2 Q}{\partial X^i \partial X^j} \langle d_{-}X^i d_{-}X^j \rangle + \frac{1}{6} \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} \langle d_{-}X^i d_{-}X^j d_{-}X^k \rangle, \quad (2.37b)$$

or more, using (2.28a) and (2.28b) with characteristics (2.29a) and (2.29b),

$$\begin{aligned} d_{+}Q &= \frac{\partial Q}{\partial t} dt + \nabla Q \cdot d_{+}X + \frac{1}{2} \frac{\partial^2 Q}{\partial X^i \partial X^j} \left(d_{+}x^i d_{+}x^j + \langle d_{+}\xi^i d_{+}\xi^j \rangle \right) \\ &\quad + \frac{1}{6} \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} \left(d_{+}x^i d_{+}x^j d_{+}x^k + \langle d_{+}\xi^i d_{+}\xi^j d_{+}\xi^k \rangle \right), \end{aligned} \quad (2.38a)$$

$$\begin{aligned} d_{-}Q &= \frac{\partial Q}{\partial t} dt + \nabla Q \cdot d_{-}X + \frac{1}{2} \frac{\partial^2 Q}{\partial X^i \partial X^j} \left(d_{-}x^i d_{-}x^j + \langle d_{-}\xi^i d_{-}\xi^j \rangle \right) \\ &\quad + \frac{1}{6} \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} \left(d_{-}x^i d_{-}x^j d_{-}x^k + \langle d_{-}\xi^i d_{-}\xi^j d_{-}\xi^k \rangle \right). \end{aligned} \quad (2.38b)$$

Even if the average value of the fractal coordinate $d_{\pm}\xi^i$ is null (see (2.29a) and (2.29b)), for higher order of fractal coordinate average, the situation can still be different. Firstly, let

us focus on the averages $\langle d_+ \xi^i d_+ \xi^j \rangle$ and $\langle d_- \xi^i d_- \xi^j \rangle$. If $i \neq j$, these averages are zero due to the independence of $d_+ \xi^i$ and $d_+ \xi^j$. So, using (2.30a) and (2.30b), we can write

$$\langle d_+ \xi^i d_+ \xi^j \rangle = \lambda_+^i \lambda_+^j (dt)^{(2/D_F)-1} dt, \quad (2.39a)$$

$$\langle d_- \xi^i d_- \xi^j \rangle = \lambda_-^i \lambda_-^j (dt)^{(2/D_F)-1} dt. \quad (2.39b)$$

Then, let us consider the averages $\langle d_+ \xi^i d_+ \xi^j d_+ \xi^k \rangle$ and $\langle d_- \xi^i d_- \xi^j d_- \xi^k \rangle$. If $i \neq j \neq k$, these averages are zero due to independence of $d_+ \xi^i$ on $d_+ \xi^j$ and $d_+ \xi^k$. Now, using (2.30a) and (2.30b), we can write

$$\langle d_+ \xi^i d_+ \xi^j d_+ \xi^k \rangle = \lambda_+^i \lambda_+^j \lambda_+^k (dt)^{(3/D_F)-1} dt, \quad (2.40a)$$

$$\langle d_- \xi^i d_- \xi^j d_- \xi^k \rangle = \lambda_-^i \lambda_-^j \lambda_-^k (dt)^{(3/D_F)-1} dt. \quad (2.40b)$$

Then, (2.38a) and (2.38b) may be written as follows:

$$d_+ Q = \frac{\partial Q}{\partial t} dt + d_+ \mathbf{x} \cdot \nabla Q + \frac{1}{2} \frac{\partial^2 Q}{\partial X^i \partial X^j} d_+ \mathbf{x}^i d_+ \mathbf{x}^j + \frac{1}{2} \frac{\partial^2 Q}{\partial X^i \partial X^j} \lambda_+^i \lambda_+^j (dt)^{(2/D_F)-1} dt \quad (2.41a)$$

$$+ \frac{1}{6} \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} d_+ \mathbf{x}^i d_+ \mathbf{x}^j d_+ \mathbf{x}^k + \frac{1}{6} \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} \lambda_+^i \lambda_+^j \lambda_+^k (dt)^{(3/D_F)-1} dt,$$

$$d_- Q = \frac{\partial Q}{\partial t} dt + d_- \mathbf{x} \cdot \nabla Q + \frac{1}{2} \frac{\partial^2 Q}{\partial X^i \partial X^j} d_- \mathbf{x}^i d_- \mathbf{x}^j + \frac{1}{2} \frac{\partial^2 Q}{\partial X^i \partial X^j} \lambda_-^i \lambda_-^j (dt)^{(2/D_F)-1} dt \quad (2.41b)$$

$$+ \frac{1}{6} \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} d_- \mathbf{x}^i d_- \mathbf{x}^j d_- \mathbf{x}^k + \frac{1}{6} \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} \lambda_-^i \lambda_-^j \lambda_-^k (dt)^{(3/D_F)-1} dt.$$

If we divide by dt and neglect the terms containing differential factors (for details on the method, see [36, 37]), (2.41a) and (2.41b) are reduced to

$$\frac{d_+ Q}{dt} = \frac{\partial Q}{\partial t} + \mathbf{V}_+ \cdot \nabla Q + \frac{1}{2} \frac{\partial^2 Q}{\partial X^i \partial X^j} \lambda_+^i \lambda_+^j (dt)^{(2/D_F)-1} + \frac{1}{6} \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} \lambda_+^i \lambda_+^j \lambda_+^k (dt)^{(3/D_F)-1}, \quad (2.42a)$$

$$\frac{d_- Q}{dt} = \frac{\partial Q}{\partial t} + \mathbf{V}_- \cdot \nabla Q + \frac{1}{2} \frac{\partial^2 Q}{\partial X^i \partial X^j} \lambda_-^i \lambda_-^j (dt)^{(2/D_F)-1} + \frac{1}{6} \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} \lambda_-^i \lambda_-^j \lambda_-^k (dt)^{(3/D_F)-1}. \quad (2.42b)$$

These relations also allow us to define the operator

$$\frac{d_+}{dt} = \frac{\partial}{\partial t} + \mathbf{V}_+ \cdot \nabla + \frac{1}{2} \frac{\partial^2}{\partial X^i \partial X^j} \lambda_+^i \lambda_+^j (dt)^{(2/D_F)-1} + \frac{1}{6} \frac{\partial^3}{\partial X^i \partial X^j \partial X^k} \lambda_+^i \lambda_+^j \lambda_+^k (dt)^{(3/D_F)-1}, \quad (2.43a)$$

$$\frac{d_-}{dt} = \frac{\partial}{\partial t} + \mathbf{V}_- \cdot \nabla + \frac{1}{2} \frac{\partial^2}{\partial X^i \partial X^j} \lambda_-^i \lambda_-^j (dt)^{(2/D_F)-1} + \frac{1}{6} \frac{\partial^3}{\partial X^i \partial X^j \partial X^k} \lambda_-^i \lambda_-^j \lambda_-^k (dt)^{(3/D_F)-1}. \quad (2.43b)$$

Under these circumstances, let us calculate $(\hat{\partial}Q/\partial t)$. Taking into account (2.43a), (2.43b), (2.31), and (2.32), we will obtain

$$\begin{aligned} \frac{\hat{\partial}Q}{\partial t} &= \frac{1}{2} \left[\frac{d_+Q}{dt} + \frac{d_-Q}{dt} - i \left(\frac{d_+Q}{dt} - \frac{d_-Q}{dt} \right) \right] \\ &= \frac{1}{2} \frac{\partial Q}{\partial t} + \frac{1}{2} \mathbf{V}_+ \cdot \nabla Q + \lambda_+^i \lambda_+^j \frac{1}{4} (dt)^{(2/D_F)-1} \frac{\partial^2 Q}{\partial X^i \partial X^j} + \lambda_+^i \lambda_+^j \lambda_+^k \frac{1}{12} (dt)^{(3/D_F)-1} \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} \\ &\quad + \frac{1}{2} \frac{\partial Q}{\partial t} + \frac{1}{2} \mathbf{V}_- \cdot \nabla Q + \lambda_-^i \lambda_-^j \frac{1}{4} (dt)^{(2/D_F)-1} \frac{\partial^2 Q}{\partial X^i \partial X^j} + \lambda_-^i \lambda_-^j \lambda_-^k \frac{1}{12} (dt)^{(3/D_F)-1} \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} \\ &\quad - \frac{i}{2} \frac{\partial Q}{\partial t} - \frac{i}{2} \mathbf{V}_+ \cdot \nabla Q - \lambda_+^i \lambda_+^j \frac{i}{2} (dt)^{(2/D_F)-1} \frac{\partial^2 Q}{\partial X^i \partial X^j} - \lambda_+^i \lambda_+^j \lambda_+^k \frac{i}{12} (dt)^{(3/D_F)-1} \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} \\ &\quad + \frac{i}{2} \frac{\partial Q}{\partial t} + \frac{i}{2} \mathbf{V}_- \cdot \nabla Q + \lambda_-^i \lambda_-^j \frac{i}{2} (dt)^{(2/D_F)-1} \frac{\partial^2 Q}{\partial X^i \partial X^j} + \lambda_-^i \lambda_-^j \lambda_-^k \frac{i}{12} (dt)^{(3/D_F)-1} \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} \\ &= \frac{\partial Q}{\partial t} + \left(\frac{\mathbf{V}_+ + \mathbf{V}_-}{2} - i \frac{\mathbf{V}_+ - \mathbf{V}_-}{2} \right) \cdot \nabla Q \\ &\quad + \frac{(dt)^{(2/D_F)-1}}{4} \left[\left(\lambda_+^i \lambda_+^j + \lambda_-^i \lambda_-^j \right) - i \left(\lambda_+^i \lambda_+^j - \lambda_-^i \lambda_-^j \right) \right] \frac{\partial^2 Q}{\partial X^i \partial X^j} \\ &\quad + \frac{(dt)^{(3/D_F)-1}}{12} \left[\left(\lambda_+^i \lambda_+^j \lambda_+^k + \lambda_-^i \lambda_-^j \lambda_-^k \right) - i \left(\lambda_+^i \lambda_+^j \lambda_+^k - \lambda_-^i \lambda_-^j \lambda_-^k \right) \right] \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k} \\ &= \frac{\partial Q}{\partial t} + \hat{\mathbf{V}} \cdot \nabla Q + \frac{(dt)^{(2/D_F)-1}}{4} \left[\left(\lambda_+^i \lambda_+^j + \lambda_-^i \lambda_-^j \right) - i \left(\lambda_+^i \lambda_+^j - \lambda_-^i \lambda_-^j \right) \right] \frac{\partial^2 Q}{\partial X^i \partial X^j} \\ &\quad + \frac{(dt)^{(3/D_F)-1}}{12} \left[\left(\lambda_+^i \lambda_+^j \lambda_+^k + \lambda_-^i \lambda_-^j \lambda_-^k \right) - i \left(\lambda_+^i \lambda_+^j \lambda_+^k - \lambda_-^i \lambda_-^j \lambda_-^k \right) \right] \frac{\partial^3 Q}{\partial X^i \partial X^j \partial X^k}. \end{aligned} \quad (2.44)$$

This relation also allows us to define the fractal operator

$$\begin{aligned} \frac{\hat{\partial}}{\partial t} &= \frac{\partial}{\partial t} + \hat{\mathbf{V}} \cdot \nabla + \frac{(dt)^{(2/D_F)-1}}{4} \left[\left(\lambda_+^i \lambda_+^j + \lambda_-^i \lambda_-^j \right) - i \left(\lambda_+^i \lambda_+^j - \lambda_-^i \lambda_-^j \right) \right] \frac{\partial^2}{\partial X^i \partial X^j} \\ &\quad + \frac{(dt)^{(3/D_F)-1}}{12} \left[\left(\lambda_+^i \lambda_+^j \lambda_+^k + \lambda_-^i \lambda_-^j \lambda_-^k \right) - i \left(\lambda_+^i \lambda_+^j \lambda_+^k - \lambda_-^i \lambda_-^j \lambda_-^k \right) \right] \frac{\partial^3}{\partial X^i \partial X^j \partial X^k}. \end{aligned} \quad (2.45)$$

Particularly, by choosing

$$\lambda_+^i \lambda_+^j = -\lambda_-^i \lambda_-^j = 2\mathfrak{D}\delta^{ij}, \quad (2.46a)$$

$$\lambda_+^i \lambda_+^j \lambda_+^k = -\lambda_-^i \lambda_-^j \lambda_+^k = 2\sqrt{2}\mathfrak{D}^{3/2}\delta^{ijk}, \quad (2.46b)$$

the fractal operator (2.45) takes the usual form

$$\hat{\partial} = \frac{\partial}{\partial t} + \hat{\mathbf{V}} \cdot \nabla - i\mathfrak{D}(dt)^{(2/D_F)-1}\Delta + \frac{\sqrt{2}}{3}\mathfrak{D}^{3/2}(dt)^{(3/D_F)-1}\nabla^3. \quad (2.47)$$

We now apply the principle of scale covariance and postulate that the passage from classical (differentiable) to “fractal” mechanics can be implemented by replacing the standard time derivative operator, d/dt , with the complex operator $\hat{\partial}/\partial t$ (this results in a generalization of Nottale’s [28, 29] principle of scale covariance). Consequently, we are now able to write the diffusion equation in its covariant form

$$\frac{\hat{\partial}Q}{\partial t} = \frac{\partial Q}{\partial t} + (\hat{\mathbf{V}} \cdot \nabla)Q - i\mathfrak{D}(dt)^{(2/D_F)-1}\Delta Q + \frac{\sqrt{2}}{3}\mathfrak{D}^{3/2}(dt)^{(3/D_F)-1}\nabla^3 Q = 0. \quad (2.48)$$

This means that at any point on a fractal path, the local temporal $\partial_t Q$, the nonlinear (convective), $(\hat{\mathbf{V}} \cdot \nabla)Q$, the dissipative, ΔQ , and the dispersive, $\nabla^3 Q$, terms keep their balance.

The dissipative approximation was applied for the drug release processes, and the result was a Weibull type function that was analyzed in [38, 39]. In what follows, we will focus on dispersive approximation.

2.3. The Dispersive Approximation

Let us now consider that, in comparison with dissipative processes, convective and dispersive processes are dominant ones. Consequently, we are now able to write the diffusion equation in its covariant form, as a Korteweg de Vries type equation

$$\frac{\hat{\partial}Q}{\partial t} = \frac{\partial Q}{\partial t} + (\hat{\mathbf{V}} \cdot \nabla)Q + \frac{\sqrt{2}}{3}D^{3/2}(dt)^{(3/D_{FD})-1}\nabla^3 Q = 0. \quad (2.49)$$

If we separate the real and imaginary parts from (2.49), we will obtain

$$\frac{\partial Q}{\partial t} + \mathbf{V} \cdot \nabla Q + \frac{\sqrt{2}}{3}D^{3/2}(dt)^{(3/D_F)-1}\nabla^3 Q = 0, \quad (2.50a)$$

$$-\mathbf{U} \cdot \nabla Q = 0. \quad (2.50b)$$

By adding them, the fractal diffusion equation is

$$\frac{\partial Q}{\partial t} + (\mathbf{V} - \mathbf{U}) \cdot \nabla Q + \frac{\sqrt{2}}{3} D^{3/2} (dt)^{(3/D_F)-1} \nabla^3 Q = 0. \quad (2.51)$$

From (2.50b), we see that, at fractal scale, there will be no Q field gradient.

Assuming that $|\mathbf{V} - \mathbf{U}| = \sigma \cdot Q$ with $\sigma = \text{constant}$ (in systems with self-structuring processes, the speed fluctuations induced by fractal/nonfractal are proportional with the concentration field [22]), in the particular one-dimensional case, (2.51) will lack parameters

$$\bar{\tau} = \omega t, \quad (2.52a)$$

$$\bar{\xi} = kx, \quad (2.52b)$$

$$\Phi = \frac{Q}{Q_0}, \quad (2.52c)$$

and normalizing conditions

$$\frac{\sigma Q_0 k}{6\omega} = \frac{\sqrt{2} D^{3/2} (dt)^{(3/D_F)-1} k^3}{3\omega} = 1 \quad (2.53)$$

take the form

$$\partial_{\bar{\tau}} \phi + 6\phi \partial_{\bar{\xi}} \phi + \partial_{\bar{\xi} \bar{\xi} \bar{\xi}} \phi = 0. \quad (2.54)$$

In relations (2.52a), (2.52b), (2.52c), and (2.53), ω corresponds to a characteristic pulsation, k to the inverse of a characteristic length, and Q_0 to balanced concentration.

Through substitutions

$$w(\theta) = \phi(\bar{\tau}, \bar{\xi}), \quad (2.55a)$$

$$\theta = \bar{\xi} - u\bar{\tau}, \quad (2.55b)$$

(2.54), by double integration, becomes

$$\frac{1}{2} w'^2 = F(w) = -\left(w^3 - \frac{u}{2} w^2 - gw - h\right), \quad (2.56)$$

with g, h are two integration constants, and u is the normalized phase velocity. If $F(w)$ has real roots, (2.54) has the stationary solution

$$\phi(\bar{\xi}, \bar{\tau}, s) = 2a \left(\frac{E(s)}{K(s)} - 1 \right) + 2a \cdot \text{cn}^2 \left[\frac{\sqrt{a}}{s} \left(\bar{\xi} - \frac{u}{2} \bar{\tau} + \bar{\xi}_0 \right); s \right], \quad (2.57)$$

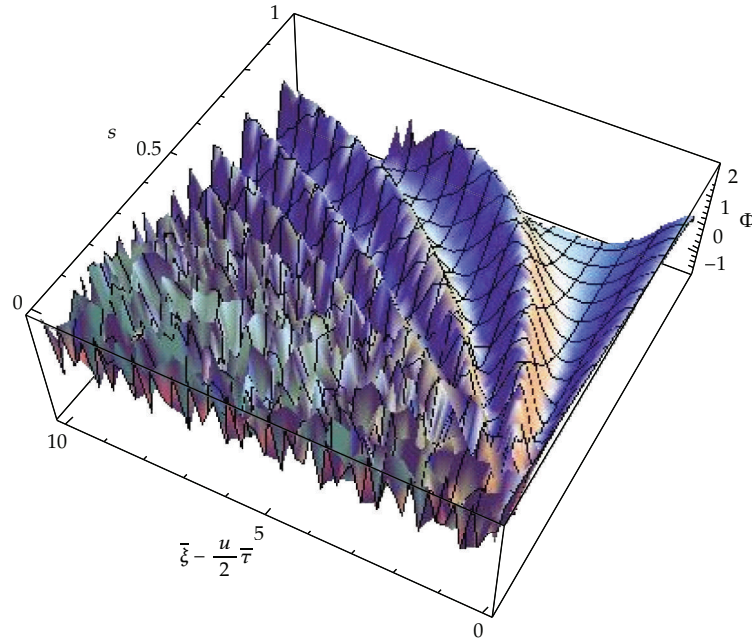


Figure 3: One-dimensional cnoidal oscillation modes of the field Φ .

where cn is Jacobi's elliptic function of s modulus [39], a is the amplitude, $\bar{\xi}_0$ is a constant of integration, and

$$K(s) = \int_0^{\pi/2} (1 - s^2 \sin^2 \varphi)^{-1/2} d\varphi, \quad (2.58a)$$

$$E(s) = \int_0^{\pi/2} (1 - s^2 \sin^2 \varphi)^{1/2} d\varphi, \quad (2.58b)$$

are the complete elliptic integrals [40].

Parameter s represents measure characterizing the degree of nonlinearity in the system. Therefore, the solution (2.57) contains (as subsequences for $s = 0$) one-dimensional harmonic waves, while for, $s \rightarrow 0$ one-dimensional wave packet. These two subsequences define the nonquasiautonomous regime of the drug release process [22, 36, 37], that is, the system should receive external energy in order to develop. For $s = 1$, the solution (2.57) becomes one-dimensional soliton, while for $s \rightarrow 1$, one-dimensional soliton packet will be generated. The last two imply a quasiautonomous regime (self-evolving and independent [22]) for drug particle release process [22, 36, 37].

The three-dimensional plot of solution (2.57) shows one-dimensional cnoidal oscillation modes of the concentration field, generated by similar trajectories of the drug particles (see Figure 3). We mention that cnoidal oscillations are nonlinear ones, being described by the elliptic function cn , hence the name (cnoidal).

It is known that in nonlinear dynamics, cnoidal oscillation modes are associated with nonlinear lattice of oscillators (the Toda lattice [41]). Consequently, large-time-scale drug

particle ensembles can be compared to a lattice of nonlinear oscillators which facilitates drug release process.

3. Experimental Results

Most of the experimental data in the literature reveal that, on average, drug release from polymeric matrices takes place according to a power law in the first 60% of the release curve and/or to exponential Weibull function on the entire drug release curve, reaching an average constant balanced value. The majority of these experimental results are carried out on relatively short time intervals, dissolution and diffusion being the dominant systems. The system exhibits a “burst effect” due to highly concentrated gradient. The phenomenon is followed by linear evolution on a constant value that corresponds to the balanced state with diluted gradient.

Nevertheless, some experimental results with long enough time intervals allow complete evolution of the process (polymer degradation stage is included here) and show unusually strong fluctuating behaviour.

Experimental data of drug release, at short and long time scales, for polymeric microparticles (as polymeric matrices) are presented below.

3.1. Materials

The following materials were used: low-molecular-weight chitosan-CS, deacetylation degree 75–80% (Aldrich), type B gelatin-GEL (Aldrich), glutaraldehyde-GA (Aldrich)-25% aqueous solution, sodium tripolyphosphate-TPP (Sigma), Levofloxacin-LEV (Sigma), Tween 80 (Aldrich), and Span 80 (Aldrich).

3.2. Preparation of Microparticles by Ionic Gelation and Covalent Cross-Linking in O/W/O Emulsion

Microparticles were prepared using an original double cross-linking method of a CS-GEL mixture. Different weight ratios CS/GEL (in terms of amino groups of both polymers) were dissolved in acetic acid solution 2%, and then Tween 80 was added to make a 2% (w/w) surfactant in the solution. The mixture was magnetically stirred until the surfactant was completely dissolved. Two different solutions of 2% Span 80 in toluene were prepared according to $O1/W (v/v) = 1/4$ and $O2/(O1 + W) = 4/1$. The organic phase O1 was dripped within the aqueous polymer phase, W under homogenization with an Ultraturax device at 9000 rpm. The primary emulsion was transformed into a double one through dripping in the second organic phase O2, according to the same hydrodynamic regime. The emulsion was then gelled by slowly adding a TPP solution at a rate of 2 mL/min with continuous stir for extra 10 min.

The suspension was then transferred to a round-bottom flask and mechanically stirred at 500 rpm. A certain amount of a saturated solution of GA in toluene was consecutively added and stirred for 60 min. The particles were separated by centrifugation (6000 rpm) and repeatedly washed with acetone and water in order to eliminate residual compounds. After hexane wash, the particles were dried at room temperature.

Table 1: The variable parameters in the preparation protocol.

Sample code	CS/GEL (w/w)	Conc. of TPP sol. (%)	NH ₂ /TPP (mols/mols)
C3	1/1	1	2.4/1
C1		5	
C4		10	
C2		15	
C7	1/1	5%	1.17/1
C1			2.4/1
C5			4.8/1
C6			11.7/1
C5	1/1	5%	4.8/1
C8	1/0		
C9	3/1		

3.3. Preparation Parameters

A two-step solidification method was used. The first step, which has critical influence over the subsequent particle shape and size, included ionic cross-linking with TPP effect through phosphate bridges among amino functionalities in both types of polymeric chains. The GA covalent cross-linking (also taking place in NH₂ groups) was performed with the purpose of stabilizing gel capsules. Our study analyzes the influence of the following cross-linking reaction parameters on the levofloxacin release kinetics:

- (i) concentration of the ionic cross-linker,
- (ii) ratio among amino functionalities of the two polymers and the ionic crosslinker,
- (iii) polymer composition of the polymer mixture.

Table 1 shows the variable parameters in the preparation protocol that have been grouped according to the variable parameters.

3.4. Levofloxacin Release Kinetics

3.4.1. Levofloxacin Release Kinetics at Small-Time Scales

If the experimental time scale is of minutes order, the evolution of the released drug concentration will be described by Peppas law. In this case, the correlated factor ranges between 0.8413 and 0.9983. Experimental and Peppas curves can be observed in Figure 4 (the Peppas parameters and the correlation coefficient R^2 for each sample are given in Table 2). The plots group according to variable preparation parameters (for a better observation of the first points, time scale is 500 min, although the fitting was made on the points up to 1440 min (one day)). Relative errors range between 1% and 5%, with no important influence on release kinetic evolution.

Previous works have shown the form dependence [38, 39] between the value of parameter n in Peppas equation (considered as short-time approximation of Weibull function) and the fractal dimension of the drug particle during the release process (D_f)

$$n = \frac{2}{D_f}. \quad (3.1)$$

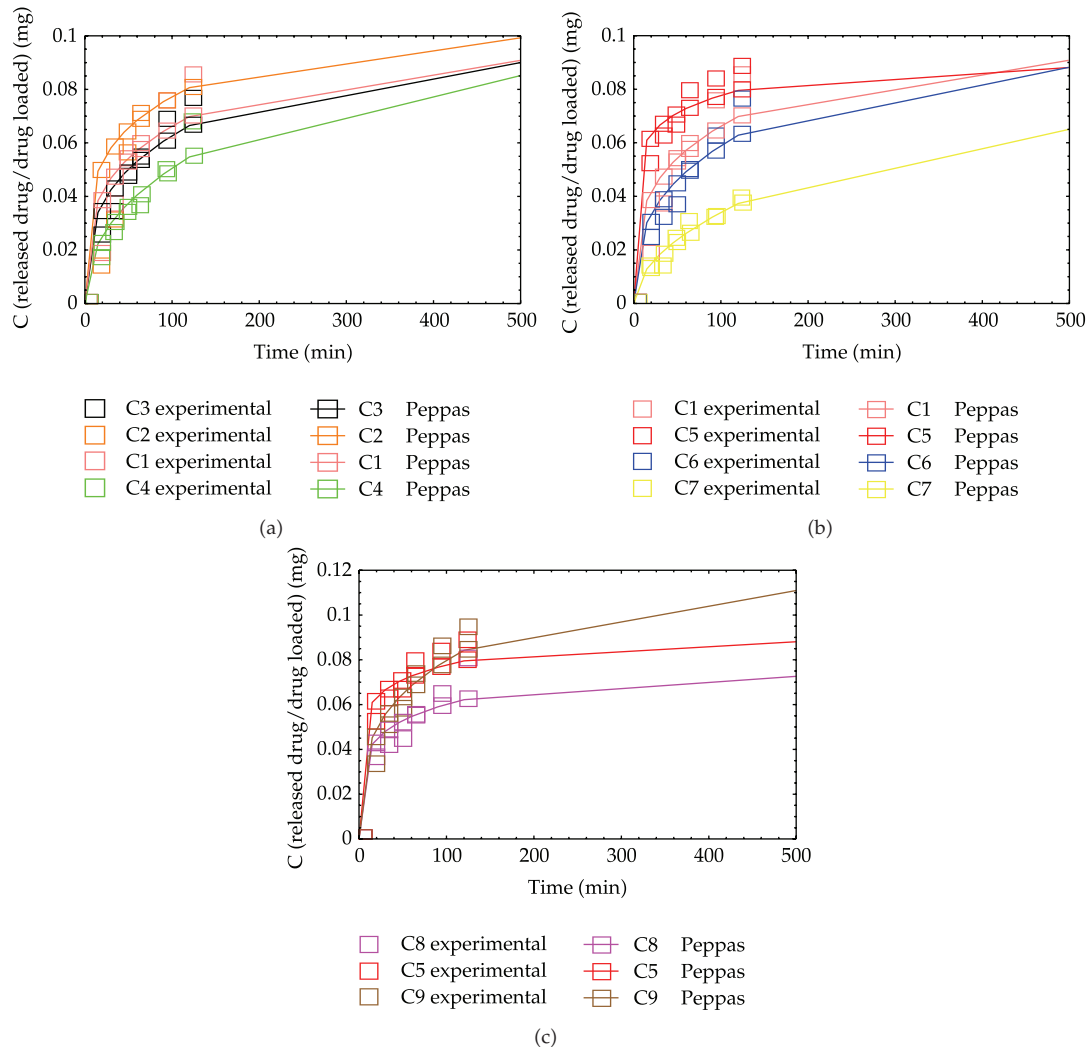


Figure 4: Levofloxacin release kinetics (experimental and Peppas fitting), at time scale of hours order, at different concentration of TPP solution (a), NH_2 : TPP mol ratios (b), and NH_2 mols (c).

Thus, according to experimental data, the following values were obtained in Table 2.

One first observation refers to the proportional dependence among experimental variable, on one hand, and Peppas parameters, on the other, in the particular case of the third sample group, that is, n increases with the chitosan/gelatin ratio. This proves to be experimentally useful if we want to obtain a Fickian diffusion. At the same time, the concentration of the released drug proves to be very low. This could be explained by drug crystallization inside the microparticle and the dependence of its release (dissolution followed by diffusion) on polymer degradation.

In our opinion, the value of the fractal dimension is important as long as its values are unusually high and indicate that either fractal dimension must be considered as function of structure "classes," or drug release processes (implicitly drug particle trajectories) have high

Table 2: The variable parameters of the preparation protocol.

Sample code	Conc. of TPP sol. (%)	k	n	Fractal dimension	R^2
C3	1	0.0142	0.3225	6.20	0.9884
C1	5	0.0176	0.2879	6.95	0.9645
C4	10	0.0069	0.4326	4.62	0.9932
C2	15	0.0261	0.2357	8.49	0.8413
Sample code	NH ₂ /TPP (mols/mols)	k	n	Fractal dimension	R^2
C7	1.17/1	0.0032	0.5131	3.90	0.9983
C1	2.4/1	0.0176	0.2879	6.95	0.9645
C5	4.8/1	0.0431	0.128	15.63	0.9787
C6	11.7/1	0.0116	0.3529	5.67	0.9879
Sample code	CS/GEL (w/w)	k	n	Fractal dimension	R^2
C5	1/1	0.0431	0.128	15.63	0.9787
C8	1/0	0.0256	0.1854	10.79	0.9539
C9	3/1	0.0201	0.2994	6.68	0.9895

degrees of complexity and nonlinearities, implying many freedom degrees in the phase space [42].

This analysis (small concentration of the released drug and high fractal dimensions) made us continue the experiment until the system reached a stationary state.

3.4.2. Levofloxacin Release Kinetics at Large-Time Scales

The experiments at large-time scales (of days order) revealed unusual behavior characterized by large variations. The release kinetics of levofloxacin is plotted in Figure 5. The relative errors range between 1% and 5%, and, for better visualization, the error bars are plotted in Figure 6.

Experiments have been performed for 28 days, the concentration of the released drug being measured daily, at the same hour. The general characteristic of the above kinetics refers to strong variations of concentration in time, approximately at the same moment.

In the following section, we will explain the evolution of these systems through the theoretical model (developed in Section 2) based on fractal approximation of motion.

3.5. The Correspondence between Theoretical Model and Experimental Results

In what follows, we identify the field Φ from relation (2.57) with normalized concentration field of the released drug from microparticles.

For best correlation between experimental data and the theoretical model (for each sample), we used a planar intersection of the graph in Figure 3, where the two variables are $y = (\bar{\xi} - \bar{\tau}u)/2$ and $x = s$. With these variables, (2.57) becomes

$$\phi_1(x, y) = 2a \left(\frac{E(x)}{K(x)} - 1 \right) + 2a \cdot \text{cn}^2 \left[\frac{\sqrt{a}}{x} (y + \xi_0); x \right]. \quad (3.2)$$

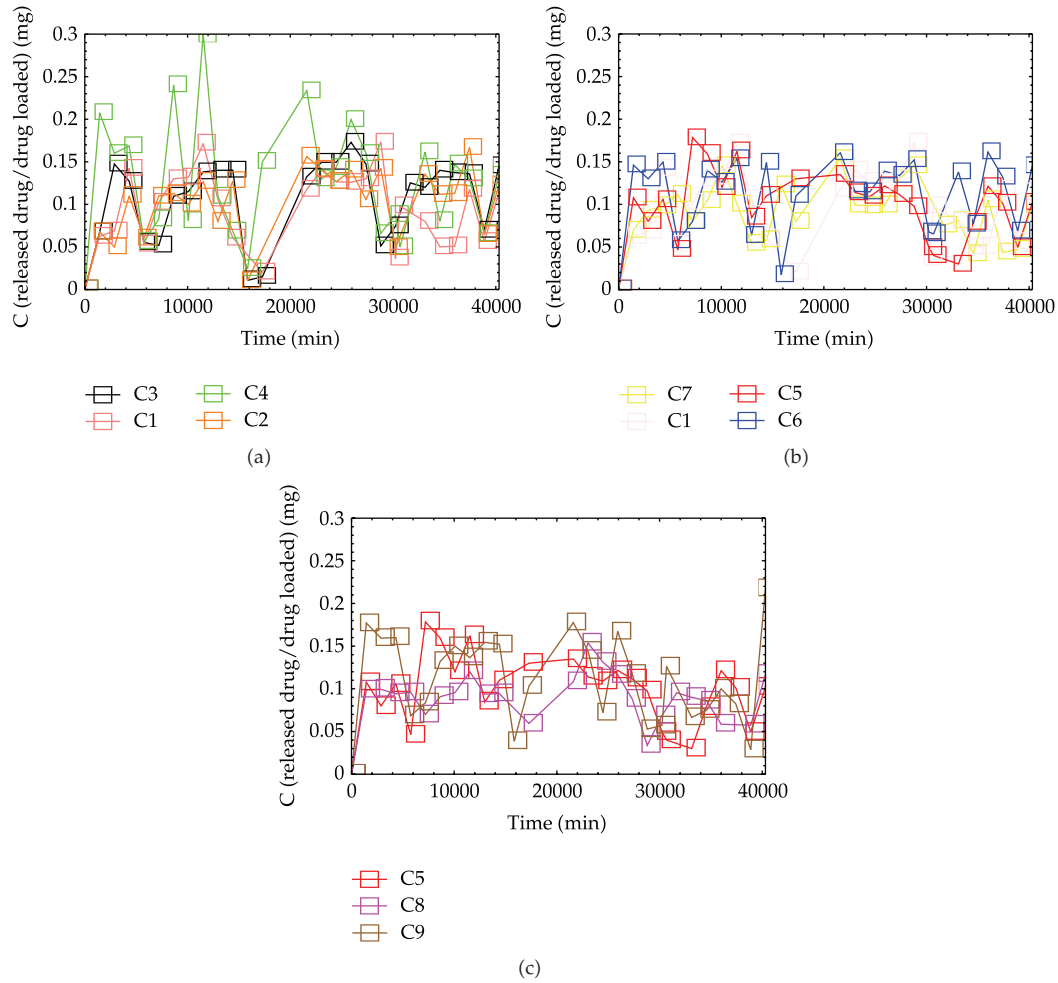


Figure 5: Experimental release kinetics of levofloxacin, at time scale of days order, at different concentration of TPP solution (a), NH₂:TPP mol ratios (b), and NH₂ mols (c).

Thus, in order to find the one-dimensional equation for a planar intersection, perpendicular to plane xOy , we used $y = mx + n$ (linear function equation), where m and n are two parameters. This equation is transformed into a parametric equation by means of the following substitutions:

$$x = \frac{l}{\sqrt{m^2 + 1}}, \quad (3.3a)$$

$$y = n + m \frac{l}{\sqrt{m^2 + 1}}, \quad (3.3b)$$

in (3.2).

Table 3: Parameters of the planar intersections.

Sample code	Conc. of TPP sol. (%)	n	m
C3	1	3.342	9.17
C1	5	6.902	8.024
C4	10	8.125	13.486
C2	15	9.479	12.25
Sample code	NH ₂ /TPP (mols/mols)	n	m
C7	1.17/1	7.322	4.297
C1	2.4/1	6.902	8.024
C5	4.8/1	2.414	8.665
C6	11.7/1	4.24	12.747
Sample code	CS/GEL (w/w)	n	m
C5	1/1	2.414	8.665
C8	1/0	8.303	5.941
C9	3/1	8.678	3.738

Afterwards, we obtain one-dimensional function

$$\phi_2(t, m, n) = \phi_1\left(\frac{t}{\sqrt{m^2 + 1}}, n + m \frac{t}{\sqrt{m^2 + 1}}\right). \quad (3.4)$$

The highest value of the correlation coefficient (for two vectors: one obtained from this very function, the other from experimental data) for different values of m and n (in the particular experimental case) will represent the best approximation of experimental data with the theoretical model.

Our goal was to find the right correlation coefficient which should be higher than 0.6-0.7, in order to demonstrate the relevance of the model we had in view. Figure 6 shows experimental and theoretical curves that were obtained through our method, where R^2 represents the correlation coefficient and η a normalized variable which is simultaneously dependent on normalized time and on nonlinear degree of the system (s parameter). Geometrically, η represents the congruent angle formed by the time axis and the vertical intersection plane.

Parameters m and n of the planar intersections for the above theoretical curves are shown in Table 3.

We must mention that for each sample the fitting process was an independent one. The corresponding intersection plane that offers best correlation factors had to be identified by each sample in turn.

A first observation refers to the correlation among plane and variable parameters (within experimental protocol) differ from Peppas small-time-scale fitting.

We consider that this could be a starting point in establishing dependence among experimental parameters involved in the protocol. The purpose of this analysis is to obtain polymer matrices together with characteristics of release kinetics, taking into account that until now, this type of dependence had to pass through intermediary stages of the physical and chemical characterization of polymer matrices.

The few experimental data could not sustain a general conclusion on the existing dependence among plane and experimental parameters, but this will be the purpose of a next paper.

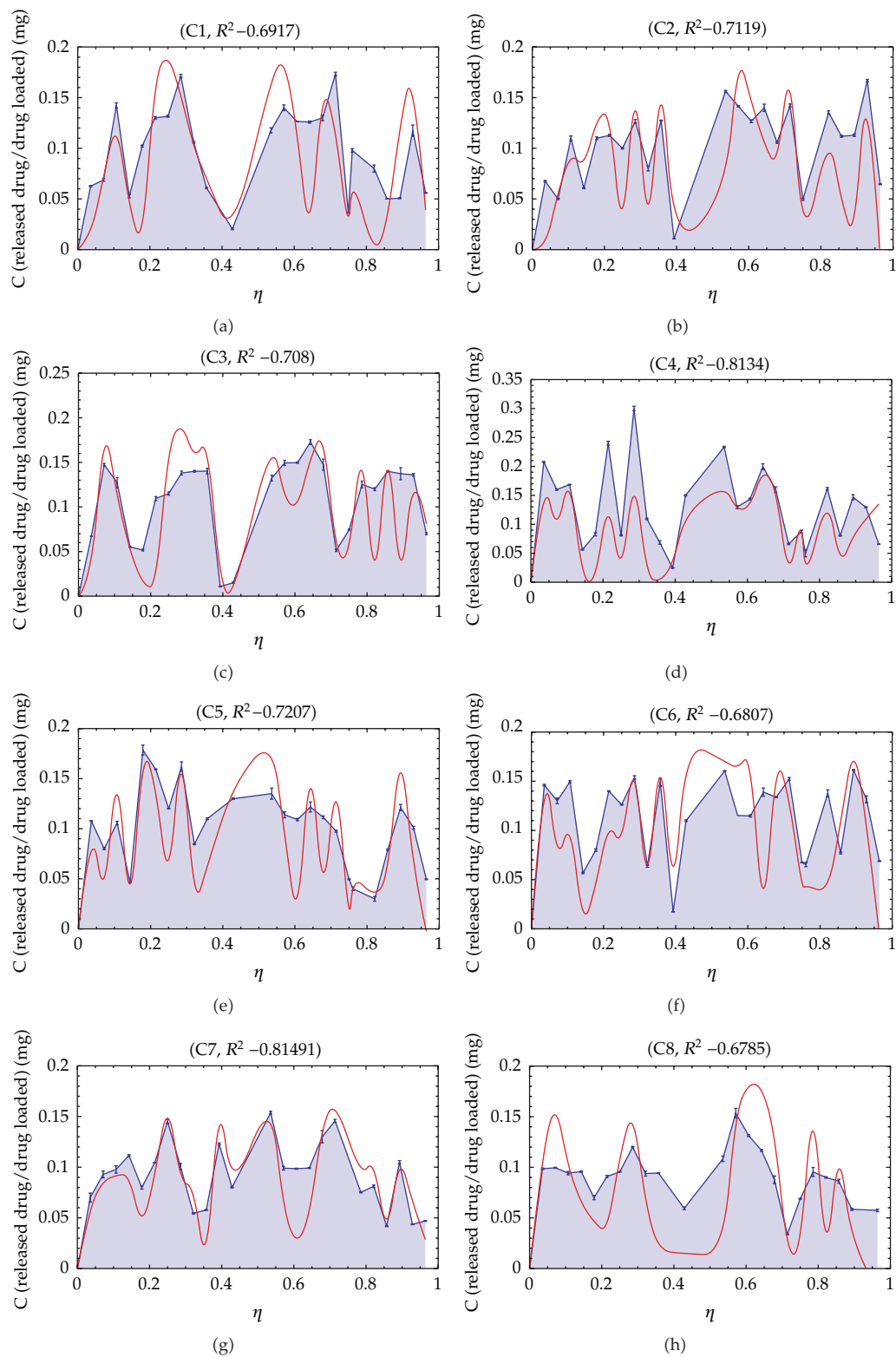


Figure 6: Continued.

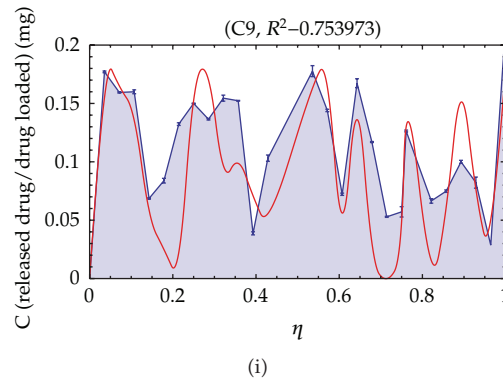


Figure 6: The best correlations among experimental and theoretical curves (blue line—experimental curve, red line—theoretical curve).

4. Conclusions

If the particle moves on fractal curves, a new model for drug release mechanism from polymer matrix (namely, polymeric particles) is obtained. This model offers new alternatives for the theoretical study of drug release process (on large time scale) in the presence of all phenomena and considering a highly complex and implicitly nonlinear system. Consequently, the concentration field has cnoidal oscillation modes, generated by similar trajectories of drug particles. This means that the drug particle ensemble (at time large scale) works in a network of nonlinear oscillators, with oscillations around release boundary. Moreover, the normalized concentration field simultaneously depends on normalized time nonlinear system (through s parameter).

The fitting procedure among experimental and theoretical curves revealed the existing correlation of some characteristics of the release kinetics (the parameters of the intersection plane) with variable experimental parameters.

At the same time, we consider that this could be a starting point in establishing dependence among experimental parameters, taking into account that until now, this type of dependence had to pass through intermediary stages of physical and chemical characteristics of polymer matrices.

Acknowledgments

This paper was supported by the project PERFORM-ERA “Postdoctoral Performance for Integration in the European Research Area” (ID-57649), financed by the European Social Fund and the Romanian Government, and by the project POSDRU/88/1.5/S/47646 of The European Social Fund.

References

- [1] A. J. Shukla and J. C. Price, “Effect of drug loading and molecular weight of cellulose acetate propionate on the release characteristics of theophylline microspheres,” *Pharmaceutical Research*, vol. 8, no. 11, pp. 1396–1400, 1991.

- [2] N. M. Najib, M. Suleiman, and A. Malakh, "Characteristics of the in vitro release of ibuprofen from polyvinyl pyrrolidone solid dispersions," *International Journal of Pharmaceutics*, vol. 32, no. 2-3, pp. 229–236, 1986.
- [3] F. J. Wang and C. H. Wang, "Etanidazole-loaded microspheres fabricated by spray-drying different poly(lactide/glycolide) polymers: effects on microsphere properties," *Journal of Biomaterials Science*, vol. 14, no. 2, pp. 157–183, 2003.
- [4] J. M. Bezemer, R. Radersma, D. W. Grijpma, P. J. Dijkstra, C. A. Van Blitterswijk, and J. Feijen, "Microspheres for protein delivery prepared from amphiphilic multiblock copolymers 2," *Journal of Controlled Release*, vol. 67, no. 2-3, pp. 249–260, 2000.
- [5] G. Crotts, H. Sah, and T. G. Park, "Adsorption determines in-vitro protein release rate from biodegradable microspheres: quantitative analysis of surface area during degradation," *Journal of Controlled Release*, vol. 47, no. 1, pp. 101–111, 1997.
- [6] H. Robson, D. Q. M. Craig, and D. Deutsch, "An investigation into the release of cefuroxime axetil from taste-masked stearic acid microspheres—III. The use of DSC and HSDSC as means of characterising the interaction of the microspheres with buffered media," *International Journal of Pharmaceutics*, vol. 201, no. 2, pp. 211–219, 2000.
- [7] J. Siepmann, N. Faisant, J. Akiki, J. Richard, and J. P. Benoit, "Effect of the size of biodegradable microparticles on drug release: experiment and theory," *Journal of Controlled Release*, vol. 96, no. 1, pp. 123–134, 2004.
- [8] P. Borgquist, A. Körner, L. Piculell, A. Larsson, and A. Axelsson, "A model for the drug release from a polymer matrix tablet-effects of swelling and dissolution," *Journal of Controlled Release*, vol. 113, no. 3, pp. 216–225, 2006.
- [9] B. Narasimhan and N. A. Peppas, "Molecular analysis of drug delivery systems controlled by dissolution of the polymer carrier," *Journal of Pharmaceutical Sciences*, vol. 86, no. 3, pp. 297–304, 1997.
- [10] I. Katzhendler, A. Hoffman, A. Goldberger, and M. Friedman, "Modeling of drug release from erodible tablets," *Journal of Pharmaceutical Sciences*, vol. 86, no. 1, pp. 110–115, 1997.
- [11] C. Raman, C. Berkland, K. Kim, and D. W. Pack, "Modeling small-molecule release from PLG microspheres: effects of polymer degradation and nonuniform drug distribution," *Journal of Controlled Release*, vol. 103, no. 1, pp. 149–158, 2005.
- [12] R. S. Harland, C. Dubernet, J. P. Benoit, and N. A. Peppas, "A model of dissolution-controlled, diffusional drug release from non-swellable polymeric microspheres," *Journal of Controlled Release*, vol. 7, no. 3, pp. 207–215, 1988.
- [13] M. I. Cabrera, J. A. Luna, and R. J. A. Grau, "Modeling of dissolution-diffusion controlled drug release from planar polymeric systems with finite dissolution rate and arbitrary drug loading," *Journal of Membrane Science*, vol. 280, no. 1-2, pp. 693–704, 2006.
- [14] T. Higuchi, "Mechanism of sustained-action medication. Theoretical analysis of rate released of solid drugs dispersed in solid matrices," *Journal of Pharmaceutical Sciences*, vol. 52, no. 12, pp. 1145–1149, 1963.
- [15] B. B. Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman and Co., San Francisco, Calif, USA, 1982.
- [16] D. Stauffer and H. E. Stanley, *From Newton to Mandelbrot*, Academic Press, New York, NY, USA, 1996.
- [17] G. V. Kozlov and G. E. Zaikov, *Fractals and Local Order in Polymeric Materials*, Nova Science, New York, NY, USA, 2001.
- [18] V. U. Novikov and G. V. Kozlov, "Structure and properties of polymers in terms of the fractal approach," *Russian Chemical Reviews*, vol. 69, no. 6, pp. 523–549, 2000.
- [19] P. L. Ritger and N. A. Peppas, "A simple equation for description of solute release I. Fickian and non-Fickian release from non-swellable devices in the form of slabs, spheres, cylinders or discs," *Journal of Controlled Release*, vol. 5, no. 1, pp. 23–36, 1987.
- [20] N. A. Peppas and J. J. Sahlin, "A simple equation for the description of solute release—III. Coupling of diffusion and relaxation," *International Journal of Pharmaceutics*, vol. 57, no. 2, pp. 169–172, 1989.
- [21] T. Alfrey, E. F. Gurnee, and W. G. Lloyd, "Diffusion in glassy polymers," *Journal of Polymer Science Part C*, vol. 12, pp. 249–261, 1966.
- [22] S. Popescu, *Actual Issues in the Physics of Self-Structured Systems*, Tehnopress, Iasi, Romania, 2003.
- [23] K. Kosmidis, P. Argyrakakis, and P. Macheras, "Fractal kinetics in drug release from finite fractal matrices," *Journal of Chemical Physics*, vol. 119, no. 12, pp. 6373–6377, 2003.
- [24] R. Villalobos, S. Cordero, A. Maria Vidales, and A. Dominguez, "In silico study on the effects of matrix structure in controlled drug release," *Physica A*, vol. 367, pp. 305–318, 2006.

- [25] R. Villalobos, A. M. Vidales, S. Cordero, D. Quintanar, and A. Dominguez, "Monte carlo simulation of diffusion-limited drug release from finite fractal matrices," *Journal of Sol-Gel Science and Technology*, vol. 37, no. 3, pp. 195–199, 2006.
- [26] R. Villalobos, A. Dominguez, A. Ganem, A. M. Vidales, and S. Cordero, "One-dimensional drug release from finite Menger sponges: in silico simulation," *Chaos, Solitons and Fractals*, vol. 42, no. 5, pp. 2875–2884, 2009.
- [27] P. Costa and J. M. Sousa Lobo, "Modeling and comparison of dissolution profiles," *European Journal of Pharmaceutical Sciences*, vol. 13, no. 2, pp. 123–133, 2001.
- [28] L. Nottale, *Fractal Space-Time and Microphysics: Towards a Theory of Scale Relativity*, World Scientific Publishing, Singapore, 1993.
- [29] L. Nottale, *Scale Relativity and Fractal Space-Time—A New Approach to Unifying Relativity and Quantum Mechanics*, Imperial College Press, London, UK, 2011.
- [30] L. Nottale, "Fractals and the quantum theory of space time," *International Journal of Modern Physics A*, vol. 4, no. 19, pp. 5047–5117, 1989.
- [31] N. A. Peppas, "Analysis of Fickian and non-Fickian drug release from polymers," *Pharmaceutica Acta Helveticae*, vol. 60, no. 4, pp. 110–111, 1985.
- [32] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics*, Butterworth Heinemann, Oxford, UK, 2nd edition, 1987.
- [33] J. F. Gouyet, *Physique et Structures Fractals*, Masson, Paris, France, 1992.
- [34] M. S. El Naschie, O. E. Röessler, and I. Prigogine, *Quantum Mechanics, Diffusion and Chaotic Fractals*, Elsevier, Oxford, UK, 1995.
- [35] P. Weibel, G. Ord, and O. E. Rösler, *Space Time Physics and Fractality*, Springer, Dordrecht, The Netherlands, 2005.
- [36] M. Agop, N. Forna, I. Casian-Botez, and C. Bejenariu, "New theoretical approach of the physical processes in nanostructures," *Journal of Computational and Theoretical Nanoscience*, vol. 5, no. 4, pp. 483–489, 2008.
- [37] I. Casian-Botez, M. Agop, P. Nica, V. Paun, and G. V. Munceleanu, "Conductive and convective types behaviors at nano-time scales," *Journal of Computational and Theoretical Nanoscience*, vol. 7, no. 11, pp. 2271–2280, 2010.
- [38] S. Bacaita, C. Uritu, M. Popa, A. Uliniuc, C. Peptu, and M. Agop, "Drug release kinetics from polymer matrix through the fractal approximation of motion," *Smart Materials Research*, vol. 2012, Article ID 264609, 8 pages, 2012.
- [39] D. Magop, S. Bacaita, C. Peptu, M. Popa, and M. Agop, "Non-differentiability at mesoscopic scale in drug release processes from polymer microparticles," *Materiale Plastice*, vol. 49, no. 2, pp. 101–105, 2012.
- [40] F. Bowman, *Introduction to Elliptic Functions: With Applications*, Dover, New York, NY, USA, 1961.
- [41] M. Toda, *Theory of Nonlinear Lattices*, Springer, Berlin, Germany, 1989.
- [42] A. J. Lichtenberg, *Phase-Space Dynamics of Particle*, John Wiley and Sons, New York, NY, USA, 1969.

Research Article

Application of the Poor Man's Navier-Stokes Equations to Real-Time Control of Fluid Flow

James B. Polly¹ and J. M. McDonough²

¹ Department of Civil Engineering, City College of the City University of New York,
New York, NY 10031, USA

² Departments of Mechanical Engineering and Mathematics, University of Kentucky,
Lexington, KY 40506, USA

Correspondence should be addressed to James B. Polly, jamespolly@gmail.com

Received 3 May 2012; Revised 17 July 2012; Accepted 17 July 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 J. B. Polly and J. M. McDonough. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Control of fluid flow is an important, underutilized process possessing potential benefits ranging from avoidance of separation and stall on aircraft wings to reduction of friction in oil and gas pipelines to mitigation of noise from wind turbines. But the Navier-Stokes (N.-S.) equations, whose solutions describe such flows, consist of a system of time-dependent, multidimensional, nonlinear partial differential equations (PDEs) which cannot be solved in real time using current computing hardware. The poor man's Navier-Stokes (PMNS) equations comprise a discrete dynamical system that is algebraic—hence, easily (and rapidly) solved—and yet which retains many (possibly all) of the temporal behaviors of the PDE N.-S. system at specific spatial locations. Herein, we outline derivation of these equations and discuss their basic properties. We consider application of these equations to the control problem by adding a control force. We examine the range of behaviors that can be achieved by changing this control force and, in particular, consider controllability of this (nonlinear) system *via* numerical experiments. Moreover, we observe that the derivation leading to the PMNS equations is very general and may be applied to a wide variety of problems governed by PDEs and (possibly) time-delay ordinary differential equations such as, for example, models of machining processes.

1. Introduction

Given the ever-increasing capabilities of computers and electronic technology, the ability to control a fluid with a range of actuators is already being investigated [1–5] both in laboratory experiments and in numerical simulations. State-of-the-art industrial applications do not exhibit the full potential of control that researchers may envision. Fortunately, it is likely only a matter of time before phenomena such as separation on airfoils, effects of friction factors,

and mitigation of turbulence effects are able to be attenuated or amplified to achieve desired outcomes. For example, ability to maintain laminar flow over an airfoil is clearly important, yet not only maintaining, but regaining laminar behavior from a separated flow *via* real-time control is most desirable. In the petroleum industry, the speed at which oil may be pumped is always subject to frictional losses in the pipelines and pumps. Reducing friction factors through real-time control adds to the general efficiency of such systems. On the other end of the power generation spectrum, control of blade interaction (e.g., pitch and rotation rate) with ambient wind can help reduce chatter and other efficiency losses as well as noise generation in wind farms.

Control schemes may be open loop or closed loop, having different methods of actuation or correction and different algorithms to provide instruction and commands—largely depending on whether open or closed. At present, some methods include microelectromechanical systems (MEMSs) sensors and actuators (as in [2]), local blowing and suction controllers (see, e.g., [4] and [6]), microscale flaps (e.g., [5]), controlled waving plates (as in Techet et al. [7]), and many others. A review of feedback-based methods is provided in Moin and Bewley [8]. It is not our goal to present a comprehensive fluid flow control system herein, or to restrict the study to a particular type of sensor or actuator. Rather, we focus on control of turbulence generically and consider application of forces to a low-order model of the N.-S. equations, the poor man's Navier-Stokes (PMNS) equations, to accomplish this.

The notion of turbulence control has been researched for roughly the past 20 years, with a substantial mathematical formalism contributed by Abergel and Temam [9] in 1990. Though numerical simulations of turbulence were becoming more feasible at that time, the computing capability of even current modern supercomputers does not allow direct numerical simulations (DNSs) at the smallest of length scales for practical flow geometries and Reynolds numbers. Furthermore, it is well known that analytical solutions of the Navier-Stokes equations are still very much an open problem. Even with the increased processor speed and multicore parallel computing available to researchers today, wall functions and nonphysical subgrid-scale (SGS) quantities are still used to obtain solutions. In order to control turbulent behaviors in real time, speed and simplicity are essential. Furthermore, the amount of (physical) correction and actuation needed in order to produce the desired outcome should be easily determined and provided.

We propose herein the direct modeling of flow velocities through a discrete dynamical system (DDS) with the addition of an adjustable body—or control—force to achieve a desired system response. The behavior of this system has been shown to exhibit many—possibly all—of the temporal behaviors found in the full (PDE) N.-S. system at specific spatial locations. Through a straightforward process, the full N.-S. system is reduced to a coupled discrete dynamical system of equations resembling the much studied logistic map ([10]). It has been shown in 2D (by McDonough and Huang [11]) that the coupled system undergoes the usual bifurcation sequence exhibited by the one-dimensional map, and furthermore that the coupled 2D system contains additional regimes which are not possessed by one-dimensional maps and at the same time retains behaviors found in N.-S. type differential systems. In 3D, the DDS appears to be equally suitable for use as a velocity model (see [12]). The system has been shown to be a practical and realistic model of SGS velocities for large-eddy simulation as well as for the complete velocity. The system of PMNS equations can be fit to local velocimetry measurements of flow over a backward-facing step (see [13]); thus, the system response is sufficiently sophisticated to reproduce laboratory observations. Computationally, the equations are algebraic, and thus very efficiently evaluated, lending themselves naturally for application to control of small-scale turbulence.

Using a DDS to model naturally occurring phenomena is not a new idea—ecologist Robert May constructed the logistic map to model population dynamics [10]. Furthermore, the work of Lorenz [14] led to a “low-order” model consisting of ordinary differential equations (ODEs) *via* a procedure similar to what is employed herein. The seminal work of Ruelle and Takens [15] viewed the N.-S. equations as a dynamical system capable of generating chaotic, yet deterministic, solutions associated with a “strange attractor.” It has been generally accepted for at least the past decade that the description of turbulence as a deterministically chaotic system is justified (see, e.g., Frisch [16]). Thus, it is from this treatment of the N.-S. equations as a dynamical system that we consider the problem of turbulence control *via* a closely related DDS.

To implement a control force in the system of PMNS equations, we add a body force to the right-hand side of the N.-S. equations. Since the PMNS equations model local velocities directly, a positive constant control force generally increases the value of the velocity in the respective direction. From studies such as [6], it is apparent that wall-bounded turbulence can be controlled *via* suction and blowing in the direction normal to the wall. The study herein proposes a numerically analogous approach by varying the value of this additional control force, thus adding to, or taking from, the velocity component magnitude generated by the PMNS equations corresponding to wall-normal directions.

Turbulent flows are known to have relatively high dimension for even low Reynolds number due to the spatial and temporal scales over which they occur (see [8]). Furthermore, without considering variable control forces, the PMNS dynamical system contains nine bifurcation parameters. The codimension of the system is therefore sufficiently high that analytical approaches to studying its behavior quickly become intractable. Though it has already been shown that the PMNS system is capable of chaotic behavior (see [11, 12]), it is worthwhile to mention that chaotic attractors are known to exist within systems with codimension as low as one (e.g., Hopf bifurcations of the N.-S. equations). The system is therefore very complex, and though previous studies have shed light on some of its features, this investigation continues an ongoing effort to explore the gamut of potentially useful characteristics it possesses.

In the following, sections we will outline the derivation of the PMNS equations beginning with the full PDE system of equations. The behavior of the DDS in general is the subject of continuous research, a precis of which will be included herein. We then present regime maps illustrating the types of behaviors that the PMNS equations are capable of reproducing and the corresponding bifurcation parameter and body force values for each regime. The aforementioned system complexity due to high codimension motivates the creation of the regime maps. In doing so, the system behavior is understood throughout the domain of bifurcation parameter space. From these regime maps, we deduce appropriate control forces and then implement these while iterating the PMNS equations to demonstrate controllability of the system toward a desired behavior. Time series of these results are shown.

2. Analysis

Here, we describe the treatment applied to the N.-S. equations by which we derive the PMNS equations. After the derivation, we present a brief discussion of the features that are common to PMNS and full N.-S. equations, specifically the symmetry between equations and highly coupled nature of the full PDE system. Then we discuss the extension of these equations to a control context.

2.1. Derivation of PMNS Equations

We begin with the incompressible N.-S. equations,

$$\mathbf{U}_t + \mathbf{U} \cdot \nabla \mathbf{U} = -\nabla P + \nu \Delta \mathbf{U}, \quad \forall (\mathbf{x}, t) \in \Omega \times (t_0, t_f], \quad \Omega \subseteq \mathbb{R}^3, \quad (2.1)$$

and the divergence-free constraint,

$$\nabla \cdot \mathbf{U} = 0, \quad (2.2)$$

so that there are as many equations as necessary to obtain the velocity components and pressure P . In (2.1), the velocity vector in the 3D domain Ω is given as $\mathbf{U} = (u, v, w)^T$; ν is the kinematic viscosity; the Laplace and gradient operators are given as Δ and ∇ , respectively, within the appropriate coordinate system. The t subscript denotes partial differentiation with respect to time; t_0 and t_f are initial and final times, respectively.

As is frequently used in theoretical analysis of N.-S. equations, we employ a Leray projection to a divergence-free subspace of solutions, thus eliminating the pressure gradient in (2.1). We refer the reader to Constantin and Foias [17] or Foias et al. [18] for details. After the typical scaling of dependent and independent variables and Leray projection, the 3D N.-S. equations take the dimensionless form

$$u_t + uu_x + vv_y + ww_z = \frac{1}{\text{Re}} \Delta u, \quad (2.3)$$

$$v_t + uv_x + vv_y + ww_z = \frac{1}{\text{Re}} \Delta v, \quad (2.4)$$

$$w_t + uw_x + vv_y + ww_z = \frac{1}{\text{Re}} \Delta w, \quad (2.5)$$

which is a 3D system of Burgers' equations. The x , y , and z subscripts denote partial differentiation with respect to the spatial variables, and Re is the Reynolds number: $\text{Re} = UL/\nu$, with U and L denoting the appropriate velocity and length scales, respectively.

Consistent with the mathematical understanding of the N.-S. equations (see, e.g., [17]), we begin with representing velocity components in Fourier space:

$$u(\mathbf{x}, t) = \sum_{\mathbf{k}=-\infty}^{\infty} a_{\mathbf{k}}(t) \varphi_{\mathbf{k}}(\mathbf{x}), \quad (2.6)$$

$$v(\mathbf{x}, t) = \sum_{\mathbf{k}=-\infty}^{\infty} b_{\mathbf{k}}(t) \varphi_{\mathbf{k}}(\mathbf{x}), \quad (2.7)$$

$$w(\mathbf{x}, t) = \sum_{\mathbf{k}=-\infty}^{\infty} c_{\mathbf{k}}(t) \varphi_{\mathbf{k}}(\mathbf{x}), \quad (2.8)$$

with the 3D wavevector $\mathbf{k} \equiv (k_1, k_2, k_3)^T$. We assume that the tensor product basis set $\{\varphi_{\mathbf{k}}\}$ is complete in $L^2(\Omega)$ and orthonormal. Furthermore, we note that construction of such a basis

could be difficult for some computational purposes—for example, in constructing proper orthogonal decomposition (POD) algorithms. It is, however, not necessary to do so for the study herein (which distinguishes our approach from POD). For brevity, the following details of the derivation will be provided for only the x -momentum equation, with treatment of the y - and z -momentum equations being identical.

The Galerkin procedure is applied to (2.3), (2.4), (2.5) resulting in an infinite system of ODEs for the Fourier coefficients contained within. Substitution of (2.6), (2.7), and (2.8) into (2.3) leads to

$$\begin{aligned} \sum_1 \dot{a}_1 \varphi_1 + i \sum_{l,m} (l_1 + m_1) a_l a_m \varphi_l \varphi_m + i \sum_{l,m} (l_2 + m_2) a_l b_m \varphi_l \varphi_m \\ + i \sum_{l,m} (l_3 + m_3) a_l c_m \varphi_l \varphi_m = -\frac{1}{\text{Re}} \left[\sum_1 (l_1^2 + l_2^2 + l_3^2) a_1 \varphi_1 \right]. \end{aligned} \quad (2.9)$$

Considering the equations without the imaginary factor (which can be removed *via* methods analogous to those of elementary ODE analysis), inner products of the above equation and each basis function are formed. Employing orthonormality then reduces the x -momentum equation to

$$\dot{a}_k + \sum_{l,m} A_{lm,k}^{(1)} a_l a_m + \sum_{l,m} A_{lm,k}^{(2)} a_l b_m + \sum_{l,m} A_{lm,k}^{(3)} a_l c_m = -\frac{\eta^{(a)}}{\text{Re}} |\mathbf{k}|^2 a_k, \quad (2.10)$$

where $\mathbf{k} = -\infty, \dots, \infty$. The factor $\eta^{(a)}$ is a normalization constant arising from the fact that derivatives of basis functions may not possess the same normalization as the functions themselves. Similarly, there are $\eta^{(b)}$ and $\eta^{(c)}$ terms in the omitted y - and z -equations, respectively. The $A_{lm,k}^{(i)}$ factors, with $i = 1, 2, 3$, are the Galerkin triple products, for example,

$$A_{lm,k}^{(1)} = (l_1 + m_1) \int \varphi_l \varphi_m \varphi_k dx, \quad (2.11)$$

with analogous terms $B_{lm,k}^{(i)}$ and $C_{lm,k}^{(i)}$ used in the y - and z -equations, respectively.

Removing all but a single arbitrary wavevector in (2.10) results in a single equation, with analogous results for the y - and z -momentum equations also shown:

$$\dot{a} = -A^{(1)} a^2 - A^{(2)} ab - A^{(3)} ac - \frac{\eta^{(a)}}{\text{Re}} |\mathbf{k}|^2 a, \quad (2.12)$$

$$\dot{b} = -B^{(1)} ba - B^{(2)} b^2 - B^{(3)} bc - \frac{\eta^{(b)}}{\text{Re}} |\mathbf{k}|^2 b, \quad (2.13)$$

$$\dot{c} = -C^{(1)} ca - C^{(2)} cb - C^{(3)} c^2 - \frac{\eta^{(c)}}{\text{Re}} |\mathbf{k}|^2 c, \quad (2.14)$$

with wavevector subscript notation suppressed, and dot ($\dot{\cdot}$) notation indicating ordinary differentiation with respect to time. The above equations are a system of three nonlinear

ODEs. We discretize these using an explicit single-step forward Euler integration method in time and denote the time-step parameter as τ . This leads to

$$a^{n+1} = a^n - \tau \left[A^{(1)}(a^n)^2 + A^{(2)}a^n b^n + A^{(3)}a^n c^n + \frac{\eta^{(a)}}{\text{Re}} |\mathbf{k}|^2 a^n \right], \quad (2.15)$$

$$b^{n+1} = b^n - \tau \left[B^{(1)}b^n a^n + B^{(2)}(b^n)^2 + B^{(3)}b^n c^n + \frac{\eta^{(b)}}{\text{Re}} |\mathbf{k}|^2 b^n \right], \quad (2.16)$$

$$c^{n+1} = c^n - \tau \left[C^{(1)}c^n a^n + C^{(2)}c^n b^n + C^{(3)}(c^n)^2 + \frac{\eta^{(c)}}{\text{Re}} |\mathbf{k}|^2 c^n \right]. \quad (2.17)$$

Again for brevity, we omit analysis of the y - and z -direction equations and rearrange the x -momentum equation as

$$a^{n+1} = \tau A^{(1)} a^n \left(\frac{1 - (\eta^{(a)} / \text{Re}) |\mathbf{k}|^2 \tau}{\tau A^{(1)}} - a^n \right) - \tau A^{(2)} a^n b^n - \tau A^{(3)} a^n c^n. \quad (2.18)$$

Recalling the logistic map

$$a^{n+1} = \beta a^n (1 - a^n), \quad (2.19)$$

we note that to recover this form in the above equation requires

$$\frac{1 - (\eta^{(a)} / \text{Re}) |\mathbf{k}|^2 \tau}{\tau A^{(1)}} = 1, \quad (2.20)$$

which implies

$$1 - \frac{\eta^{(a)}}{\text{Re}} |\mathbf{k}|^2 \tau = \tau A^{(1)}. \quad (2.21)$$

Having recovered the form of the logistic map *via* the requirement in (2.20), it is observed that as $\text{Re} \rightarrow \infty$, the left-hand side of (2.21) approaches unity from below. As this term approaches one, the first term on the right-hand side of (2.18) becomes the logistic map, as shown in the following expressions.

Thus, the advanced time step equations can be expressed as

$$a^{n+1} = \left(1 - \frac{\eta^{(a)}}{\text{Re}} |\mathbf{k}|^2 \tau \right) a^n (1 - a^n) - \tau A^{(2)} a^n b^n - \tau A^{(3)} a^n c^n, \quad (2.22)$$

$$b^{n+1} = \left(1 - \frac{\eta^{(b)}}{\text{Re}} |\mathbf{k}|^2 \tau \right) b^n (1 - b^n) - \tau B^{(1)} b^n a^n - \tau B^{(3)} b^n c^n, \quad (2.23)$$

$$c^{n+1} = \left(1 - \frac{\eta^{(c)}}{\text{Re}} |\mathbf{k}|^2 \tau \right) c^n (1 - c^n) - \tau C^{(1)} c^n a^n - \tau C^{(2)} c^n b^n. \quad (2.24)$$

Now define the following bifurcation parameters:

$$\beta_1 = 1 - \frac{\eta^{(a)}}{\text{Re}} |\mathbf{k}|^2 \tau, \quad (2.25)$$

$$\beta_2 = 1 - \frac{\eta^{(b)}}{\text{Re}} |\mathbf{k}|^2 \tau, \quad (2.26)$$

$$\beta_3 = 1 - \frac{\eta^{(c)}}{\text{Re}} |\mathbf{k}|^2 \tau, \quad (2.27)$$

$$\gamma_{12} = \tau A^{(2)} \quad \gamma_{13} = \tau A^{(3)}, \quad (2.28)$$

$$\gamma_{21} = \tau B^{(1)} \quad \gamma_{23} = \tau B^{(3)}, \quad (2.29)$$

$$\gamma_{31} = \tau C^{(1)} \quad \gamma_{32} = \tau C^{(2)}. \quad (2.30)$$

Then the equations of motion collapse to the algebraic system

$$a^{(n+1)} = \beta_1 a^{(n)} (1 - a^{(n)}) - \gamma_{12} a^{(n)} b^{(n)} - \gamma_{13} a^{(n)} c^{(n)}, \quad (2.31)$$

$$b^{(n+1)} = \beta_2 b^{(n)} (1 - b^{(n)}) - \gamma_{21} b^{(n)} a^{(n)} - \gamma_{23} b^{(n)} c^{(n)}, \quad (2.32)$$

$$c^{(n+1)} = \beta_3 c^{(n)} (1 - c^{(n)}) - \gamma_{31} c^{(n)} a^{(n)} - \gamma_{32} c^{(n)} b^{(n)}. \quad (2.33)$$

Here, we employ parentheses with the time-level indices to emphasize that these equations comprise an iterated map. Following [11], and in deference to Frisch [16], we term (2.31), (2.32), and (2.33) the “poor man’s Navier-Stokes (PMNS) equations.”

For the study performed herein, we consider the isotropic case, which means that the γ_{ij} values are all equal to one another, and the β_i values are also equal to one another. Note that if $\gamma_{ij} = 0$, for all i, j , then we recover the logistic map (2.19). Also notice that as Re approaches infinity, the value of β_i approaches unity from below. In the DDS of PMNS equations, we limit the starting values between zero and one. Noting that the function $f(a) \equiv a(1 - a)$ has a maximum range of $1/4$ for $a \in [0, 1]$, we rescale the β_i values by a factor of 4 (as done by May [10]) to obtain the typical scaling for DDSs. Thus, as in the logistic map, the range of β_i values is between zero and four. It is well known that as the bifurcation parameter of the logistic map increases, the system behavior becomes increasingly chaotic. This is in accord with our formulation of the bifurcation parameter β , which is a function of the Reynolds number. As Re increases, so does the value of β , and the system behaves correspondingly chaotically.

Note the similar structure of all three components of (2.31), (2.32), and (2.33). This feature is shared with the full N.-S. system of PDEs. Moreover, each equation is nonlinear, and this property is not shared with other often-studied DDSs shown to exhibit chaotic behaviors, for example, discretization of the Lorenz equations (see [14]) or the related 2D DDS due to Hénon [19]. An in-depth investigation of the coupled nature of the PMNS equations is not performed herein, yet may be found in [11, 12]. At present, it is simply important to note that “symmetrical” structure and nonlinearity are desirable qualities preserved in the derivation of the PMNS system from the N.-S. equations.

2.2. Control of PMNS Equations

In order to carry out the intended computational experiments, we must begin with valid ranges for the bifurcation parameters β_i and γ_{ij} . Consideration of isotropy, mentioned above, greatly reduces the space of bifurcation parameter values to consider. Instead of a nine-dimensional space of parameters, we set all β_i , $i = 1, 2, 3$, values equal to one another. The bifurcation parameters γ_{ij} , $i, j = 1, 2, 3$ and $i \neq j$, are also set equal. From the studies contained in [11, 12], we know which values of bifurcation parameters produce behavior that is of interest in the context of this study of turbulence control and collect data beginning in this subset of bifurcation parameter space.

The following section contains results collected *via* numerical experiments allowing bifurcation parameters to vary while simultaneously applying a “control force.” The control force is simply a constant added to the right-hand side of the PMNS equations. This is analogous to adding a force at specified wavevectors as is commonly done in DNS and may be viewed as suction or blowing, akin to the study performed by Moin and Bewley [6]. With a_{CF} , b_{CF} , and c_{CF} , the control—or body—forces for each equation, and including the simplification of isotropy, we will solve the following DDS:

$$a^{(n+1)} = \beta a^{(n)} (1 - a^{(n)}) - \gamma a^{(n)} b^{(n)} - \gamma a^{(n)} c^{(n)} + a_{CF}, \quad (2.34)$$

$$b^{(n+1)} = \beta b^{(n)} (1 - b^{(n)}) - \gamma b^{(n)} a^{(n)} - \gamma b^{(n)} c^{(n)} + b_{CF}, \quad (2.35)$$

$$c^{(n+1)} = \beta c^{(n)} (1 - c^{(n)}) - \gamma c^{(n)} a^{(n)} - \gamma c^{(n)} b^{(n)} + c_{CF}. \quad (2.36)$$

3. Results and Discussion

Here, we present results from iteration of the DDS of (2.34), (2.35), and (2.36) for a range of bifurcation parameter values. We will focus on the isotropic case as already noted above. The bifurcation parameters will vary as $\beta \in [0, 4]$, with assigned values of γ (which can typically range from -0.8 to 0.8 , with the range depending on values of β). With varying β (related to Re) and fixed γ , we will allow the control force a_{CF} to vary and observe the range of behavior types which the system of PMNS equations will produce. Following the analysis techniques in [11, 12], we first describe the types of discernible behavior the system exhibits. Our method of categorizing behaviors—or regime types—is *via* processing and analysis of the temporal power spectra. Observing the power spectra of the time series output from the DDS and classifying the behavior types into regimes will allow a comparative study of the system behavior, as well as provide insight into the transitions between the regimes as the bifurcation parameter and control force are varied. Having discriminated the regime types, we further explore the ability to control the time series of the DDS by “turning on” the control force to obtain a desired system behavior. Thus, given a specific behavior type corresponding to a pair of bifurcation parameters, we attempt to control the system by means of the control force a_{CF} .

The majority of the data presented in this work was calculated on a 376-node Dell high-performance computing cluster at the University of Kentucky Computing Center. Regime map calculations were performed in parallel using OpenMP with three cores from a 2.66 GHz Intel Xeon X5650 six-core processor. Less intensive calculations (*i.e.*, calculating a single time series) were performed on a desktop computer with a 3.60 GHz Intel Pentium 4 and are

computed nearly instantaneously. Computations were performed using double precision (64-bit) Fortran. In regime map cases, for each point corresponding to a pair of β and a_{CF} values, the DDS was iterated 1×10^4 times. A radix-2 fast-Fourier transform was applied to the last 2^{12} (or 4096) iterations to produce power spectra for regime identification. In cases where a single time series was computed, the system was iterated 2×10^4 times with the control force activated at iteration number 10^4 . Though the computing resources used to conduct these experiments are somewhat powerful, it is important to observe that the PMNS equations may be iterated $\mathcal{O}(10^5)$ times nearly instantly using a modern laptop computer. Advances in mobile computing hardware and chip technology will soon render the computation of numerous iterations of an algebraic map (such as the PMNS equations) a trivial task for even the smallest and most ubiquitous of devices.

3.1. Regime Maps

As in the studies [11, 12], we present regime maps containing the behavior types—or regimes—which the system of PMNS equations is capable of producing. In the aforementioned investigations, various regime types were distinguished from one another in an automatic and objective fashion. Thus, given a pair of β and γ bifurcation parameters and a value of the control force a_{CF} , the system response can be characterized. Furthermore, from those studies, the range of values for which to allow the control force to vary, given a pair of bifurcation parameters, is known *via* numerical experimentation. The behaviors exhibited by the 3D PMNS equations include all of those behaviors produced by the 1D family of logistic maps and the types of N.-S. regimes found in laboratory experiments. Regime types are distinguished from one another *via* the power spectral density (PSD), and thus their classification requires human definition of power and frequency criteria for which one regime is distinguished from another. Some discussion of this procedure is appropriate here; however, more detail is included in the two aforementioned studies.

Our goal is to distinguish the various behavior types which the DDS will produce, and do so automatically and objectively. Though some insight may be gained by calculating the fractal dimension, Lyapunov exponents, or other statistical quantities (or by constructing bifurcation diagrams), these will not allow automatic regime classification of the chaotic behaviors of which the DDS is capable. (Some comparison of these methods and their effectiveness is given in [12], and we refer the reader to this citation for more detail, though it suffices to say that the method we present is the most effective for our intents and purposes.) Thus, in order to understand the system behavior as bifurcation parameters—including control forces—are varied, we perform computations across the full range of values with $\mathcal{O}(10^{10})$ realizations of the DDS contained within some regime maps. Furthermore, we also compute regime maps over restricted subdomains of bifurcation parameter and control force such that we can clearly study the transitions and localized—in bifurcation parameter space—behaviors. With so many instances of the equations, our means of classifying regime types must be efficient, automatic, and objective.

Identification of the regime type needs to be only qualitative. Thus, the PSD appears to contain sufficient information to permit distinguishing one regime from an inherently different one. The following is a list of the regime types which we have identified in the time series produced by the DDS *via* this technique: (i) steady, (ii) periodic, (iii) periodic with different fundamental frequency, (iv) subharmonic, (v) phase locked, (vi) quasiperiodic, (vii) noisy subharmonic, (viii) noisy phase locked, (ix) noisy quasiperiodic with fundamental

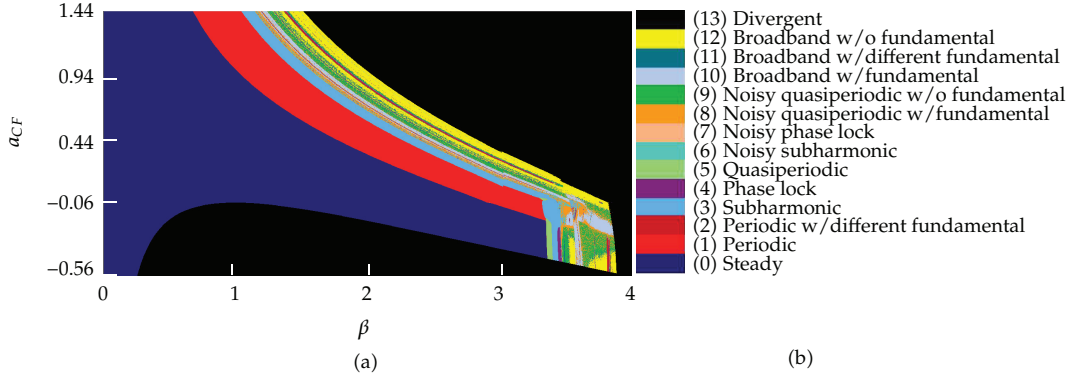


Figure 1: (a) 2501×1251 regime map for $\gamma = -0.05$ with $\Delta\beta = 0.0016 = \Delta a_{CF}$, (b) legend of regime types.

frequency, (x) noisy quasiperiodic without fundamental frequency, (xi) broadband with fundamental frequency, (xii) broadband with different fundamental frequency, (xiii) broadband without fundamental frequency, and (xiv) divergent, as displayed in Figure 1.

Some discussion of the nomenclature and criteria for categorizing the regimes is due. Several of the regimes are referred to as “noisy” along with another description which identifies a distinguishing or primary characteristic of the behavior. This is to imply some broadband features of the PSD along with a salient distinguishing characteristic. Differing from laboratory experiments, the numerical DDS evaluation is of course done entirely within the computer. Thus, there is no “noise” in the sense of the type of signal fluctuation coming from instrumentation or sensors as would be a concern in laboratory experiments. The broadband noise which is observed in the PSDs we consider herein is an actual feature of the DDS. It has been shown in [11, 12] that this noisy behavior exhibits at least a mild sensitivity to initial conditions (SIC) which is known as the “hallmark of chaos” (see, e.g., [20]), thus implying existence of a strange attractor in dissipative systems as considered herein. For further information on the nature of the PMNS equations themselves, the regime types, and the process by which these regimes are distinguished, we defer to the studies in [11, 12].

Figure 1(a) shows the regime map corresponding to a fixed value of $\gamma = -0.05$ for β ranging from zero to four with the forcing term a_{CF} between -0.56 and 1.44 . Other values of γ will also be used, but the results presented herein focus on bifurcation parameter γ near zero. This is not done for simplification; rather it is a choice which permits a more complete investigation of the behaviors achievable *via* control (and is typical in LES runs where the β_i and γ_{ij} are computed rather than assigned). In the studies of [11, 12], no control force terms were used, and the regime maps were computed allowing both β and γ to vary. The types of behavior which the PMNS equations exhibit over the range of β values with γ fixed near zero include nearly all possible regimes, so this choice of γ results in a thorough investigation.

A key of the colors corresponding to each behavior type is presented in Figure 1(b). There are several features of the regime map worth indicating specifically. Attention should be given to the bifurcation sequence which occurs as the β value and the a_{CF} value are increased. For most of the map ($\beta < 3.3$), the sequence corresponds to that of the usual 1D logistic map, (2.19):

$$\text{steady} \longrightarrow \text{periodic} \longrightarrow \text{subharmonic} \longrightarrow \text{chaotic}. \quad (3.1)$$

For β values larger than three, the system begins to show more erratic behavior with a less clearly defined bifurcation sequence. This region of the map for values of $\beta > 3.3$ will be of special interest later as we consider regime maps corresponding to different values of γ . Notice that the regions of the map corresponding to steady and periodic, even including the regions of subharmonic and quasiperiodic, are clearly demarcated. The boundaries between the more noisy and chaotic regime types are less clearly defined and thus complicate the matter of choosing a control parameter value.

After the system bifurcates from subharmonic behavior, as the value of control force and bifurcation parameter are increased, there are several locations where the boundaries between the regimes are not well defined. This result is shared with the 2D and 3D studies of the PMNS equations in [11, 12] and is not unique to the implementation of a control force. The unclear boundary between regimes is what will be referred to as a “fractal region” and is an interesting and important result which is addressed in the two aforementioned investigations. This unclear boundary persists regardless of how finely we are able to resolve bifurcation parameter space, that is, very small steps in the β and a_{CF} directions. We will provide a high-resolution “zoom-in” of fractal and other interesting regions below.

The same type of behavior is observed in Figures 2(a), 2(b), and 2(c), where the same values of β and a_{CF} are allowed to vary with γ fixed at $\gamma = 0.0$, $\gamma = 0.05$, and $\gamma = 0.1$, respectively. We present these regime maps to show the numerous possibilities of controlling the PMNS equations—within a small range of control force values—given a particular pair of bifurcation parameters. To better illustrate this, it is clear from Figure 1, corresponding to $\gamma = -0.05$, that for $\beta > 3.3$ there are at least eight unique regime types available within a small range of β and a_{CF} values. The sequence of bifurcations, as β and/or a_{CF} is increased, is somewhat unstructured. This region for $\beta > 3.3$ becomes much more ordered when γ is held fixed at zero, as shown in Figure 2(a). In this case—as shown in the studies of [11, 12]—the aforementioned bifurcation sequence corresponding to the logistic map (2.19) is sustained across the entire range of β values.

The ordered sequence of bifurcations presented in Figure 2(a) shows that there are well-defined regions for which the values of control force and bifurcation parameter can be determined and thus used in a control scheme. Given a pair of bifurcation parameters, the control force required to achieve a particular behavior type is immediately known from the regime maps presented herein. As the value of γ is increased, as in Figures 2(b) and 2(c), a “window” of periodic behavior appears near $\beta = 3.0$, and for $\beta = 3.5$, there is an interesting bifurcation sequence as β is increased:

$$\text{steady} \rightarrow \text{quasiperiodic} \rightarrow \text{subharmonic} \rightarrow \text{phaselock} \rightarrow \text{chaos}. \quad (3.2)$$

Figure 2(c), however, shows that the aforementioned region of subharmonic behavior is consumed by the adjacent quasiperiodicity and even some “islands” of phase-lock behavior. The change from $\gamma = 0.05$ to $\gamma = 0.1$ shown in Figures 2(b) and 2(c) illustrates the wide variety of achievable behaviors which will be available *via* changing the control force.

To better exhibit the range of control possibilities within a small range of bifurcation parameter and control force, a localized regime map is presented in Figure 3. It should be noted that this regime map is not simply a visual magnification of the corresponding subregion in Figure 1 for which $\beta > 3.3$. Rather, it is produced from a higher-resolution calculation consisting of 1121^2 points with $\Delta\beta = 0.000625 = \Delta a_{CF}$, that is, the bifurcation parameter grid is much finer. Herein, many fractal boundaries exist between the regimes; yet

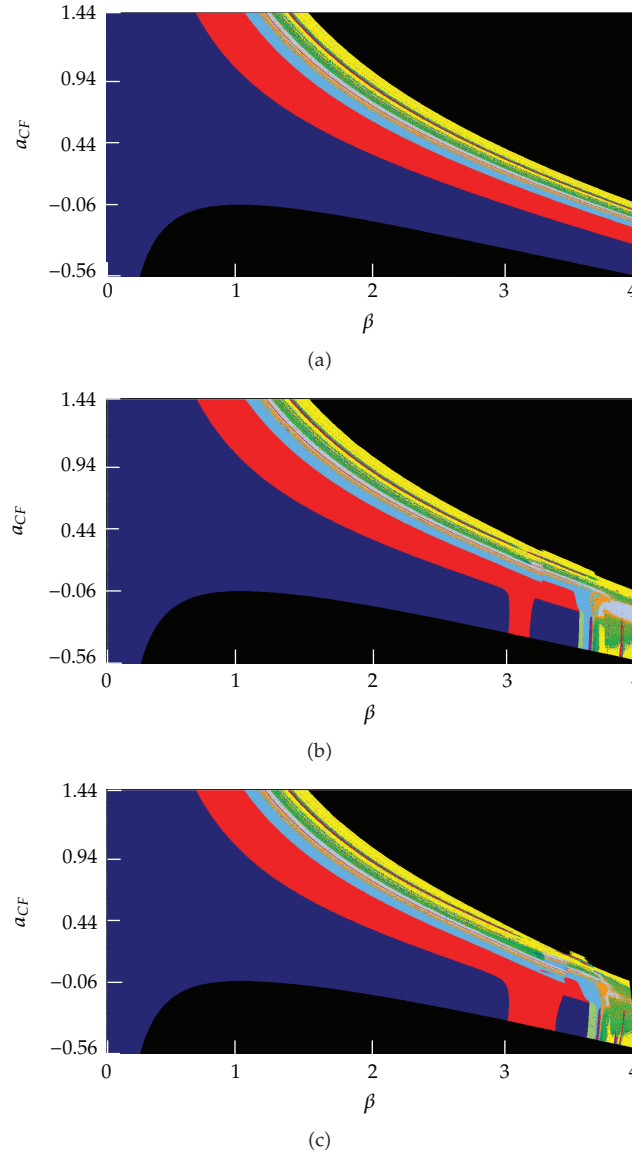


Figure 2: 2501×1251 regime maps (a), (b), and (c) for which $\gamma = 0.00$, $\gamma = 0.05$, and $\gamma = 0.10$, respectively, with $\Delta\beta = 0.0016 = \Delta a_{CF}$.

there are still very clear regions of behavior types, and the SIC characteristic of the PMNS equations does not present a problem with respect to determining a suitable control force. Furthermore, there exist ample regions of noisy behavior which may be useful for triggering turbulent activity when desired (e.g., increased thermal advection, avoidance of boundary-layer separation, etc.).

There is a tendency for regime types to orient themselves into regions aligned along the direction of varying control force (vertically in the regime maps). When attempting to vary the control force to achieve a different regime type, this presents a problem. In these regions,

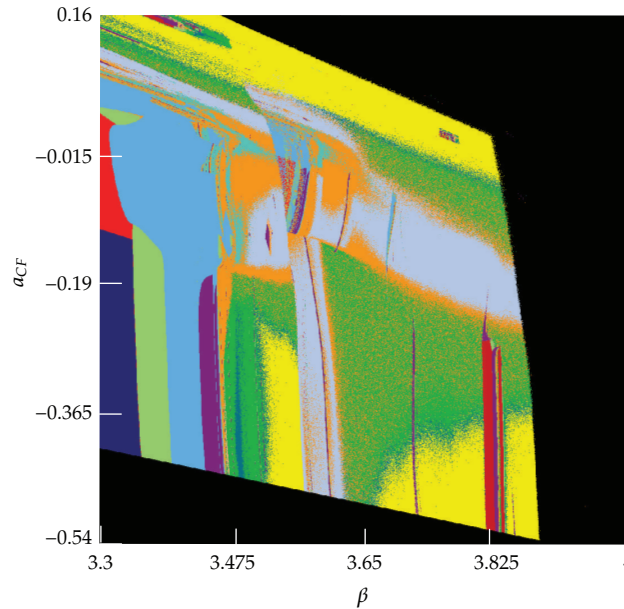


Figure 3: 1121×1121 regime map for $\gamma = -0.05$ with $\Delta\beta = 0.000625 = \Delta a_{CF}$.

shown clearly in Figure 3, adjusting the control force a_{CF} will lead to a bifurcation in behavior due to the “shape” of the regime in bifurcation parameter space. In addition to this problem, a change of control force will often times lead from one chaotic regime to another, or, worse yet, to a nonphysical result such as divergence. Note that this pattern of regime types oriented along the direction of the control force occurs for ranges of β corresponding to chaotic behaviors. That said, the degree to which the regime types are structured in this manner is not consistent for all combinations of bifurcation parameters. The case of Figure 3 considers $\gamma = -0.05$. The tendency for regimes to be structured in this manner depends largely on the value of γ and needs to be studied in further detail. It may in fact be the case that certain values of γ exhibit different bifurcation sequences, and thus control is limited to certain regions of bifurcation parameter space in which more desirable regime combinations and bifurcation sequences may be found.

Despite this obstacle, there are still regions where the control force may be varied to achieve a particular regime with *a priori* knowledge of the appropriate control force for a given set of bifurcation parameters. This is shown in the preceding discussion. From (2.25), (2.26), (2.27), (2.28), (2.29), and (2.30), it is clear that the bifurcation parameters contain physical quantities which govern the behavior of the N.-S. equations, that is, the bifurcation parameters correspond to physical flow scenarios. The regime maps presented herein show that given a set of bifurcation parameters—or a physical scenario—employing a control force will alter the system, allowing for a robust control to nearly any desired outcome.

3.2. Control of Time Series

In this subsection, we use contents of the regime map in Figure 1 to help select bifurcation parameters leading to a chaotic behavior. Given the bifurcation parameters corresponding to

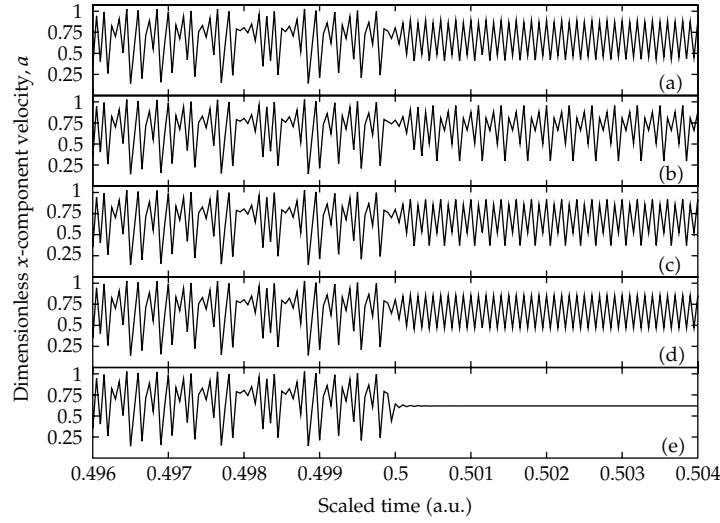


Figure 4: Time series response of the PMNS equations. Control force a_{CF} is changed half way through 2×10^4 iterations with $\beta = 3.3$, $\gamma = -0.05$, and $a_{CF} = 0.17$. Control of broadband w/o fund. freq. to (a) quasiperiodicity, (b) phase lock, (c) subharmonic, (d) periodic, and (e) steady, by setting a_{CF} equal to 0.035, 0.098, 0.06, 0.00, and -0.2 , respectively.

a particular chaotic regime type, we successfully employ a control force to induce bifurcation to other less chaotic regimes. In the context of applying the PMNS equations to a real-time control system, the bifurcation parameters corresponding to a given scenario can, in principle, be computed *via* physical data collected from sensors. Thus, the range of control force values appropriate for achieving the desired behavior type could be easily computed *via* on-board hardware in many specific systems.

Figure 4 shows the time series of the PMNS equations while allowing the value of the control force to change during system evolution. The DDS was iterated 1×10^4 times prior to applying the control force required to obtain the desired system response. Prior to application of the external force, the methods we employ for detecting the regime type distinguish the behavior as broadband without a fundamental frequency. In starting with this regime type, the PMNS equations produce time series which are erratic and chaotic—essentially “white noise.” Controlling the time series of this most chaotic regime is considered herein, though it should be noted that the PMNS equations can be controlled between essentially any regime types. Thus, the results presented in Figure 4 are only a small fraction of the possible control scenarios that can be achieved *via* this technique. We focus on controlling the broadband without fundamental frequency regime to show that our procedure is capable of handling even the most chaotic behaviors.

The regime map in Figure 1 was constructed with $\gamma = -0.05$ and varying values of β and a_{CF} . Though there are many combinations possible, we choose $\beta = 3.3$ and $a_{CF} = 0.17$ so that the behavior of the x -component of velocity is broadband without a fundamental frequency. Many other combinations of β and a_{CF} would have produced similar results; thus, there is nothing special about these values. Given the known behavior of the PMNS equations using these values of β and a_{CF} , we then allow the DDS to evolve additional 1×10^4 iterations after changing the value of a_{CF} . The second value of a_{CF} also comes from the regime map in Figure 1. For example, to produce a quasiperiodic behavior from the chaotic broadband,

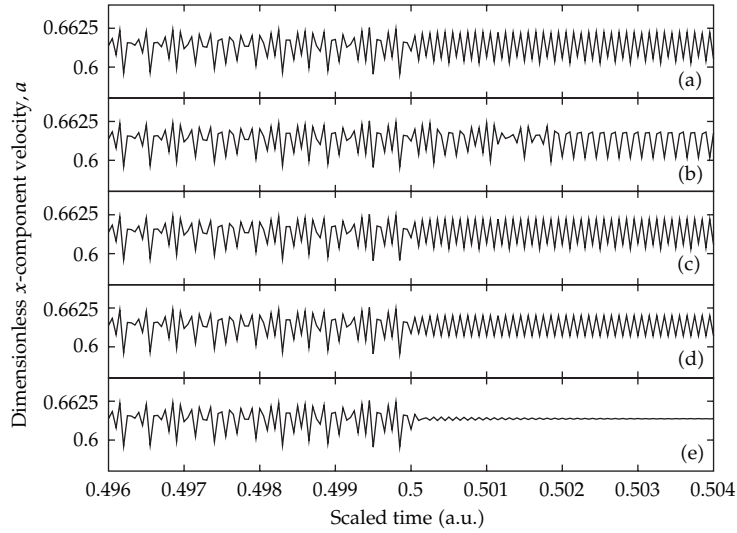


Figure 5: Time series response (x -direction) of the PMNS equations with y -direction control force b_{CF} is changed half way through 2×10^4 iterations with $\beta = 2.552$, $\gamma = -0.05$, and $b_{CF} = 0.5872$. Control of broadband w/o fund. freq. to (a) quasiperiodicity, (b) phase lock, (c) subharmonic, (d) periodic, and (e) steady, by setting b_{CF} equal to 0.03888, 0.548, 0.38, 0.25, and 0.12, respectively.

the value of a_{CF} is changed to 0.035. Again, there are many other values to which the control force can be changed so that the DDS will produce quasiperiodicity from the broadband signal, but a value must be assigned *a priori*.

The ability to control the x -direction component of velocity is clearly demonstrated in Figure 4, and the same would also be true of the y - and z -direction components. Ability to control a velocity component with wall-normal direction forcing as has been experimentally investigated *via* blowing and suction (see [4, 6]) and is easily simulated with the PMNS equations, the results of which are presented in Figure 5. Regime maps as constructed in Figures 2 and 3 can be constructed for the case of varying wall-normal control force b_{CF} , as performed in previous studies (see [12] for details). From those regime maps, appropriate values of the control force b_{CF} are obtained and used to calculate the time series shown in Figure 5.

It is clear from both Figures 4 and 5 that the system responds very quickly to a change in control force and does not require substantial time to stabilize. This is in contrast with the initial behavior of the DDS, where many iterations may be required before initial transient behavior decays and the salient behavior of the system can be determined (as explained in [11, 12]). Not only can velocity be controlled within very few iterations, but the efficiency of the PMNS equations allows these sorts of computations to be performed quickly in wall-clock time.

A noticeable difference between control of the PMNS system with streamwise control force (Figure 4) and wall-normal control force (Figure 5) is the increased time required to stabilize in the cases of phase-locked and steady behaviors. Notice that among each behavior type, the response to change in streamwise control force (Figure 4) requires similar times to stabilize. The wall-normal control force case differs from this in that the time required for the phase-locked and steady behaviors to stabilize is slightly longer than in other cases. This difference is quite negligible from the perspective of system control as only few additional

iterations are required. That having been said, this observation may allude to more interesting characteristics of the phase lock regime which has been previously studied in [11, 12] and maybe important from the perspective of dynamical systems analysis.

Significant research has been conducted in recent years to control dynamical systems (e.g., [21–23]). The majority of this work, however, does not pertain to systems which model physics, especially fluid motion. The results presented herein are unique in two ways. First, the dynamical system modeled is directly derived from the equations of fluid motion. Second, the control parameter analogously corresponds to a method which is frequently used in laboratory experiments of turbulence control, namely, suction and blowing.

Though the PMNS dynamical system models the motions of fluids, the attempts to control this system may be compared with other studies in which a more general dynamical system is controlled. In [22], a 4D chaotic system is controlled by means of recursive backstepping, and it is shown that chaotic behavior may be eradicated in a similar fashion as in the present study. Control of the system from chaotic to stable states of steady or periodic behavior was exhibited, and the system response times were relatively short. A control problem of a different nature is studied in [23], where a finite-time control scheme is implemented to synchronize two different dynamical systems with chaotic behavior. Effective control is demonstrated as the trajectories of one system converge to those of the other. Furthermore, convergence occurs rapidly once control is implemented.

Results from those general dynamical systems previously discussed have much in common with the time series results presented in Figures 4 and 5. The primary shared features are the ability to control the system to periodic or other stable trajectories. Moreover, rapid system response is observed in each. With the exception of the transition to phase-locked and steady behaviors in the case using a wall-normal control force, the response times herein only span a few iterations. These are expected to be quicker than the results contained in [22, 23], where an algorithm depending on the system output is implemented. No such algorithm is yet developed for the PMNS equations, and the control is achieved with *a priori* knowledge of the appropriate control force required for a given set of bifurcation parameters from the regime maps shown in Figures 2 and 3. In the following section, we discuss a simple outline for a potential control scheme.

3.3. A Control Algorithm

Here, we present one (of probably many) possible algorithms by means of which the PMNS equations might be used to control turbulent fluid flow. It is important to observe that these equations can be rapidly evaluated, as already emphasized, and their simple algebraic structure is such as to permit easy implementations on microprocessors.

For simplicity, we assume that the desired flow regime is known, although one can easily envision situations where this might not be true. Then to achieve this desired behavior, carry out the following steps.

- (1) Collect flow data, possibly (but not necessarily) at many locations.
- (2) Use these data to construct PMNS equation bifurcation parameters.
- (3) Run the PMNS equations and use the regime map algorithm to identify the current physical flow regime.

- (4) If regime is the desired one, return to 1. If it is not, begin systematic perturbation of control force parameters with PMNS equation and regime map calculations until desired flow regime is obtained.
- (5) Convert PMNS equation control force to physical one and send corresponding signal to controller.

We remark that there are a number of details still to be treated in this control procedure, not the least of which is the need to perform implementations on microprocessors. But in this regard, the only part of the algorithm involving more than a few lines of code is the regime map program, and even this is relatively small and should easily fit on modern microprocessors. The more difficult parts of the algorithm are physically collecting sufficient data to build PMNS equation bifurcation parameters with adequate accuracy to be of use, and similarly inverting the process to convert PMNS equation control parameters to physical ones. Clearly, these issues are system dependent; some work is currently in progress to permit investigation of these ideas for a particularly simple system.

4. Summary and Conclusions

In this study, we have derived a discrete dynamical system directly from the 3D, incompressible Navier-Stokes equations and investigated the use of these equations within the context of turbulence control. We have studied the possible behaviors which the PMNS equations may exhibit and included computational results for $\mathcal{O}(10^{10})$ instances. Having *a priori* knowledge of the system behavior, we are able to select bifurcation parameters and control forces such that time series of the PMNS equations are controlled during their evolution. In doing this, we demonstrated the ability of the equations to control a velocity field, and we propose that due to their efficiency, the PMNS equations are well suited for turbulence control.

Derivation of the PMNS equations is, in principle, quite general and can be applied to a wide variety of problems governed by PDEs and (possibly) time-delay ODEs (e.g., models of machining processes). The derivation does not introduce any nonphysical quantities or attempt to model any physical ones. The PMNS equations have been shown to have significant potential as a “synthetic velocity” model, and herein, the ability to manipulate velocity fields across a wide variety flow behaviors was shown. The subject of ongoing investigations and future work will be implementation of the PMNS equations into hardware for use within a real-time control system similar to that described above.

Nomenclature

$\mathbf{U} = (u, v, w)^T$:	3D velocity vector
ν :	Kinematic viscosity
P :	Pressure
Re :	Reynolds number
t :	Time
Ω :	Spatial domain in \mathbb{R}^3
\mathbf{k} :	Wavevector
η^i :	Normalization constant, $i = a, b, c$
τ :	Time step parameter
β_i :	Bifurcation parameter, $i = 1, 2, 3$

γ_{ij} : Bifurcation parameter, $i, j = 1, 2, 3$ with $i \neq j$
 n : Iteration counter
 a_{CF} : u -direction forcing term
 b_{CF} : v -direction forcing term
 c_{CF} : w -direction forcing term.

References

- [1] M. Gad-El-Hak and H. Tsai, *Transition and Turbulence Control*, World Scientific Books, Hackensack, NJ, USA, 2005.
- [2] N. Kasagi, Y. Suzuki, and K. Fukagata, "Microelectromechanical systems-based feedback control of turbulence for skin friction reduction," *Annual Review of Fluid Mechanics*, vol. 41, pp. 231–251, 2009.
- [3] K. Fukagata and N. Kasagi, "Drag reduction in turbulent pipe flow with feedback control applied partially to wall," *International Journal of Heat and Fluid Flow*, vol. 24, no. 4, pp. 480–490, 2003.
- [4] S. F. Tardu, "Active control of near-wall turbulence by local oscillating blowing," *Journal of Fluid Mechanics*, vol. 439, pp. 217–253, 2001.
- [5] K. Angele, N. Kurimoto, Y. Suzuki, and N. Kasagi, "Evolution of the streamwise vortices in a coaxial jet controlled with micro flap actuators," *Journal of Turbulence*, vol. 7, no. 73, pp. 1–19, 2006.
- [6] P. Moin and T. Bewley, "Application of control theory to turbulence," in *20th Australian Fluid Mechanics Conference*, pp. 109–117, Springer, New York, NY, USA, 1995.
- [7] A. H. Techet, F. S. Hover, and M. S. Triantafyllou, "Separation and turbulence control in biomimetic flows," *Flow, Turbulence and Combustion*, vol. 71, no. 1–4, pp. 105–118, 2003.
- [8] P. Moin and T. Bewley, "Feedback control of turbulence," *Applied Mechanics Reviews*, vol. 47, no. 6, pp. S3–S12, 1994.
- [9] F. Abergel and R. Temam, "On some control problems in fluid mechanics," *Theoretical and Computational Fluid Dynamics*, vol. 1, no. 6, pp. 303–325, 1990.
- [10] R. M. May, "Simple mathematical models with very complicated dynamics," *Nature*, vol. 261, no. 5560, pp. 459–467, 1976.
- [11] J. M. McDonough and M. T. Huang, "A "poor man's Navier-Stokes equation": derivation and numerical experiments—the 2-D case," *International Journal for Numerical Methods in Fluids*, vol. 44, no. 5, pp. 545–578, 2004.
- [12] J. B. Polly, *Numerical experiments of the 3-D poor man's Navier-Stokes equations [M.S. thesis]*, University of Kentucky, Lexington, Ky, USA, 2012.
- [13] T. Yang, J. M. McDonough, and J. D. Jacob, "Two-dimensional "poor man's Navier-Stokes equation" model of turbulent flows," *AIAA Journal*, vol. 41, no. 9, pp. 1690–1696, 2003.
- [14] E. Lorenz, "Deterministic nonperiodic flow," *Journal of the Atmospheric Sciences*, vol. 20, no. 2, pp. 130–141, 1963.
- [15] D. Ruelle and F. Takens, "On the nature of turbulence," *Communications in Mathematical Physics*, vol. 20, pp. 167–192, 1971.
- [16] U. Frisch, *Turbulence: The Legacy of A. N. Kolmogorov*, Cambridge University Press, Cambridge, UK, 2001.
- [17] P. Constantin and C. Foias, *Navier-Stokes Equations*, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, Ill, USA, 1988.
- [18] C. Foias, O. Manley, R. Rosa, and R. Temam, *Navier-Stokes Equations and Turbulence*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, UK, 1988.
- [19] M. Hénon, "A two-dimensional mapping with a strange attractor," *Communications in Mathematical Physics*, vol. 50, no. 1, pp. 69–77, 1976.
- [20] K. T. Alligood, T. D. Sauer, and J. A. Yorke, *Chaos: An Introduction to Dynamical Systems*, Textbooks in Mathematical Sciences, Springer, New York, NY, USA, 1996.
- [21] E. Ott, C. Grebogi, and J. A. Yorke, "Controlling chaos," *Physical Review Letters*, vol. 64, no. 11, pp. 1196–1199, 1990.
- [22] D. Y. Chen, L. Shi, H. T. Chen, and X. Y. Ma, "Analysis and control of a hyperchaotic system with only one nonlinear term," *Nonlinear Dynamics*, vol. 67, no. 3, pp. 1745–1752, 2012.
- [23] M. P. Aghababa and H. P. Aghababa, "A general nonlinear adaptive control scheme for finite-time synchronization of chaotic systems with uncertain parameters and nonlinear inputs," *Nonlinear Dynamics*, vol. 69, no. 4, pp. 1903–1914, 2012.

Research Article

Data Fusion Based Hybrid Approach for the Estimation of Urban Arterial Travel Time

S. P. Anusha, R. A. Anand, and L. Vanajakshi

Department of Civil Engineering, Indian Institute of Technology Madras, Chennai 600036, India

Correspondence should be addressed to L. Vanajakshi, lelitha@iitm.ac.in

Received 2 May 2012; Revised 23 July 2012; Accepted 24 July 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 S. P. Anusha et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Travel time estimation in urban arterials is challenging compared to freeways and multilane highways. This becomes more complex under Indian conditions due to the additional issues related to heterogeneity, lack of lane discipline, and difficulties in data availability. The fact that most of the urban arterials in India do not employ automatic detectors demands the need for an effective, yet less data intensive way of estimating travel time. An attempt has been made in this direction to estimate total travel time in an urban road stretch using the location based flow data and sparse travel time data obtained using GPS equipped probe vehicles. Three approaches are presented and compared in this study: (1) a combination of input-output analysis for mid-blocks and Highway Capacity Manual (HCM) based delay calculation at signals named as base method, (2) data fusion approach which employs Kalman filtering technique (nonhybrid method), and (3) a hybrid data fusion HCM (hybrid DF-HCM) method. Data collected from a stretch of roadway in Chennai, India was used for the corroboration. Simulated data were also used for further validation. The results showed that when data quality is assured (simulated data) the base method performs better. However, in real field situations, hybrid DF-HCM method outperformed the other methods.

1. Introduction

Characterization of traffic systems is complex in nature due to the dynamic interaction between the system components, namely, the vehicles, road, and the road users. The uncertainties associated with human behavior makes the system more complex making modeling of the system a challenging task. Estimation and prediction of various parameters associated with this system is also difficult due to the associated uncertainties. The usual parameters used for characterizing the system include flow, speed, density, and travel time. The present study is dealing with the estimation of one of these parameters, namely, travel time. To obtain travel time information of all vehicles in a stream by direct measurement is both time consuming and costly, and it is impractical to collect this information from

all the road stretches in a network. Travel time in urban roads experience high short-term variability and hence cannot be measured using point detection. Being a spatial parameter, direct measurement of it needs either vehicle tracking devices or vehicle reidentification feature. However, majority of the vehicle tracking or reidentification techniques available such as automatic vehicle locators (AVL) and automatic vehicle identifiers (AVI) require participation, which limits the sample size. This underlines the need to estimate travel time from other easily measurable location based parameters such as flow and speed and has been an important research topic for many years. However, majority of researches on travel time estimation and prediction were reported for freeways, where traffic flow is not much affected by external factors such as traffic signals and conflicting movements. Travel time estimation and prediction is more complex and challenging on an urban network due to the influence of signals, presence of opposing movements, mid-link sources and sinks, and random fluctuations in travel demand. The situation is grave for Indian conditions because of the additional complexities related to heterogeneity, lack of lane discipline, and nonavailability of a reliable historic data base. Hence, methods which are cost effective and less demanding in terms of data base need to be explored.

Travel time, being a spatial parameter, is difficult to be measured directly from field. Most of the direct travel time measurement techniques such as test car methods or vehicle re-identification are expensive, immature, or involve privacy concerns and hence majority of the studies depend on indirect methods for travel time estimation. Most of the indirect travel time estimation and forecasting methods can be grouped under extrapolation techniques [1], regression models [2, 3], pattern recognition techniques [4], time series analysis [5], use of filtering techniques [6, 7], neural networks [8], methods based on traffic flow theory [9–14], data fusion techniques [15], and combination of above methods [16–19].

Many of the above methods require a good data base and may not be feasible for locations where automated data collection is not yet functional. Under such conditions, methods which demand less amount of data are required. Indian traffic characterised by its heterogeneity and lack of lane discipline poses additional challenges in terms of automated data collection. Most of the existing location based sensors are lane based and will fail under less or no lane disciplined traffic. Thus, accurate measurement of all the traffic parameters automatically is still a difficult task under Indian traffic conditions. Preliminary developments in this area are showing some promise in terms of traffic counts and hence the present study assumes that traffic count is the only location based data available. On the other hand, spatial data collection using GPS is a proven technology and is applicable under Indian conditions too. However, due to less participation, data from only a sample of the entire population, mainly from public transit, can be obtained using this technology. Thus, there is a need to have multiple sensors to characterise the entire traffic stream.

The present study develops a methodology for estimating stream travel time for an urban arterial using flow data obtained from location based sensors and GPS data obtained from limited number of probe vehicles. This approach known as data fusion (DF) is not explored under Indian traffic conditions. To improve the estimation accuracy, a hybrid DF-HCM method using data fusion for mid-block sections and HCM approach for the delay calculation at intersection is attempted. To compare the performance, a base method which employs input-output analysis for mid-blocks and HCM for intersection is also carried out. The usefulness of analysing separately the delay at signals is tested by comparing with the total travel time till intersection being estimated by using the nonhybrid method which employs data fusion alone for the whole stretch. A brief literature review on these approaches is given below.

Data fusion is a broad area of research in which data from several sensors are combined to provide comprehensive and accurate information [20]. The advantages of using data fusion include increased confidence, reduced ambiguity, improved detection, increased robustness, enhanced spatial and temporal coverage, and decreased costs [20–22]. The basic idea of data fusion is to estimate parameters by using more than one measurement from different sources or sensors. This may be due to lack of availability of enough data from a single source or to capture the advantages of different data sources. Some specific applications of data fusion in the field of transportation engineering are discussed below.

Kwon et al. [23] proposed a linear regression model for travel time prediction by combining both loop detector and probe vehicle data. They showed that linear regression on current flow, occupancy measurements, departure time, and day of week is beneficial for short-term travel time prediction while historical method is better for long-term travel time prediction. Zhang and Rice [24] used a linear model with varying coefficients to predict the travel time on freeways using loop detector and probe vehicle data. The coefficients vary as smooth functions of departure time. The coefficients have to be estimated offline and stored and after that the model can be used real-time. El-Faouzi et al. [25] put forward a model based on the Dempster-Shafer theory. They used travel time from loop detector and toll collection data to estimate travel time. The model required the likelihood that the data sources are giving the correct data. El-Faouzi [22] carried out a similar work using Bayesian method using travel time data from loop detector and probe vehicle to estimate travel time. The results showed that the travel time estimate using data fusion approach was better than the estimate obtained if the data sources were used individually. Chu et al. [21] used simulated loop detector and probe vehicle data to estimate travel time using a model based approach with Kalman filtering technique. Ivan [26] used the ANN technique to detect traffic incidents on signalized arterials using simulated travel time data from loop detector and probe vehicle data.

Another simple analytical model that uses readily available count data from upstream and downstream ends of a link for the estimation of travel time is the cumulative counts (input-output) method [9, 10, 12–14, 27]. However, a major drawback of the input-output method is its dependency on the accuracy of flow counts for travel time estimation [9–13, 27, 28]. Some of the other reported approaches include traffic flow theory based [11, 29, 30]. It can be seen from the above literature review that majority of the models discussed above are limited to freeways, and it may not be feasible to apply them directly on urban networks without further calibration due to differences in behaviour of traffic on the freeway and urban facilities. Moreover, the models developed for freeways generally provide average travel time for the link as a whole, which may not be a true representation in case of links with intersections, turning movements, and so forth. Thus, for better performance, intersection delays may have to be dealt separately. The present study analyse the validity of this assumption by comparing the accuracy of the estimated travel time with and without considering the intersection separately.

The first and one of the most popular methods for intersection delay estimation was developed by Webster [31] from a combination of theoretical and numerical simulation approaches that became the basis for all subsequent delay models. Modifications to the above model under varying traffic conditions were reported by Miller [32] and Newell [33]. The delay model suggested in Highway Capacity Manual usually known as HCM model [34] is a modified Webster's model incorporating the effect of progression and platooning. Attempts to overcome the assumption of steady-state condition by using time-dependent functions are reported in [35]. Other reported studies include deterministic queueing method [28, 36],

modified input-output technique [18], shock-wave theory based models [37, 38], and the use of Markov Chain processes [39, 40].

Overall, it can be seen that most of the reported studies on travel time and delay estimation used data collected from homogeneous and lane-disciplined traffic, either directly from the field or indirectly through simulation models. The traffic conditions existing in India is complex and different with its heterogeneity and lack of lane discipline. There are only limited studies [7, 41, 42] which addressed heterogeneous traffic characteristics. None of those studies estimated the stream travel time in an urban arterial taking signals into account. Also, lack of automated data collection methods in India makes it difficult to explore many of the statistical, time series, and machine learning techniques which are data driven since a good data base is required for applying such techniques. Thus, the application of a new approach for urban arterial travel time estimation with less data requirement is an area that will be of interest for countries like India and requires additional research and is discussed in this study.

The present study compares the performance of three different travel time estimation methodologies, which uses flow data as the main input. The estimation methodologies include two hybrid methods namely base method and data fusion-HCM and a non-hybrid method. The study stretch consists of a midblock and an intersection. The total travel time of the stretch is considered as the sum of travel time in the midblock section, without being influenced by the intersections, and the travel time at intersection, taking into account the delays at signals. Mid-block travel time is estimated using two approaches, namely, input-output analysis and data fusion approach. Input-output analysis is a popular approach and utilizes the cumulative count at entry and exit to find the travel time of vehicles within the section. HCM method is the most popular approach for estimating delay at signalized intersections. The method of applying input-output analysis for the midblock and HCM method for the intersection to obtain the total travel time of vehicles in the study stretch can be considered as a base approach and is entitled as base method. The other approach presents a data fusion method for mid-block travel time estimation and HCM analysis for the intersection area and will be called as hybrid DF-HCM method. The data fusion approach utilizes the location based flow data and the sparse travel time data obtained from probe vehicles for estimating the travel time of the stream. The total travel time of the stretch is then obtained by summing up the mid-block travel time and delay incurred at the bounding intersection. The necessity for analyzing the intersections separately was validated using a non-hybrid approach, where the total travel time is estimated using data fusion alone for the whole stretch without separating into mid-block and intersection area. In this approach, the total travel time is directly estimated using the data fusion approach without separately analysing the delay at intersection. The data till intersection stop line is used in this case for the data fusion approach assuming that the delay is implicitly captured by the flow and travel time data till stop line. A comparison of these three methodologies is carried out to understand the best method for travel time estimation under heterogeneous traffic conditions. This is one of the first study under Indian conditions that have applied data fusion techniques as well as hybrid technique for the estimation of travel time. The study has illustrated an efficient method to estimate the stream travel time in urban arterials with limited GPS data and location based flow data. The results of the study stressed the necessity of analyzing the intersections separately for more reliable estimates of travel time in urban roads.

2. Data Collection

Under Indian traffic conditions, a ready-to-use data archive is not available and hence the study relied upon field data collected manually and simulated data using VISSIM simulation package for corroborating the estimation methods.

2.1. Field Data

The test bed selected for the present study is a six-lane busy arterial road, namely, Rajiv Gandhi Salai in Chennai, India. Traffic in one direction only was considered for the present study. Data requirements based on the selected methodology included flow data from three locations as shown in Figure 1. The distance between location 1 and location 2, which is before the influence of intersection, is 1.72 km and between location 2 and intersection (location 3) is 0.1 km, making the total section under consideration to be of 1.82 km.

Two pedestrian over bridges were identified to mount the cameras for covering location 1 and 2. The intersection area covering the stop line at location 3 was recorded using a camera mounted on a convenient luminaire support. To ensure that the three-video recordings could be synchronized during replay in the lab, the time clocks in all the three cameras were set to a common time at the start of the data collection. Data were collected over three days for a total of six hours for the complete analysis and another two days for a total of three and a half hours from the mid-block alone for comparing input-output method with data fusion approach.

The required location based data, namely, the flow was collected using videographic technique. Initial snap shots of the traffic inside the study stretch were taken from elevated points to get the initial count of vehicles in the section required for input-output analysis. Photographs taken from entry and exit points along with additional photographs taken from in-between elevated points were required to capture the whole length of the section. Classified flow data at the three-data collection points were extracted manually for two-wheeler, three-wheeler, and four-wheeler categories from the videos in a temporal resolution of one minute. The required flow data were extracted manually due to lack of automated procedures. Travel time data required for validation was also extracted manually from videos by reidentifying vehicles at various locations.

The limited travel time data required for the data fusion model were collected using test vehicles equipped with GPS units. The test vehicles comprised of two cars, two three-wheelers, and two two-wheelers which provided travel time data. These vehicles were travelling back and forth between entry and exit points continuously during the data collection period. Moreover, GPS data were available from route number 5C of the public transport bus passing through the stretch. The GPS raw data included time, latitude, and longitude at every five seconds interval. From this, travel time data were extracted using the software package ArcGIS [43].

Due to the lack of automated video data extraction, all the above field data collection and extraction were carried out manually, which was laborious and time consuming. The data collection procedure required lot of coordination for collecting the video simultaneously from three different locations, along with initial snapshots, and GPS data from all different types of vehicles. An error in video recording even in one location makes the entire data not useful for analysis. Due to these difficulties, it was decided to simulate the field traffic conditions using collected field data, and carry out further analysis using the data generated using the calibrated simulated network. The details of the simulation are detailed below.

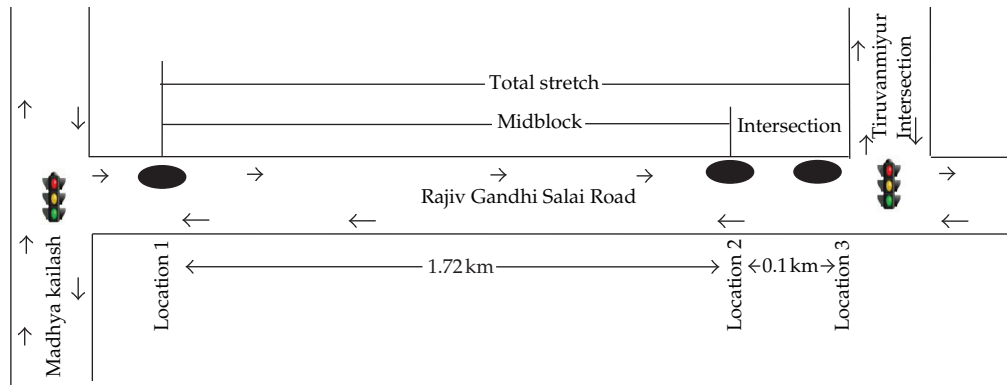


Figure 1: Schematic representation of the study corridor.

2.2. Simulated Data

VISSIM 5.3 from the PTV vision [44] was used in the present study to simulate the traffic conditions for testing the accuracy of the travel time estimation model under varying traffic conditions. A road stretch similar to the field test bed was created in VISSIM using a satellite image. For realistic representation of field conditions, data on intersection geometry, signal timing and phasing, vehicle types, traffic composition, vehicle input, proportion of turning traffic, and speed distribution were entered from field. Less lane disciplined, traffic movement was achieved by placing the vehicles anywhere on the lane, by setting the option for over-taking through left and right side of vehicle and allowing a diamond-shaped queuing at the intersection. To account for the nonstandard vehicles types, static, and dynamic characteristics of most of the regular vehicle types in terms of length, width, acceleration, and deceleration, and speed ranges were defined based on field values.

Signal timing and phase change data from the field were used in the simulation with a total cycle time of 145 s, red time of 98 s, green time 45 s, and amber time 2 s. Classified flow data for two-wheeler, three-wheeler, and four-wheeler categories extracted manually for the three locations is used for calibration and validation of the simulation. Five-minute aggregated flow and composition at location 1 were used as dynamic inputs for calibration. Data generated at the other two locations were compared with the field values for validation.

During calibration, several parameters in VISSIM were adjusted to match the field scenario. Simulations were performed with different random seeds with an average of five to ten values for each influencing parameter. Parameters were calibrated such that the error in flow, density, and travel time/speed was reduced. The errors were quantified in terms of mean absolute percentage error and were comparable with other similar studies in VISSIM [45, 46].

3. Methodology

3.1. Estimation Schemes

As mentioned already, in this study, the total travel time is estimated using a hybrid DF-HCM method making use of data fusion approach for the mid-block and HCM approach for intersection. A comparison is carried out using the base method employing input-output analysis which uses a simple deductive principle of cumulative counts (input-output method) for estimating the link travel time and HCM analysis for the intersection

area. The total travel time of the stretch is then estimated as the sum of mid-block and intersection travel time. Also, the need for analysing intersection delay separately is verified by comparing with a direct estimation of travel time till intersection using non-hybrid method which employs data fusion approach alone for the whole section. The basic approach of the above methods, namely, model based approach using data fusion, input-output method, and HCM approach are discussed below.

3.2. Data Fusion Method

This section details the methodology adopted for fusing both Eulerian video data and Lagrangian GPS data for estimating the travel time. The methodology was motivated from the study of Chu et al. [21]. The estimation scheme is based on the conservation equation and the fundamental traffic flow equation given in (3.1), and (3.2), respectively:

$$\frac{\partial q}{\partial x} + \frac{\partial k}{\partial t} = 0, \quad (3.1)$$

$$q = kV, \quad (3.2)$$

where q is the flow in PCU/hour, k is the density in PCU/km, and V is the space mean speed in km/hour, with x being the distance and t being the time.

By discretising (3.1), the density at time t can be represented as

$$k(t) = k(t-1) + \frac{\Delta t \times \{q_{\text{entry}}(t-1, t) - q_{\text{exit}}(t-1, t)\}}{\Delta x}, \quad (3.3)$$

where $q_{\text{entry}}(t-1, t)$ and $q_{\text{exit}}(t-1, t)$ are, respectively, the flow in PCU/h at the entry and exit points during the time interval $(t-1)$ to t . Δt is the data aggregation interval (1 minute in this study).

A filtering technique is used to estimate the density by assuming a value for the initial density, $k(0)$. Then the average travel time taken by the vehicles to reach the exit point from the entry is given by

$$\text{tt}(t) = \frac{\Delta x}{V(t)} = \frac{\Delta x \times k(t)}{q(t-1, t)}, \quad (3.4)$$

where $\text{tt}(t)$ is the travel time at time t , $q(t-1, t)$ is the flow along the section during $(t-1)$ to t which is given by

$$q = \begin{cases} q_{\text{entry}}, & \text{if } (q_{\text{exit}} - q_{\text{entry}}) > q_{\text{critical}} \\ q_{\text{exit}}, & \text{if } (q_{\text{entry}} - q_{\text{exit}}) > q_{\text{critical}} \\ \frac{q_{\text{entry}} + q_{\text{exit}}}{2}, & \text{if } |q_{\text{entry}} - q_{\text{exit}}| < q_{\text{critical}}. \end{cases} \quad (3.5)$$

Average of the flows at entry and exit points is used under normal conditions, when the flows at both ends are comparable without any shock-wave propagation. When

the flows at the entry and exit are not comparable, minimum of the two was adopted to capture the density variation within the stretch. In the present study, q_{critical} was selected as 20 PCU/minute.

In the above formulation, since the initial density in the section, $k(0)$ is unknown, there is a need for a parameter estimation scheme. The use of techniques such as Kalman Filtering or High Gain Observer (HGO) based parameter identification are reported in literature for similar applications [21, 47, 48]. In the present study, a Kalman filter based estimation scheme is adopted.

The Kalman filter (KF) is a recursive algorithm [49] and is usually applicable to system models which can be written in the state space representation. It is a model based tool for estimation and prediction and incorporates the stochastic nature of parameters. The KF can be of different types such as discrete Kalman filter, extended Kalman filter, and adaptive Kalman filter. The selection of the filter depends on the nature of the governing equations. In the present problem, as state equation (3.3) and the measurement equation (3.4) are linear, the Discrete Kalman Filter (DKF) is used. The present study uses flow data from the video and travel time from limited test vehicles to estimate average stream travel time. The state variable used is traffic density and the travel time is the measurement variable. The state (process) and measurement (observation) equations of DKF can be derived from (3.3) and (3.4) and are given below.

State equation:

$$k(t) = k(t-1) + u(t-1, t) + w(t-1), \quad (3.6)$$

Measurement equation:

$$tt(t) = H(t) \times k(t) + z(t), \quad (3.7)$$

where $u(t-1, t)$ is the input which depends on the flow and is given by

$$u(t-1, t) = \frac{q_{\text{entry}}(t-1, t) - q_{\text{exit}}(t-1, t)}{\Delta x}, \quad (3.8)$$

$H(t)$ is the transition matrix which converts the density to travel time and is given by

$$H(t) = \frac{\Delta x}{q(t-1, t)}, \quad (3.9)$$

and $w(t-1)$ and $z(t)$ are the process disturbance and the measurement noise, respectively. These are assumed to be Gaussian with zero mean and variances Q and R , respectively.

The Kalman filter algorithm is given by

$$\begin{aligned}
 \bar{k}(t) &= \hat{k}(t-1)^+ + u(t-1, t), \\
 \bar{P}(t) &= P(t-1)^+ + Q, \\
 G(t) &= \bar{P}(t)H(t)^T \left[H(t)\bar{P}(t)H(t)^T + R \right]^{-1}, \\
 \hat{k}(t)^+ &= \bar{k}(t) + G(t) \left[tt(t) - H(t)\bar{k}(t) \right], \\
 P(t)^+ &= \bar{P}(t) - G(t)H(t)\bar{P}(t),
 \end{aligned} \tag{3.10}$$

where $\bar{k}(t)$ is the a priori estimate of density calculated using the measurements prior to the instant t and $\bar{P}(t)$ is the a priori error covariance associated with $\bar{k}(t)$. $\hat{k}(t)^+$ and $P(t)^+$ are the a posteriori density estimate and its covariance, respectively, after incorporating the measurements till time t . $G(t)$ is the Kalman gain which is used in the correction process.

The above steps are repeated at every time steps, and the correction step was carried out only at the intervals when a measurement of GPS travel time is available.

3.3. Input-Output Method

The input-output (cumulative count) method as given by Nam and Drew [50] involves constructing the cumulative vehicle counts (N) on the y -axis and time on the x -axis, as shown in Figure 2.

The classical analytical procedure for travel time estimation considers cumulative flow plots $N(X_1, t)$ and $N(X_2, t)$ at upstream entrance and downstream exit of the link. The total travel time of vehicles during a given time interval, say between t_n and t_{n-1} , is then given by the area between the two curves for that time period, represented by the shaded region in Figure 2. The area can be calculated considering all vehicles that are entering, exiting, or entering and exiting. In this study, all vehicles that are exiting in the time period are considered, and the area is calculated accordingly. Corresponding analytical expression for total travel time (area of trapezoid) is given by

$$T(t_n) = \frac{1}{2} [(t_{n-1} - t'') + (t_n - t')] \times m(t_n), \tag{3.11}$$

where, t' = time of entry of the last vehicle that exits the link during t_{n-1} to t_n , t'' = time of entry of the first vehicle that exits the link during the t_{n-1} to t_n , $m(t_n)$ = the total number of vehicles that exit the link during t_{n-1} to t_n .

Under the first-in first-out condition $m(t_n)$ can be given as

$$m(t_n) = Q(X_2, t_n) - Q(X_2, t_{n-1}), \tag{3.12}$$

where, $Q(X_2, t_n)$ = cumulative number of vehicles exiting at t_n , $Q(X_2, t_{n-1})$ = cumulative number of vehicles exiting at t_{n-1} .

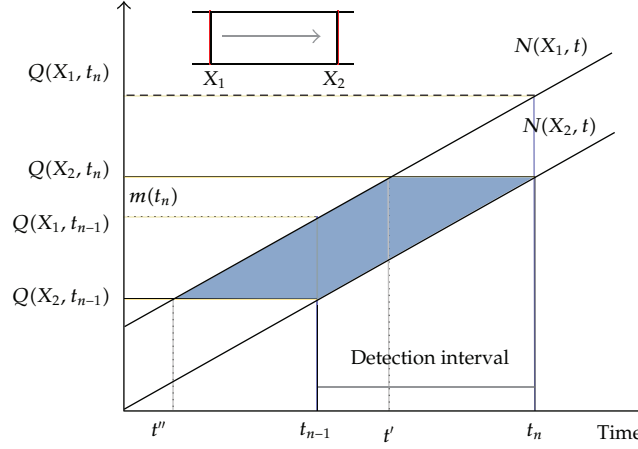


Figure 2: Illustration of input output analysis.

Interpolating the values of t' and t'' and substituting them in (3.11), the total travel time $T(t_n)$ can be calculated. The average travel time of vehicles exiting the link during the given interval (\bar{T}) is then obtained by dividing the total travel time $T(t_n)$ by the number of vehicles exiting the link for the same period as $T(t_n)/m(t_n)$ where $m(t_n)$ is the number of vehicles that exit during the interval.

3.4. HCM Delay Method

HCM delay method [34] is for estimating delay at an intersection over a given time period. Using this method, the average delay per vehicle for a lane group can be calculated using (3.13).

$$\begin{aligned}
 d &= (d_1 \times f_{PF}) + d_2 + d_3, \\
 d_1 &= \frac{0.5 \times C [1 - g/C]^2}{1 - [\min(1, X)(g/C)]}, \\
 d_2 &= 900 \times T \left[(X - 1) + \sqrt{(X - 1)^2 + \frac{8kIX}{cT}} \right], \\
 f_{PF} &= \frac{(1 - P)f_p}{1 - g/C},
 \end{aligned} \tag{3.13}$$

where, d = average overall delay per vehicle (seconds/vehicles); d_1 = uniform delay (s/veh); d_2 = incremental or random delay (s/veh); d_3 = residual demand delay or initial queue delay (s/veh); PF = progression adjustment factor; X = volume to capacity ratio of the lane group; C = traffic signal cycle time (seconds); c = capacity of the lane group (veh/h); g = effective green time for the through lane group (seconds); T = duration of analysis period (hours); k = incremental delay factor (0.50 for pre timed signals); I = upstream filtering/metering adjustment factor (1.0 for an isolated intersection); P = proportion of vehicles arriving during the green interval; f_{PF} = progression adjustment factor.

The above delay calculation of HCM requires flow values at location 3, free flow running time between the location 2 and location 3, cycle timings, capacity of the lane group, vehicle arrival type, and progression adjustment factor as input values. Out of these, the flow values and cycle timings were directly obtained from the field. The capacity of the lane group was calculated from the field using saturation flow rate and green cycle time ratio as given by $c = s \times g/C$ [34]. A value of 3564 vehicles per hour was obtained as capacity value which closely matched the value given in IRC: 106-1990 [51] for three-lane arterial road. Hence, the standard value of 3600 vehicles per hour as per IRC: 106-1990 for three-lane arterial road was taken for analysis. HCM Exhibits 16-11 and 16-12 [34] were used to determine the arrival type and progression adjustment factor for the known volume condition and vehicle distribution over green time. The value corresponding to arrival type 4 and green cycle time ratio of 0.3 was chosen for the progression adjustment factor.

The additional free flow running time for the delay stretch (constant value for a particular link) is added to the estimated delay to obtain the total travel time between the two locations. The free flow running time is obtained by dividing the distance between locations 2 and 3, $L_{a,s}$ by the free flow speed ff_s of the stretch. Thus, the total travel time of the delay stretch is obtained as

$$\overline{TT} = d + \frac{L_{a,s}}{ff_s}, \quad (3.14)$$

where, \overline{TT} is the total travel time of the delay stretch and d is the estimated delay value.

The total link travel time of 1.82 km stretch is then computed as the sum of mid-block travel time (1.72 km) and the travel time in the delay stretch (0.1 km). Each of the modules of base, hybrid DF-HCM, and non-hybrid method is corroborated using field data and simulated data and is detailed in the section below.

4. Corroboration of the Estimation Scheme

In order to evaluate the performance of the above methods, the estimated mean link travel times using these methods were plotted against the actual travel times using both field data and simulated data. The actual travel time data required for the validation of the results while using field data were obtained through GPS equipped test vehicles as well as by manually re-identifying vehicles from videos. GPS data were collected using three cars, three auto rickshaws, two motor bikes, and five buses as representative samples of each classification. Throughout the data collection period, these GPS test vehicles, except the buses, were made to travel along the study section repeatedly. Figures 3, 4, and 5 shows the plots of travel times predicted by the three methods compared to the actual travel time from the field data and simulated data. It can be seen that the hybrid DF-HCM model is able to capture the variations in the actual travel time better than the other two methods.

The errors in travel time estimation in the above cases were quantified using mean absolute percentage error (MAPE) and root mean squared error (RMSE). The mean absolute percentage error is obtained using

$$MAPE = \left(\frac{1}{N} \sum_{k=1}^N \frac{|N_{meas}(k) - N_{est}(k)|}{N_{meas}(k)} \right) \times 100\%, \quad (4.1)$$

where $N_{est}(k)$ and $N_{meas}(k)$ are the estimated and the measured average travel time of the study stretch during the k th interval of time with N being the total number of time intervals.

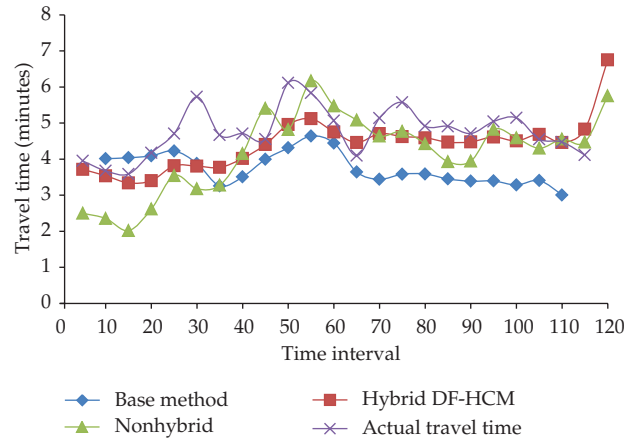


Figure 3: Sample travel time results using field data.

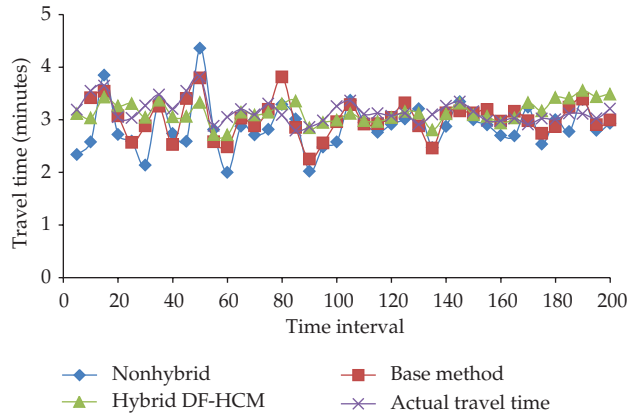


Figure 4: Travel time comparison for a 2 hrs simulated data.

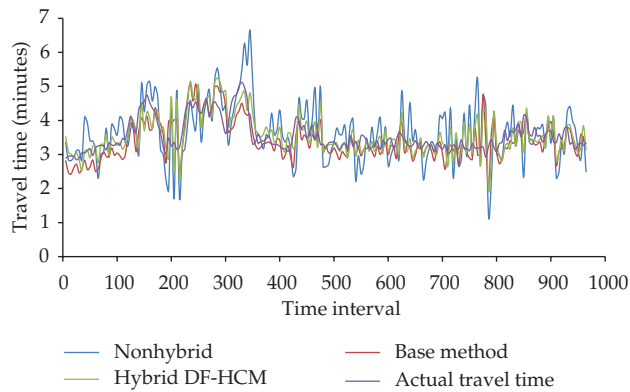


Figure 5: Sample plot for travel time comparison from 6:00 A.M–10:00 P.M simulated data.

MAPE meets most of the criteria required for a summary measure such as measurement validity, reliability, ease of interpretation, clarity of presentation, and support of statistical evaluation. However, as noted by most researchers [52], the distribution of MAPE is often asymmetrical or right skewed and undefined for zero values. Hence, a scale dependent measure called root mean square error (RMSE) is also used which is often helpful when different methods applied to the same set of data are compared. However, there is no absolute criterion for a “good” value of any of the scale dependent measures as they are on the same scale as the data [53]. The lesser the value of RMSE, the better is the forecast obtained. RMSE expresses the expected value of the error and has the same unit as the data which makes the size of a typical error visible. the root mean square error is given by

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{k=1}^N [N_{\text{meas}}(k) - N_{\text{est}}(k)]^2}, \quad (4.2)$$

where $N_{\text{est}}(k)$ and $N_{\text{meas}}(k)$ are the estimated and the measured average travel time of the study stretch during the k th interval of time with N being the total number of time intervals.

The MAPE and RMSE values obtained are shown in Figures 6 and 7 and can be seen that the errors are within acceptable ranges [52, 53]. According to Lewis’ scale of judgment of forecasting accuracy [49], any forecast with a MAPE value of less than 10% can be considered highly accurate, 11%–20% is good, 21%–50% is reasonable, and 51% or more is inaccurate.

It can be observed that in the case of field data, hybrid DF-HCM method performed better than the other two methods. The results using simulated data show both base and hybrid DF-HCM methods performing comparably and both performing better than the non-hybrid method. Thus, the results clearly show that analysing the delays at intersections separately brings in more accuracy to travel time estimation. Also, it can be observed that input-output method is too constrained by the flow data quality and can be used only when the flow data accuracy is guaranteed such as from simulation. This is further checked by comparing the performance of two more days of field data for mid-block section, and the MAPE and RMSE is as shown in Table 1. It can be seen that in these cases, the data fusion method outperformed the input-output method unlike the case of using data from simulation confirming that with uncertainty in flow values as obtained from field, the data fusion method outperforms the input-output approach.

Overall, it can be observed that the proposed data fusion method is a better candidate for travel time estimation compared to input-output method in the mid-block sections. Input-output method can be considered in cases where the accuracy of flow values is guaranteed. Also, it can be clearly observed that, analysing the intersection delay separately brings in more accuracy to the travel time estimation. This leads to the conclusion that intersection delay needs to be analysed separately to determine the total travel time of urban arterials. Thus, the proposed data fusion method can be used in mid-block sections along with HCM method for delay estimation at intersections to determine the total travel time of urban arterials.

5. Summary and Conclusions

The negative impacts of growth in vehicular population include congestion and delays, and are much debated topics currently all over the world. India, with its rapid growth in

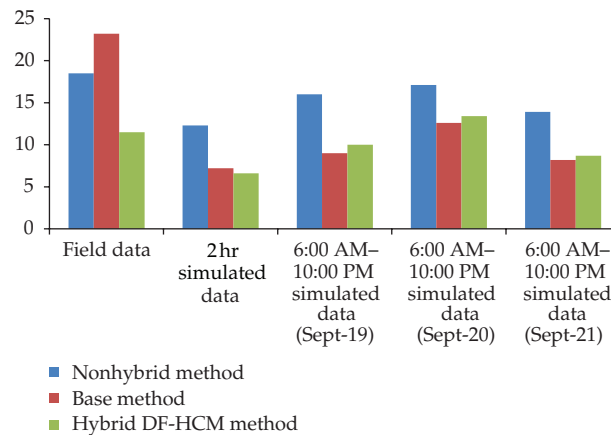


Figure 6: Performance of models in terms of MAPE for field and simulated data.

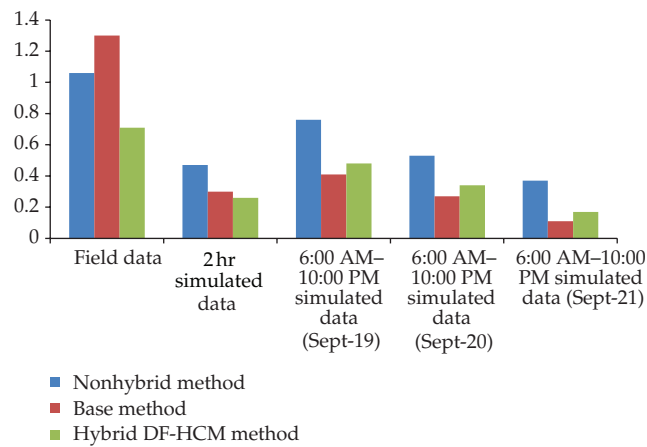


Figure 7: Performance of models in terms of RMSE for field and simulated data.

economy and corresponding growth in vehicular population, is no exception to this. The heterogeneous nature of traffic and lack of lane discipline makes these issues more complex. Characterization and analysis of this type of traffic existing in developing countries demand a different approach than what is followed in western countries with homogenous and lane disciplined traffic. While dealing with urban arterials, there are additional challenges due to the high variability in traffic characteristics. In countries such as India, another challenge is in terms of data availability and hence less data intensive methods for travel time estimation are to be employed.

The present study analysed the application of data fusion technique for travel time estimation in an Indian urban arterial. The study attempted a hybrid DF-HCM method and compared its performance with a base method and a non-hybrid data fusion model to obtain the total travel time of the whole section. The performance of the models for varying traffic flow conditions was tested using field and simulated data. The results showed hybrid DF-HCM model as the best candidate for travel time estimation in urban arterials. The non-hybrid method was found to have the highest error stressing the need for analysis

Table 1: Performance of models in terms of MAPE and RMSE for mid-block.

Method/Date	MAPE (%)		RMSE (min)	
	Input-output	Data fusion	Input-output	Data fusion
June 30	21.9	18.9	0.132	0.096
Sept 23	13.6	9.0	0.28	0.12

of the intersections separately for better performance. Among the hybrid models, data fusion method gave promising results under field conditions. When the accuracy of flow value was guaranteed, such as using simulated data, both the base and hybrid DF-HCM methods showed comparable performance with a slightly better performance from the input-output method. Hence, for real-time field implementations such as ATIS for urban arterials, the hybrid approach using data fusion for mid-block sections along with separate delay estimation at signalized intersections brings in the maximum accuracy of the predicted travel time information showing its potential for any such real time ITS implementations.

Acknowledgments

The authors acknowledge the support for this study by Ministry of Communication and Information Technology, Government of India grant through letter no. 23(1)/2009-IEAD and the Indo-US Science and Technology Forum for the support through Grant no. IUSSTF/JC-Intelligent Transportation Systems//95-2010/2011-12.

References

- [1] S. M. Turner, W. L. Eisele, R. J. Benz, and D. J. Holdener, "Travel time data collection handbook," Tech. Rep. FHWA-PL-98-035, Federal Highway Administration, Washington, DC, USA, 1998.
- [2] V. P. Sisiopiku, N. M. Rouphail, and A. Santiago, "Analysis of correlation between arterial travel time and detector data from simulation and field studies," *Transportation Research Record*, no. 1457, pp. 166–173, 1994.
- [3] H. M. Zhang, "Link-journey-speed model for arterial traffic," *Transportation Research Record*, no. 1676, pp. 109–115, 1999.
- [4] S. Robinson and J. W. Polak, "Modeling urban link travel time with inductive loop detector data by using the k-NN method," *Transportation Research Record*, no. 1935, pp. 47–56, 2005.
- [5] J. Anderson and M. Bell, "Travel time estimation in urban road networks," in *Proceedings of the 1997 IEEE Conference on Intelligent Transportation Systems, ITSC*, pp. 924–929, November 1997.
- [6] W. H. Lin, A. Kulkarni, and P. Mirchandani, "Short-term arterial travel time prediction for advanced traveler information systems," *Journal of Intelligent Transportation Systems*, vol. 8, no. 3, pp. 143–154, 2004.
- [7] L. D. Vanajakshi, S. C. Subramanian, and R. Sivanandan, "Short-term prediction of travel time for Indian traffic conditions using buses as probe vehicles," in *Proceedings of the 87th Annual Meeting on Transportation Research Board*, Washington, DC, USA, 2008.
- [8] D. Park, L. R. Rilett, and G. Han, "Spectral basis neural networks for real-time travel time forecasting," *Journal of Transportation Engineering*, vol. 125, no. 6, pp. 515–523, 1999.
- [9] D. H. Nam and D. R. Drew, "Traffic dynamics: method for estimating freeway travel times in real time from flow measurements," *Journal of Transportation Engineering*, vol. 122, no. 3, pp. 185–191, 1996.
- [10] D. H. Nam and D. R. Drew, "Analyzing freeway traffic under congestion: traffic dynamics approach," *Journal of Transportation Engineering*, vol. 124, no. 3, pp. 208–212, 1998.
- [11] A. Skabardonis and N. Geroliminis, "Real-time estimation of travel times on signalized arterials," in *Proceedings of the 16th International Symposium on Transportation and Traffic Theory*, pp. 387–406, 2005.
- [12] K. Krishnan, *Travel time estimation and forecasting on urban roads [Ph.D. thesis]*, Imperial College, London, UK, 2008.

- [13] L. D. Vanajakshi, B. M. Williams, and L. R. Rilett, "Improved flow-based travel time estimation method from point detector data for freeways," *Journal of Transportation Engineering*, vol. 135, no. 1, pp. 26–36, 2009.
- [14] A. Bhaskar, E. Chung, and A. G. Dumont, "Analysis for the use of cumulative plots for travel time estimation on signalized network," *International Journal of Intelligent Transportation Systems Research*, vol. 8, no. 3, pp. 151–163, 2010.
- [15] K. Choi and Y. Chung, "A data fusion algorithm for estimating link travel time," *ITS Journal*, vol. 7, no. 3-4, pp. 235–260, 2002.
- [16] J. You and T. J. Kim, "Development and evaluation of a hybrid travel time forecasting model," *Transportation Research C*, vol. 8, no. 1–6, pp. 231–256, 2000.
- [17] C. Xie, R. L. Cheu, and D. H. Lee, "Improving arterial link travel time estimation by data fusion," in *Proceedings of the 83rd Annual Meeting on Transportation Research Board*, Washington, DC, USA, 2004.
- [18] A. Sharma, D. M. Bullock, and J. A. Bonneson, "Input-output and hybrid techniques for real-time prediction of delay and maximum queue length at signalized intersections," *Transportation Research Record*, no. 2035, pp. 69–80, 2007.
- [19] J. Bonneson, A. Sharma, and D. Bullock, "Measuring the performance of automobile traffic on urban streets," NCHRP Report, Transportation Research Board of the National Academies, Washington, DC, USA, 2008.
- [20] D. J. Dailey, P. Harn, and P. J. Lin, "ITS data fusion," ITS Research program Final Research Report, Washington State Transportation Commission, Washington, DC, USA, 1996.
- [21] L. Chu, J. Oh, and W. Recker, "Adaptive kalman filter based freeway travel time estimation," in *Proceedings of the 84th Annual Meeting on Transportation Research Board*, Washington, DC, USA, 2005.
- [22] N. E. El-Faouzi, "Bayesian and evidential approaches for traffic data fusion," in *Proceedings of the 85th Annual Meeting on Transportation Research Board*, Washington, DC, USA, 2006.
- [23] J. Kwon, B. Coifman, and P. Bickel, "Day-to-day travel-time trends and travel-time prediction from loop-detector data," *Transportation Research Record*, no. 1717, pp. 120–129, 2000.
- [24] X. Zhang and J. A. Rice, "Short-term travel time prediction," *Transportation Research C*, vol. 11, no. 3-4, pp. 187–210, 2003.
- [25] N. E. El-Faouzi, L. A. Klein, and O. De Mouzon, "Improving travel time estimates from inductive loop and toll collection data with dempster-shafer data fusion," *Transportation Research Record*, no. 2129, pp. 73–80, 2009.
- [26] J. N. Ivan, "Neural network representations for arterial street incident detection data fusion," *Transportation Research C*, vol. 5, no. 3-4, pp. 245–254, 1997.
- [27] C. F. Daganzo, *Fundamentals of Transportation and Traffic Operations*, Pergamon-Elsevier, Oxford, UK, 1997.
- [28] A. D. May, *Traffic Flow Fundamentals*, Prentice Hall, Englewood Cliffs, NJ, USA, 1990.
- [29] R. Pueboobpaphan and T. Nakatsuji, "Real-time traffic state estimation on urban road network: the application of unscented Kalman filter," in *Proceedings of the Ninth International Conference on Applications of Advanced Technology in Transportation*, pp. 542–547, usa, August 2006.
- [30] H. X. Liu, W. Ma, X. Wu, and H. Hu, "Real time estimation of arterial travel time under congested conditions," *Transportmetrica*, vol. 8, no. 2, pp. 87–104, 2012.
- [31] F. V. Webster, "Traffic signal settings," Road Research Technical Paper 39, Department of Scientific and Industrial Research, London, UK, 1958.
- [32] A. J. Miller, "Settings for fixed-cycle traffic signals," *Operational Research Quarterly*, no. 14, pp. 373–386, 1963.
- [33] G. F. Newell, "Stochastic delays on signalized arterial highways," in *Transportation and Traffic Theory*, M. Koshi, Ed., pp. 598–598, Elsevier Science Publishing, 1990.
- [34] HCM, *Highway Capacity Manual*, Transportation Research Board, Washington, DC, USA, 2000.
- [35] R. M. Kimber and E. M. Hollis, "Traffic queues and delays at road junctions," TRRL Laboratory Report 909, Transport and Road Research Laboratory (TRRL), London, UK, 1979.
- [36] D. W. Strong, R. M. Nagui, and K. Courage, "New calculation method for existing and extended HCM delay estimation procedure," in *Proceedings of the 87th Annual Meeting Transportation Research Board*, Washington, DC, USA, 2006.
- [37] S. C. Wirasinghe, "Determination of traffic delays from shock-wave analysis," *Transportation Research*, vol. 12, no. 5, pp. 343–348, 1978.
- [38] W. B. Cronje, "Analysis of existing formulas for delay, overflow and stops," *Transportation Research Record*, no. 905, pp. 93–95, 1983.

- [39] W. Brilon and N. Wu, "Delays at fixed-time traffic signals under time-dependent traffic conditions," *Traffic Engineering and Control*, vol. 31, no. 12, 1990.
- [40] F. Viti and H. V. Zuylen, "Markov mesoscopic simulation model of overflow queues at multilane signalized intersections," in *Proceedings of the Joint Conference 10th EWGT Meeting & 16th Mini-EURO Conference*, Warsaw, Poland, 2005.
- [41] Y. Ramakrishna, P. Ramakrishna, V. Lakshmanan, and R. Sivanandan, "Bus travel time prediction using GPS data," <http://www.gisdevelopment.net/proceedings/mapindia/2006student%20oral/mi06stu.84.htm>, 2011.
- [42] R. P. S. Padmanaban, L. Vanajakshi, and S. C. Subramanian, "Estimation of bus travel time incorporating dwell time for APTS applications," in *2009 IEEE Intelligent Vehicles Symposium*, pp. 955–959, chn, June 2009.
- [43] ArcGIS 9.3 manual, ESRI, August 2008.
- [44] PTV, *VISSIM Version 5.3 Manual, Innovative Transportation Concepts*, Planung Transport Verkehr AG, Karlsruhe, Germany, 2010.
- [45] F. C. Fang and L. Elefteriadou, "Some guidelines for selecting microsimulation models for interchange traffic operational analysis," *Journal of Transportation Engineering*, vol. 131, no. 7, pp. 535–543, 2005.
- [46] T. V. Mathew and P. Radhakrishnan, "Calibration of microsimulation models for nonlane-based heterogeneous traffic at signalized intersections," *Journal of Urban Planning and Development*, vol. 136, no. 1, Article ID 005001QUP, pp. 59–66, 2010.
- [47] X. Dai, Z. Gao, T. Breikin, and H. Wang, "High-gain observer-based estimation of parameter variations with delay alignment," *IEEE Transactions on Automatic Control*, vol. 57, no. 3, pp. 726–732, 2012.
- [48] Z. Gao, X. Dai, T. Breikin, and H. Wang, "Novel parameter identification by using a high-gain observer with application to a gas turbine engine," *IEEE Transactions on Industrial Informatics*, vol. 4, no. 4, pp. 271–279, 2008.
- [49] G. Welch and G. Bishop, *An Introduction to the Kalman Filter*, University of North Carolina, Chapel Hill, NC, USA, 2006.
- [50] D. H. Nam and D. R. Drew, "Automatic measurement of traffic variables for intelligent transportation systems applications," *Transportation Research B*, vol. 33, no. 6, pp. 437–457, 1999.
- [51] IRC: 106, *Guidelines for Capacity of Urban Roads in Plain Areas*, Indian Roads Congress, New Delhi, India, 1990.
- [52] D. L. Kenneth and K. K. Ronald, *Advances in Business and Management Forecasting*, Emerald books, Howard House, London, UK, 1982.
- [53] C. D. Coleman and D. A. Swanson, "On MAPE-R as a measure of cross-sectional estimation and forecast accuracy," *Journal of Economic and Social Measurement*, vol. 32, no. 4, pp. 219–233, 2007.

Research Article

The Chaotic Prediction for Aero-Engine Performance Parameters Based on Nonlinear PLS Regression

Chunxiao Zhang¹ and Junjie Yue²

¹ College of Science, Civil Aviation University of China, Tianjin 300300, China

² Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China

Correspondence should be addressed to Chunxiao Zhang, cxzhang@cauc.edu.cn

Received 3 May 2012; Revised 18 July 2012; Accepted 19 July 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 C. Zhang and J. Yue. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The prediction of the aero-engine performance parameters is very important for aero-engine condition monitoring and fault diagnosis. In this paper, the chaotic phase space of engine exhaust temperature (EGT) time series which come from actual air-borne ACARS data is reconstructed through selecting some suitable nearby points. The partial least square (PLS) based on the cubic spline function or the kernel function transformation is adopted to obtain chaotic predictive function of EGT series. The experiment results indicate that the proposed PLS chaotic prediction algorithm based on biweight kernel function transformation has significant advantage in overcoming multicollinearity of the independent variables and solve the stability of regression model. Our predictive NMSE is 16.5 percent less than that of the traditional linear least squares (OLS) method and 10.38 percent less than that of the linear PLS approach. At the same time, the forecast error is less than that of nonlinear PLS algorithm through bootstrap test screening.

1. Introduction

Aeroengine condition monitoring is applied as a better way to estimate aero-engine condition and reliability amongst most of the commercial airlines [1, 2]. Predicting and detecting abnormal behavior of main aeroengine performance parameters is of great importance in the aeroengine condition monitoring. With the development of modern industry, one typical aeroengine becomes a more and more integrative and complex system, of which working condition and fault are very difficult to predict and judge, and the influence and harm caused by system faults are also more serious than ever. Therefore, there will be a growing demand for effective prediction methods of long-term trend of aeroengine system to predict

abnormal behavior of significant performance parameters and potential faults. Then we can take preventive maintenance measure to avoid faults occurrence, shorten aircraft down time, thereby reducing maintenance cost and improving aeroengine life.

The aeroengine system is a complex system with nonlinear and dynamic chaos action, so it is difficult to establish an exact analytic mathematical model of performance parameters. As a result, analysis usually depends on time series obtained from observation. But in practical monitoring the availability of obtaining characteristic data is very limited, and one fault usually shows various changes in performance parameters, and a same change may result from different faults [3, 4]. However, extracting multiple fault characteristic of parameters usually brings complexity and redundancy, which may bring difficulty for the following prediction. The theory of chaotic prediction is a mathematics method to deal with inexactness, uncertainty, and incompleteness, which is remarkable efficiency in aspects of data reduction and feature extraction.

Studies on chaos prediction can be traced back to Parker and Chua (1989) [5]. More sophisticated approaches of nonlinear chaotic time series prediction mainly rely on neural network [6], support vector theory [7], fuzzy rules [8], a simple least square algorithm [9]. In recent years, considerable interest has been aroused in the study of nonlinear chaos prediction with the actual observation of aeroengine dynamic system [10, 11]. These papers summarize the prediction model of nonlinear chaos system based on ordinary least squares (OLS); however, there are some practical problems in those models such as serious multicollinearity and poor robustness of regression function.

Partial least squares (PLS) can overcome information overlaps caused by multicollinearity since it adopts canonical correlation analysis, which improves accuracy of nonlinear system modeling [12]. What is more, PLS adopts cross validity method, which rejects redundancy principal component and enhances regression equations robustness. Hence, in our investigation the designed chaotic algorithm is based on the nonlinear partial least squares (NPLS) regression and the prediction research on the time series.

2. Data Pretreatment

Typical aeroengine performance parameters are exhaust gas temperature (EGT), low-spool rotor speed (NI), high-spool rotor speeds (N2), and fuel flow (WF) [13]. These four measurements are often called the four basic parameters, which can be recorded in the airborne equipment on most new and old aeroengines. Note that we take the most important performance parameter, EGT, as example to analyze in this paper; the other parameters can be fitted in the same method.

2.1. Similarly Correcting EGT Sequence

According to operating principle of the aeroengine, the data of main performance parameters of the same aeroengine in different external flight conditions such as the external atmosphere total temperature, total pressure, MACH numbers are quite different, so the original engine exhaust temperature EGT data usually cannot be directly used for comparison analysis. To

solve this problem, similar theory is adopted to eliminate the influence of external conditions on the aeroengine EGT sequence [14].

$$\text{EGT}_{\text{cor}} = \frac{\text{EGT}_{\text{obs}}}{\theta T_2}, \quad (2.1)$$

where EGT_{obs} is the original EGT data, EGT_{cor} is the corrected value of EGT_{obs} , $\theta T_2 = (\text{TAT} + 273.15) / 288.15$, TAT is the total air temperature.

2.2. Eliminating Outliers According to Layida Criterion

If outliers exist in the data, they will seriously affect the prediction accuracy of an algorithm. Layida criterion is the most commonly used for discrimination outliers, of which the precondition is that there is sufficient data [15]. For time series $\{X_i\} (i = 1, \dots, n)$, if n is large enough (normally $n > 80$), and only random disturbance exists in all data. Assume that the random time series is following a normal distribution whose mean is μ and deviation is σ^2 , so the probability that the deviation $(X_i - \mu)$ falls in $\pm 3\sigma$ ranges is about 0.3%. According to Layida criterion, if the absolute deviation of a time series data $X_i (i = 1, \dots, n)$ is larger than 3σ , then X_i can be judged as an outlier and should be removed.

Let the number of sample be N , the observed value of time series is $\{x_i\} (i = 1, \dots, N)$, the sample mean and standard deviation is \bar{x} and S , respectively. If the sample point x_i meets $|x_i - \bar{x}| > 3S$, then x_i is an outlier which will be eliminated. \bar{x} and S should be computed again after all the outliers are removed.

3. Chaos Prediction Model Based on Nonlinear Partial Least Squares (PLS) Regression

Suppose observed time series is $\{x_n\} (n = 1, 2, \dots, N)$, the chaotic phase space s_n is reconstructed by [16].

$$s_n = (x_n, x_{n-\tau}, \dots, x_{n-(m-1)\tau}) \in R^m, \quad n = (m-1)\tau + 1, (m-1)\tau + 2, \dots, N, \quad (3.1)$$

where τ is the delay time, m is the embedding dimension and N is the length of the time series.

3.1. Characterizing Chaos through Lyapunov Exponents

Through the years, the existence of chaos has been characterized in time series by means of several methods, among others: analysis of Fourier spectra, fractal power spectra, entropy, fractal dimension, and calculation of Lyapunov exponents. However, several of these methods have proven not to be very efficient [16]. The Lyapunov exponents provide all the information about the local and global complexity of chaos, therefore the measurement of Lyapunov exponents has been an effective method to judge whether a time series is caused by a chaotic dynamical system.

The Lyapunov exponents $\lambda_1, \lambda_2, \dots, \lambda_d$ are the average rates of expansion ($\lambda_i > 0$) or contraction ($\lambda_i < 0$) near a limit set of a dynamical system [16]. In fact, the LE is a quantitative

measure of the divergence of several trajectories in the system. There is one LE for each direction in the d -dimensional phase space where the system is embedded. The variable d represents the Lyapunov dimension or phase space dimension. The LE does not change when the initial conditions of a trajectory are modified or when some perturbation occurs in the system. If at least one LE is positive, the system presents chaotic motion [17]; hence if chaos exists in a dynamical system, the largest Lyapunov exponent must be positive.

We will calculate the largest Lyapunov exponents according to the algorithm offered by Wolf et al. (1985) [17]; the numerical results are shown in Table 1.

3.2. Chaos Prediction Model

Let state vector at time T be s_T , we obtain the nearest neighbor points $s_{\alpha_1}, s_{\alpha_2}, \dots, s_{\alpha_K}$ by calculating the Euclidean distance between the target points s_T and any one of reconstruction vectors [18], where

$$s_{\alpha_i} = (x_{\alpha_i}, x_{\alpha_i-\tau}, \dots, x_{\alpha_i-(m-1)\tau}), \quad i = 1, 2, \dots, K. \quad (3.2)$$

Let X be the sample matrix of reconstructed variable,

$$X = \begin{bmatrix} x_{\alpha_1} & x_{\alpha_1-\tau} & \cdots & x_{\alpha_1-(m-1)\tau} \\ x_{\alpha_2} & x_{\alpha_2-\tau} & \cdots & x_{\alpha_2-(m-1)\tau} \\ \vdots & \vdots & \vdots & \vdots \\ x_{\alpha_K} & x_{\alpha_K-\tau} & \cdots & x_{\alpha_K-(m-1)\tau} \end{bmatrix} \in R^{K \times m}. \quad (3.3)$$

Let y be the prediction variable, then the nonlinear chaos prediction model is described as

$$y = f(x_1, x_2, \dots, x_m), \quad (3.4)$$

where x_i is the i th column of X , $i = 1, 2, \dots, m$, $y = [x_{\alpha_1+p}, x_{\alpha_2+p}, \dots, x_{\alpha_K+p}] \in R^K$, p is the length of prediction step.

In practice, the observed data sets of independent variable and dependent variable can usually be obtained. However, the model form of their relations is unknown especially in case of higher dimensions, and the problem becomes more complicated when the relationship between the dependent variable and independent variables is nonlinear. Usually (3.4) is fitted by additive model or multiplicative model. For the sake of computation convenience, we consider the addition model of each independent variable; that is,

$$y = f_1(x_1) + f_2(x_2) + \cdots + f_m(x_m) + \varepsilon. \quad (3.5)$$

At present, the nonlinear relationship between the original variable usually is into a quasi-linear relationship, after then, the prediction model is established by an effective linear theory method. Spline function and kernel function are commonly used in the transformation of the basic functions.

Table 1: Comparative analysis of the optimal results based on various forecasting methods.

Forecasting method	Embedding dimension (m)	Delay time (τ)	Subset numbers	Largest Lyapunov exponent	NMSE
Cubic spline interpolation	9	3	5	0.25	0.6400
Uniform kernel	8	19	2	0.16	0.5529
Epanechnikov kernel	8	19	2	0.17	0.5529
Bi-weight kernel	8	19	2	0.38	0.5431
Tri-weight kernel	9	3	5	0.29	0.5918
Gaussian kernel	3	10	4	0.32	0.6325
PLS	9	3	0	0.15	0.6469
OLS	30	16	0	0.098	0.7081

(1) Cubic Spline Function Transformation

Let $\hat{f}_j(x_j) (j = 1, 2, \dots, m)$ be the cubic spline fitting function, then,

$$y = \beta_0 + \hat{f}_1(x_1) + \hat{f}_2(x_2) + \dots + \hat{f}_m(x_m) + \varepsilon, \quad (3.6)$$

where

$$\hat{f}_j(x_j) = \sum_{l=0}^{M_j+2} \beta_{j,l} \Omega_3 \left(\frac{x_j - \xi_{j,l-1}}{h_j} \right), \quad (3.7)$$

$$\Omega_3 \left(\frac{x_j - \xi_{j,l-1}}{h_j} \right) = z_{j,l} = \frac{1}{3!h_j^3} \sum_{k=0}^4 (-1)^k \binom{4}{k} (x_j - \xi_{j,l-3+k})^3, \quad (3.8)$$

$$\xi_{j,l-1} = \min(x_j) + (l-1)h_j,$$

$$h_j = \frac{\max(x_j) - \min(x_j)}{M_j} \quad (l = 0, 1, \dots, M_j + 2), \quad (3.9)$$

where $\beta_0, \beta_{j,l}$ are the model parameters to be determined; $\xi_{j,l-1}, h_j, M_j$ are range points, segment length and section number are divided on the variable x_j , respectively.

The minimum observed value of variables is recorded as $\min(x_j)$ and the maximum as $\max(x_j)$, then the nonlinear prediction function relationship between independent variables and dependent variable can be transformed as the following equation:

$$y = \beta_0 + \sum_{j=1}^m \hat{f}_j(x_j) + \varepsilon = \beta_0 + \sum_{j=1}^m \sum_{l=0}^{M_j+2} \beta_{j,l} \Omega_3 \left(\frac{x_j - \xi_{j,l-1}}{h_j} \right) + \varepsilon. \quad (3.10)$$

Equations (3.6)–(3.8) show that the relationship between y and $z_{j,l}$ is linear, and the relationship between x_j and $z_{j,l}$ is nonlinear, so the chaos prediction model with cubic spline function transformation is a quasi-linear regression model (see in (3.10)).

(2) Kernel Function Transformatio

The principle of Kernel function transformation is same as spline function transformation; that is, a kernel function transformation based on one dimensional nonlinear function $f_j(x_j)$, and the cubic spline function $\{\Omega_3((x_j - \xi_{j,l-1})/h_j)(l = 0, 1, \dots, M_j + 2)\}$ is replaced by basic function $\{K((x_j - \xi_{j,l-1})/h_j)(l = 0, 1, \dots, M_j + 2)\}$. Thus, the chaos prediction model with kernel function transformation is as follows [19]:

$$\hat{y} = \beta_0 + \sum_{j=1}^p \hat{f}_j(x_j) = \beta_0 + \sum_{j=1}^p \sum_{l=0}^{M_j+2} \beta_{j,l} K\left(\frac{x_j - \xi_{j,l-1}}{h_j}\right). \quad (3.11)$$

The commonly used kernel functions include uniform kernel function, Epanechnikov kernel function, biweight kernel function, tri-weight kernel function and Gaussian kernel function.

3.3. Nonlinear Partial Least Squares Regression (NPLS)

However, there exist some problems when establishing the above quasi-linear regression model based on ordinary least squares method, because the dimension of regression model will increase significantly after a function change, which may lead to the result that the number of sample points is less than the number of transformed variables; on the other hand, the data of reconstructed variable x_1, x_2, \dots, x_m are come from the same series, and the nonlinear correlation exists between each other. So it is not appropriate to employ the ordinary least squares method to estimate the model parameters when the multicollinear relationship among the new transformed independent variables is found. One of the effective methods to solve this problem is to use partial least squares method (PLS) to estimate parameter of the predictive model.

In this paper, a modified chaos prediction approach based on nonlinear PLS regression with basic function transformation is developed. The cubic spline function and various kernel function transformation are used in our model, and the comparative analysis between them is presented in Section 5.

3.4. Prediction Evaluation

Normally, prediction accuracy is evaluated by the mean squared error MSE,

$$\text{MSE} = \frac{1}{P} \sum_{p=1}^P (x_{T+p} - \hat{x}_{T+p})^2, \quad (3.12)$$

where \hat{x}_{T+p} is predict value, x_{T+p} is observed value. However MSE is concerned with the order of the magnitude of the observed data. In this paper, the normalized mean square error NMSE is used as the evaluation criterion of prediction.

$$\text{NMSE} = \frac{\text{MSE}}{\sigma^2}, \quad (3.13)$$

where σ^2 is sample variance of the prediction sequence. Our investigation adopts the mean of 20 predicted values NMSE in order to overcome contingency.

4. Algorithm and Algorithm Flowchart

Step 1. The observational data of aeroengine performance parameter EGT are collected and similarly corrected, and then the outliers are removed according to Layida criterion.

Step 2. For the selected embedding dimension m ($1 \leq m \leq m_{\max}$) and delay time τ ($1 \leq \tau \leq \tau_{\max}$), the chaos phase space on pretreated EGT time series is reconstructed, and the appropriate nearby points are selected to construct independent variable matrix \mathbf{X} and dependent variable vector \mathbf{y} .

Step 3. Choose section number M ($1 \leq M \leq M_{\max}$); each row vector x_j ($j = 1, 2, \dots, m$) in independent variable matrix \mathbf{X} is transformed to \mathbf{Z}_j by various basis function transformation such as (3.8), then

$$\mathbf{Z}_j = (z_{j,0}, z_{j,1}, \dots, z_{j,M+2}). \quad (4.1)$$

Step 4. The new independent variable \mathbf{Z}_j and dependent variable \mathbf{y} are standardized and marked as $(\tilde{\mathbf{Z}}, \tilde{\mathbf{y}}) = (\tilde{\mathbf{Z}}_1, \tilde{\mathbf{Z}}_2, \dots, \tilde{\mathbf{Z}}_m, \tilde{\mathbf{y}})$, so the prediction model is described as

$$\tilde{\mathbf{y}} = \sum_{j=1}^p \sum_{l=0}^{M+2} \alpha_{j,l} \tilde{z}_{j,l} + \varepsilon, \quad (4.2)$$

where $\alpha_{j,l}$ ($j = 1, 2, \dots, m; l = 0, 1, \dots, M+2$) are the parameter of regression model (4.2), ε is the stochastic error and follows the normal distribution with zero mean and the same variance.

Step 5. The partial least squares (PLS) is applied to build regression model (4.2). The number of appropriate PLS components is extracted by effective cross-method [20]; estimate the regression coefficients $\alpha_{j,l}$ ($j = 1, 2, \dots, m; l = 0, 1, \dots, M+2$), and obtain the regression coefficients $\beta_0, \beta_{j,l}$ ($j = 1, 2, \dots, m; l = 0, 1, \dots, M+2$) of original model (3.5) by returning $\alpha_{j,l}$.

Step 6. If estimation of prediction model for all the combination of M are completed, then go to step 7, otherwise go to step 3.

Step 7. Calculate the predicted value according to prediction function (3.6) and calculate the corresponding predicted NMSE.

Step 8. If the calculation of the combination of all the embedding dimension m and delay time τ is completed, then go to step 9, otherwise go to step 2.

Step 9. Find out the optimal prediction model with the smallest NMSE, and obtain the corresponding dimension m , delay time τ and section number M .

Algorithm flow chart is shown in Figure 1.

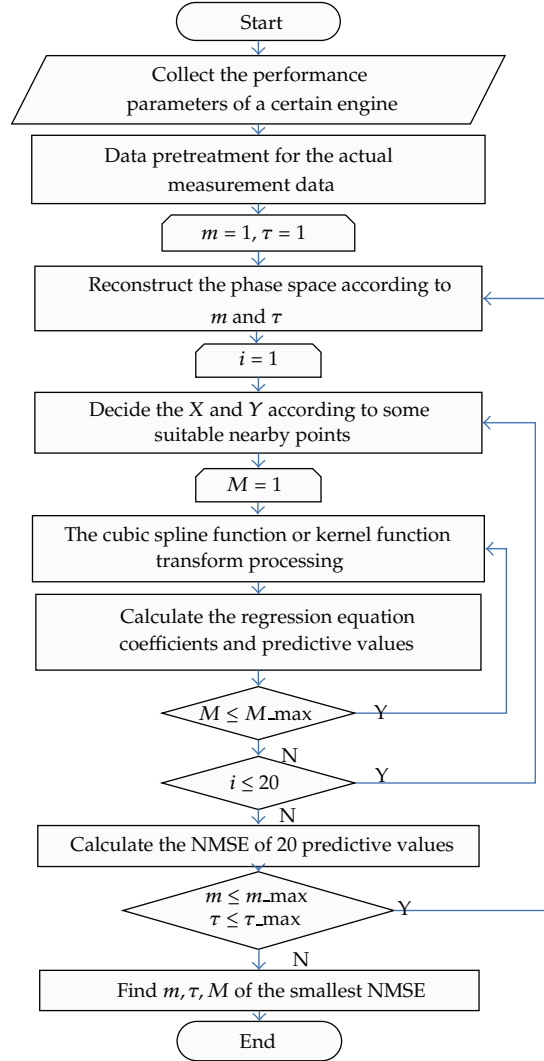


Figure 1: Algorithm flowchart.

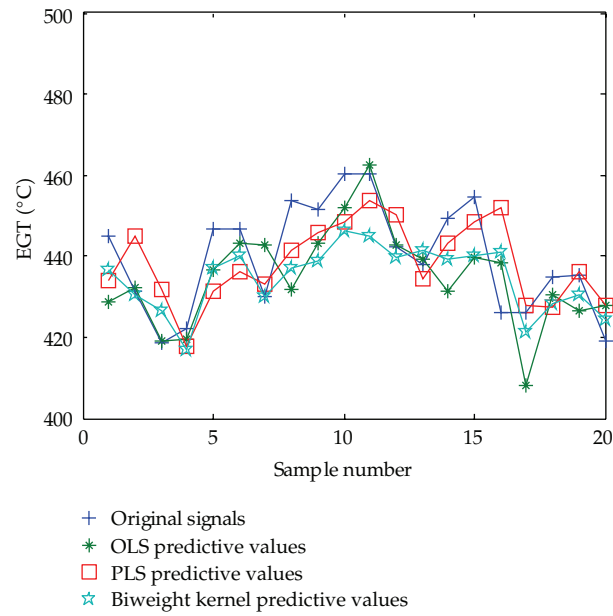
5. Results and Analysis

The in-the-wing ACARS data of four PW4077D aeroengines have been collected for one year, and about 1400 EGT data of each engine were selected to form a time series used in numerical example analysis.

The maxim value of embedding dimension m_{\max} , the delay time τ_{\max} , and the segment number M_{\max} in the above algorithm are estimated after a large amount of experiment. In our research, the range of m is among 1–30, τ is 1–20, M is 1–8, respectively. The cubic spline function and five kernel function introduced in Section 2 are employed for transformation, and we minimize chaotic prediction NMSE based on PLS algorithm.

Table 2: The optimal results after selecting independent variable by Bootstrap method based on NPLS.

Forecasting method	Embedding dimension (m)	Delay time (τ)	Subset numbers	NMSE
Cubic spline interpolation	9	3	5	0.6480
Uniform kernel	28	6	4	0.5753
Epanechnikov kernel	9	3	7	0.5917
Bi-weight kernel	8	19	2	0.5606
Triweight kernel	9	3	5	0.5786
Gaussian kernel	3	10	4	0.6510

**Figure 2:** Comparison of original values and predictive values.

The numerical results of the first engine are shown in Table 1, in which line 1 to line 6 show the prediction results based on various NPLS, line 7 to line 8 show the prediction results based on the common PLS and OLS.

As can be seen in Table 1, all the largest Lyapunov exponents are positive and it is clarified that the nonlinear and chaotic natures exist in EGT sequence. The first seven NMSE calculated based on PLS are all less than that based on OLS, which indicates that there is an obvious advantage of PLS over OLS in fitting chaotic prediction models. The NMSE of PLS based on various nonlinear transform methods is even less than that of free transformed PLS, and it indicates that there is a greater advantage of chaotic prediction based on PLS in forecasting aeroengine EGT. PLS prediction algorithm based on biweight Kernel transformation is the optimal chaotic prediction algorithm, estimated optimal embedding dimension is 8, the delay time is 19, and the number of nonlinear transformation section is 2.

PLS regression model includes all of the original selected variables, and there is multi-collinearity in the independent variables, which can affect the prediction accuracy. The model parameters estimated based on PLS is of a pretty complex nonlinear character,

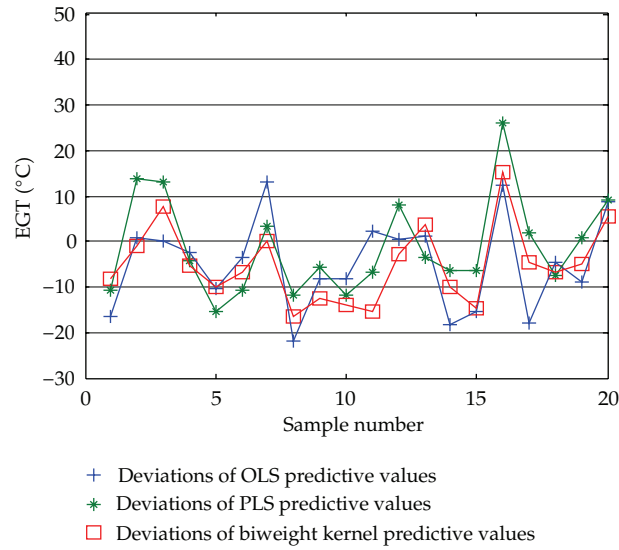


Figure 3: Deviations comparison.

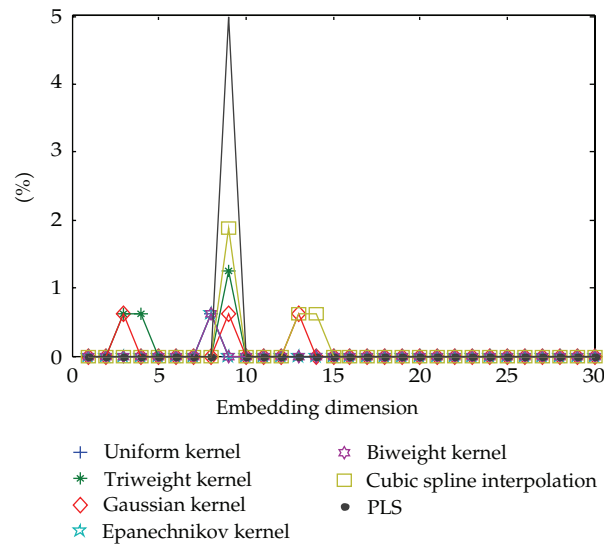


Figure 4: Embedding dimension analysis of some forecasting methods.

so the independent variables cannot be selected by the parameter test of linear methods. In this paper, the nonparametric statistical method (Bootstrap [20]) is adopted to select the independent variables transformed with kernel function and Cubic spline function, and to find the minimum NMSE; the optimal results of the first engine are shown in Table 2.

Table 2 shows the predictive results based on various NPLS algorithm after selecting independent variable by Bootstrap method. Only the optimal predictive NMSE based on triweight Kernel function transformation decreases, and the minimum NMSE in Table 2 is larger than that in Table 1; it indicates that better NMSE cannot be achieved by Bootstrap test for selecting independent variables.

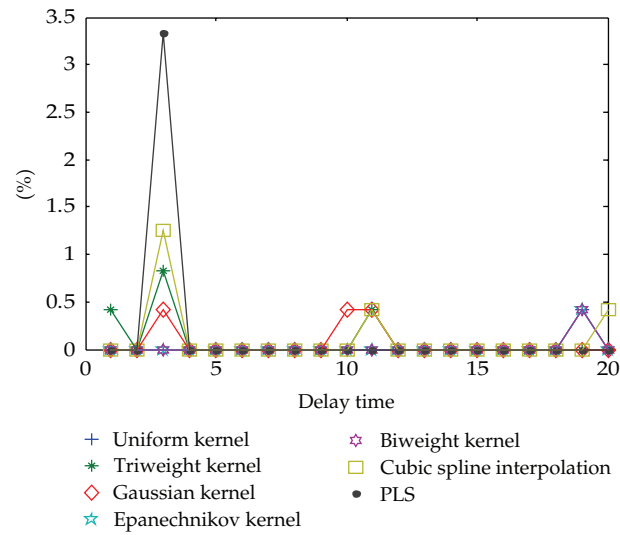


Figure 5: Delay time analysis of some forecasting method.

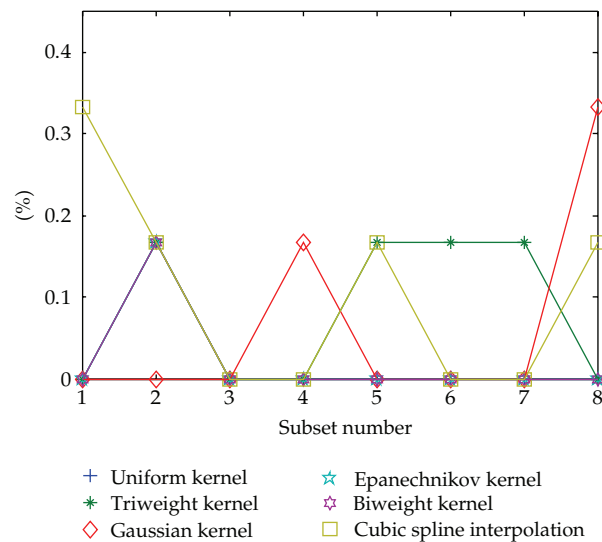


Figure 6: Subset numbers analysis of some forecasting methods.

Figure 2 shows the comparison of the prediction values based on OLS, PLS, and biweight Kernel transformation chaos algorithm with the original value. The PLS algorithm based on biweight Kernel function transformation has the optimal NMSE in Figure 2. However, not all of the predictive value has the minimum forecast NMSE.

The deviation between the predictive value and original value based on the above three algorithms are shown in Figure 3. The fluctuations of the deviations curve based on biweight Kernel method with the optimal NMSE is the smallest, and larger predicted error is not appeared in its deviation sequence.

Table 3: Optimal results of three engines based on various forecasting methods.

Forecasting method	Second aeroengine				Third aeroengine				Fourth aeroengine			
	m	τ	Subset numbers	NMSE	m	τ	Subset numbers	NMSE	m	τ	Subset numbers	NMSE
Cubic spline interpolation	8	5	4	0.6041	9	2	5	0.3354	17	3	5	0.3800
Uniform kernel	12	2	4	0.5411	7	2	2	0.3741	10	3	4	0.3529
Epanechnikov kernel	12	2	4	0.5706	9	2	2	0.3417	17	3	4	0.3329
Bi-weight kernel	12	2	4	0.5143	9	2	2	0.2574	17	3	4	0.2431
Triweight kernel	11	3	3	0.5717	9	3	4	0.2805	19	5	5	0.3918
Gaussian kernel	5	9	3	0.5675	14	2	4	0.4521	17	2	3	0.4325
PLS	26	3	0	0.6876	17	3	0	0.4688	28	3	0	0.4192
OLS	12	7	0	0.7515	8	11	0	0.5742	16	7	0	0.4937

Through the numerical results of the first engine we also find that the search range of designed algorithm is enormous and it takes a long time, so we hope to further reduce search range and the algorithm complexity. The percentages of the number of NMSE in Table 1 which falls in $[\min(\text{NMSE}), \min(\text{NMSE}) + 0.03]$ are statistically analyzed, and the results are shown in Figures 4, 5 and 6.

As can be seen from Figure 4 to Figure 6, the percentage of NMSE falling in $[\min(\text{NMSE}), \min(\text{NMSE}) + 0.03]$ is lower for different value of m , τ , M , which indicates that the probability of obtaining suboptimal solution by our algorithm is very low, and also illustrates that the selection of variables is of significant importance to improve the prediction accuracy. From Figure 4 we can see that when the dimension of predicted model is over 15, there is no optimal NMSE which falls within the range above mentioned, So the range of m in the algorithm can be selected among ranging from 1 to 15. Figures 5 and 6 also show that the predicted NMSE distribution of various delay time and section number are more dispersed, so the search range of τ and M cannot be minimized.

The brief computational results of the other three engines are as in Table 3.

From the above calculation results we can see that the algorithm based on various nonlinear PLS (Cubic Spline Interpolation, Uniform Kernel, Epanechnikov Kernel, biweight Kernel, Tri-weight Kernel, and Gaussian Kernel) is superior to that based on PLS, and the algorithm based on PLS is superior to that based on OLS. The results in Tables 1 and 3 show that the predictive results derived from the algorithm based on biweight Kernel are the optimal among the four aeroengines. And there are no significant difference between the first aeroengines and the second aeroengines in the same aircraft in predictive NMSE, so we can conclude that the proposed algorithm is stable for predicting the aeroengine parameter EGT.

Tables 1 and 3 also show that the optimal values of the embedding dimension (m), delay time (τ), subset numbers (M) searched by various NPLS algorithm are approximately the same, but they are quite different among the four aeroengines.

6. Conclusion

The chaos prediction model based on NPLS through Kernel function or Cubic Spline function transformation is developed and is applied in predicting the aeroengine performance

parameters EGT. Numerical results show that the proposed model is able to effectively predict the performance parameter EGT with a higher degree of accuracy. The optimal embedding dimension is 8, the delay time is 19, and the number of segments is 2. NPLS chaotic prediction algorithm based on biweight Kernel function transformation is the optimal EGT prediction algorithm because of the smallest predictive NMSE. Our prediction NMSE is 16.5 percent less than that of the traditional linear least squares (OLS) method and 10.38 percent less than that of the linear PLS approach. The NPLS algorithm of chaotic prediction with selection variables through Bootstrap test fails to make NMSE further reduced. And we find that reducing calculation time complexity by lessening search range of embedding dimension will not affect the prediction accuracy of EGT. It is proved that the proposed algorithm is robust by the predictive results from four aeroengines.

In particular, the conservative baselines of main aeroengine performance parameters are given in ECM program provided by jet engine manufacturers [14]. Set the baseline value of EGT be EGT_{base} , the deviation between EGT_{cor} and EGT_{base} can be defined as

$$\Delta EGT = EGT_{cor} - EGT_{base}. \quad (6.1)$$

The trend plot of the deviation (ΔEGT) sequence in continuous flights is routinely used by airlines for the performance monitoring and diagnostics [13]. On the basis of our prediction algorithm, the EGT_{cor} of upcoming flights can be predicted, and then a layout for trending of the deviation of main performance parameters for upcoming flights can be provided. Experienced power maintainer or engineer look at a given trend shift in the measurement deviations can identify significant behavior disparities, evaluate the aeroengine condition, and isolate the faulty module or component using the ECM trend plot reports and related PW4000 fingerprint.

In summary, accurately forecasting the aeroengine performance parameters is of great importance in the aeroengine condition monitoring. The anomalies of the aeroengine performance parameters can be detected in time by the predictive value calculated from our algorithm, then maintenance measures are taken early to prevent the aeroengine from failure effectively, which assists to reduce aircraft down time, and improve the reliability of the aeroengine.

Acknowledgment

This work is supported by the Fundamental Research Funds for the Central Universities under Grant no. ZXH2012 K005.

References

- [1] E. Ortiz, A. Babbar, and V. L. Syrmos, "Extreme value theory for engine health monitoring and diagnosis," in *Proceedings of the IEEE International Conference on Control Applications (CCA '09)*, pp. 1069–1074, Saint Petersburg, Russia, July 2009.
- [2] A. Babbar, E. M. Ortiz, V. L. Syrmos, and M. M. Arita, "Advanced diagnostics and prognostics for engine health monitoring," in *Proceedings of the IEEE Aerospace Conference*, March 2009.
- [3] Z. Gao, X. Dai, T. Breikin, and H. Wang, "Novel parameter identification by using a high-gain observer with application to a gas turbine engine," *IEEE Transactions on Industrial Informatics*, vol. 4, no. 4, pp. 271–279, 2008.

- [4] Z. Gao, T. Breikin, and H. Wang, "High-gain estimator and fault-tolerant design with application to a gas turbine dynamic system," *IEEE Transactions on Control Systems Technology*, vol. 15, no. 4, pp. 740–753, 2007.
- [5] T. S. Parker and L. O. Chua, *Practical Numerical Algorithms for Chaotic Systems*, Springer, New York, NY, USA, 1989.
- [6] R. Gençay and T. Liu, "Nonlinear modelling and prediction with feedforward and recurrent networks," *Physica D*, vol. 108, no. 1-2, pp. 119–134, 1997.
- [7] W. Niu, G. Wang, Z. Zhai, and J. Cheng, "Prediction of chaotic time series based on rough sets and support vector machine," *International Journal of Digital Content Technology and Its Applications*, vol. 5, no. 9, pp. 145–152, 2011.
- [8] L. X. Wang and J. M. Mendel, "Generating fuzzy rules by learning from examples," *Institute of Electrical and Electronics Engineers. Transactions on Systems, Man, and Cybernetics*, vol. 22, no. 6, pp. 1414–1427, 1992.
- [9] D. Kugiumtzis, O. C. Lingjærde, and N. Christophersen, "Regularized local linear prediction of chaotic time series," *Physica D*, vol. 112, no. 3-4, pp. 344–360, 1998.
- [10] G. Ding and S. Zhong, "Approximation capability analysis of parallel process neural network with application to aircraft engine health condition monitoring," in *Proceedings of the 4th International Symposium on Neural Networks (ISNN '07)*, vol. 4493 of *Lecture Notes in Computer Science*, pp. 66–72, 2007.
- [11] C. Zhang and J. Yue, "Application of an improved adaptive chaos prediction model in aero-engine performance parameters," *WSEAS Transactions on Mathematics*, vol. 11, no. 2, pp. 114–124, 2012.
- [12] L. Zunxiong and L. Jianhui, "Chaotic time series multi-step direct prediction with partial least squares regression," *Journal of Systems Engineering and Electronics*, vol. 18, no. 3, pp. 611–615, 2007.
- [13] E. J. Mo, M. S. Jie, C. S. Kim, and K. W. Lee, "Gas turbine engine fault isolation based fuzzy inference logic with ECM trend plot report," in *Proceedings of the International Conference on Control, Automation and Systems (ICCAS '07)*, pp. 454–458, Seoul, Korea, October 2007.
- [14] Pratt&Whitney Company, "ECM II Training manual," Section 2, 1-34, 1994.
- [15] D. M. Hawkins, *Identification of Outliers*, Chapman & Hall, London, UK, 1980.
- [16] L. D. Iasemidis, K. Tsakalis, J. C. Sackellares et al., "Comment on 'inability of Lyapunov exponents to predict epileptic seizures'," *Physical Review Letters*, vol. 94, no. 1, Article ID 019801, 2005.
- [17] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, "Determining Lyapunov exponents from a time series," *Physica D*, vol. 16, no. 3, pp. 285–317, 1985.
- [18] Z. Wang, *Of Time Series*, China Statistics Press, Beijing, China, 2005.
- [19] J. Zeyu, Y. Yue, and Z. Wenzhong, "Application of improved partial least square regressive model in power load forecasting," *Power DSM*, vol. 13, no. 1, pp. 10–14, 2011.
- [20] D. V. Pisarenko and V. F. Pisarenko, "Statistical estimation of the correlation dimension," *Physics Letters A*, vol. 197, no. 1, pp. 31–39, 1995.

Research Article

Adaptive Control of a Two-Item Inventory Model with Unknown Demand Rate Coefficients

Ahmad M. Alshamrani

Department of Statistics and Operations Research, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia

Correspondence should be addressed to Ahmad M. Alshamrani, ahmadm@ksu.edu.sa

Received 23 December 2011; Revised 8 April 2012; Accepted 22 April 2012

Academic Editor: Chuanhou Gao

Copyright © 2012 Ahmad M. Alshamrani. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper considers a multiitem inventory model with unknown demand rate coefficients. An adaptive control approach with a nonlinear feedback is applied to track the output of the system toward the inventory goal level. The Lyapunov technique is used to prove the asymptotic stability of the adaptive controlled system. Also, the updating rules of the unknown demand rate coefficients are derived from the conditions of the asymptotic stability of the perturbed system. The linear stability analysis of the model is discussed. The adaptive controlled system is modeled by a system of nonlinear differential equations, and its solution is discussed numerically.

1. Introduction

The area of adaptive control has grown rapidly to be one of the richest fields in the control theory. Many books and research monographs already exist on the topics of parameter estimation and adaptive control. Adaptive control theory is found to be very useful in solving many problems in different fields, such as management science, dynamic systems, and inventory systems [1–3].

- (i) El-Gohary and Yassen [4] used an adaptive control approach and synchronization procedures to the coupled dynamo system with unknown parameters. Based on the Lyapunov stability technique, an adaptive control laws were derived such that the coupled dynamo system is asymptotically stable and the two identical dynamo systems are asymptotically synchronized. Also the updating rules of the unknown parameters were derived;

- (ii) El-Gohary and Al-Ruzaiza [5] studied the adaptive control of a continuous-time three-species prey-predator populations. They have derived the nonlinear feedback control inputs which asymptotically stabilized the system about its steady states;
- (iii) Tadj et al. [6] discussed the optimal control of an inventory system with ameliorating and deteriorating items. They considered different cases for the difference between the ameliorating and deteriorating items;
- (iv) Foul et al. [7] studied the problem of adaptive control of a production and inventory system in which a manufacturing firm produces a single product, then it sells some of its production and stocks the remaining. They applied a model reference adaptive control system with a feedback to track the output of the system toward the inventory goal level;
- (v) Alshamrani and El-Gohary [8] studied the problem of optimal control of the two-item inventory system with different types of deterioration. They derived the optimal inventory levels and continuous rates of supply from the optimality conditions;
- (vi) Many other studies which are concerned with the production and inventory systems, multiitem inventory control, and inventory analysis can be found in references [9–14]. Such studies have discussed the optimal control of a multi-item inventory model, the stability conditions of a multi-item inventory model with different demand rates, and the optimal control of multi-item inventory systems with budgetary constraints.

This paper is concerned with a two-item inventory system with different types of deteriorating items subjected to unknown demand rate coefficients. We derive the controlled inventory levels and continuous rates of supply. Further, the updating rules of the unknown demand rate coefficients are derived from the conditions of asymptotic stability of the reference model. The resulting controlled system is modeled by a system of nonlinear differential equations, and its solution is discussed numerically for different sets of parameters and initial states.

The motivation of this study is to extend and generalize the two-item inventory system with different types of deterioration and applying an adaptive control approach to this system in order to get an asymptotic controlled system. This paper generalizes some of the models available in the literature, see for example, [6, 9].

The rest of this paper is organized as follows. In Section 2, we present the mathematical model of the two-item system. Also, stability analysis of the model is discussed in Section 2. Section 3 discusses the adaptive control problem of the system. Numerical solution and examples are presented in Section 4. Finally, Section 5 concludes the paper.

2. The Two-Item Inventory System

This section uses the mathematical methods to formulate the two-item inventory system with two different type of deteriorations. In this model, we consider a factory producing two items and having a finished goods warehouse.

2.1. The Model Assumptions and Formulation

This subsection is devoted to introduce the model assumptions and its formulation. It is assumed that the inventory supply rates are equal to the production rates, while the demand rates may adopt two different types. Throughout this paper we use $i, j = 1, 2$ for the two different types of items. Moreover, the following variables and parameters are used:

- $x_i(t)$: the i th item inventory level at time t ;
- $u_i(t)$: rate of continuous supply to x_i at time t ;
- x_{i0} : the i th item initial state inventory level;
- $d_i x_i$: linear demand rate at instantaneous level of inventory x_i , where d_i is a constant;
- θ_{ii} : the deterioration coefficient due to self-contact of the i th item inventory level x_i ;
- a_{ij} : the demand coefficient of x_i due to presence of units of x_j , where $(i \neq j)$;
- θ_i : the natural deterioration rate of the i th item inventory level x_i ;
- \bar{x}_i : the value of the i th item inventory level at the steady state;
- \bar{u}_i : the value of the i th item continuous rate of supply at the steady state;
- $\hat{a}_{ij}(t)$: the dynamic estimator of demand coefficient of x_i due to presence of units of x_j , where $(i \neq j)$.

The main problem of this paper is to present the adaptive control problem for the two-item inventory system as a control problem with two state variables and two control variables which are the inventory levels $x_i(t)$, $i = 1, 2$ and the two continuous supply rates $u_i(t)$, $i = 1, 2$, respectively.

Also, since an analytical solution of the resulting control system is nonlinear and its analytical solution is not available, we solve it numerically and display the solution graphically. We show that the solution of the adaptive controlled system covers different modes of demand rates.

2.2. The Mathematical Model and Stability Analysis

In this subsection, we present a suitable mathematical form for a two-item inventory system with two types of deteriorations. This mathematical form must be simple to deal with any response of the two-item inventory model with deterioration to any given input. The differential equations system that governs the time evolution of the two-item inventory system is found to be as follows [8]:

$$\begin{aligned}\dot{x}_1(t) &= u_1(t) - x_1(t)(d_1 + \theta_1 + a_{12}x_2(t) + \theta_{11}x_1(t)), \\ \dot{x}_2(t) &= u_2(t) - x_2(t)(d_2 + \theta_2 + a_{21}x_1(t) + \theta_{22}x_2(t)),\end{aligned}\tag{2.1}$$

with the following nonnegatively conditions:

$$x_i(t) \geq 0, \quad u_i(t) \geq 0, \quad d_i(t) > 0, \quad \theta_i(t) > 0, \quad \theta_{ii} > 0, \quad i = 1, 2,\tag{2.2}$$

and with the following boundary conditions:

$$x_i(0) = x_{i0}, \quad i = 1, 2. \quad (2.3)$$

In this paper, we consider the inventory goal levels \bar{x}_i and the goal rates of the continuous rate of supply \bar{u}_i to be their values at the steady state of the system. The advantage of this study is to prove the asymptotic stability of the two-item inventory system using the Liapunov technique about the steady state of the system.

Next, we will derive the steady state solution of (2.1). The steady state of the system (2.1) can be derived by putting both of $\dot{x}_1(t)$ and $\dot{x}_2(t)$ equal zero, that is,

$$\begin{aligned} \bar{x}_1(d_1 + \theta_1 + a_{12}\bar{x}_2 + \theta_{11}\bar{x}_1) - \bar{u}_1 &= 0, \\ \bar{x}_2(d_2 + \theta_2 + a_{21}\bar{x}_1 + \theta_{22}\bar{x}_2) - \bar{u}_2 &= 0. \end{aligned} \quad (2.4)$$

Solving (2.4), we get \bar{x}_1 as a function of \bar{x}_2 as follows:

$$\bar{x}_1 = \left[-(d_1 + \theta_1 + a_{12}\bar{x}_2) \pm \sqrt{(d_1 + \theta_1 + a_{12}\bar{x}_2)^2 + 4\theta_{11}\bar{u}_1} \right] (\theta_{11})^{-1}, \quad (2.5)$$

where the values of \bar{x}_2 are the roots of the equation:

$$\left(\theta_{22} - \frac{a_{12}a_{21}}{2\theta_{11}} \right) \bar{x}_2^2 + \frac{a_{21}}{2\theta_{11}} \left[\sqrt{(d_1 + \theta_1 + a_{12}\bar{x}_2)^2 + 4\theta_{11}\bar{u}_1} - d_1 - \theta_1 \right] \bar{x}_2 + d_2 + \theta_2 - \bar{u}_2 = 0. \quad (2.6)$$

In what follows, we discuss the numerical solution for the (2.4) for fixed values of the parameters $d_i, \theta_i, \theta_{ii}, (i = 1, 2)$ and a_{12}, a_{21} :

- (1) in this example, we discuss the numerical solution of (2.4) for constant rates of supply $\bar{u}_1 = 2.25$ and $\bar{u}_2 = 3.25$, the steady states are given in Table 1;
- (2) in this example, we discuss the numerical solution of (2.4) for the supply rates $\bar{u}_1 = 5\bar{x}_1 + 6\bar{x}_2$ and $\bar{u}_2 = 45\bar{x}_1 + 25\bar{x}_2$ of the inventory levels, the steady states are given in Table 2;
- (3) in this example, we discuss the numerical solution of (2.4) for supply rates $\bar{u}_1 = 2\bar{x}_1^2 + 3\bar{x}_2^2 + 5\bar{x}_1\bar{x}_2$ and $\bar{u}_2 = 5\bar{x}_1^2 + 15\bar{x}_2^2 + 45\bar{x}_1\bar{x}_2$ of the inventory levels, the steady states are given in Table 3.

Figure 1 displays the numerical solution for the two-item inventory system with the quadratic continuous rates of supply $\bar{u}_1 = \alpha\bar{x}_1\bar{x}_2$ and $\bar{u}_2 = \beta\bar{x}_1\bar{x}_2$, with the initial inventory levels $x_1(0) = 3$ and $x_2(0) = 15$, and with the following set of parameters in Table 4.

Figure 2 displays the numerical solution for the two-item inventory with constant continuous rates of supply $\bar{u}_1 = \alpha$ and $\bar{u}_2 = \beta$, with the initial inventory levels $x_1(0) = 3$ and $x_2(0) = 5$, and the following set of parameters in Table 5.

2.3. Linear Stability Analysis

The concept of stability is concerned with the investigation and characterization of the behavior of dynamic systems. Stability analysis plays a crucial role in system theory and

Table 1

Model parameters	Steady states (\bar{x}_1, \bar{x}_2)
$d_1 = 10, d_2 = 20, a_{12} = 2.5, a_{21} = 3.5,$	$(-1010, -0.00996)$, negative
$\theta_1 = 0.1, \theta_2 = 0.15, \theta_{11} =$	$(-6.4, -4.624)$, negative
0.01 , and $\theta_{22} = 0.01$	$(0.225, 0.156)$, positive

Table 2

Model parameters	Steady states (\bar{x}_1, \bar{x}_2)
$d_1 = 10, d_2 = 20, a_{12} = 25, a_{21} =$	$(0, 0)$, null state
$45, \theta_1 = 0.05, \theta_2 = 0.25, \theta_{11} =$	$(-149.81, 0.995)$, negative
0.2 , and $\theta_{22} = 0.2$	$(0.22, 1.93)$, positive

control engineering and has been investigated extensively in the past century. Some of the most fundamental concepts of stability were introduced by the Russian mathematician and engineer Alexandr Lyapunov in [15].

In this section, we discuss the linear stability analysis of the system (2.1) about its steady states (2.4). We classify the roots of the characteristic equation of the Jacobian matrix of the system (2.1) about its steady states (2.4).

The characteristic equation is given by:

$$\lambda^2 - b\lambda + c = 0, \quad (2.7)$$

where the coefficients b and c are:

$$b = \left\{ \sum_{i=1}^2 \left(\frac{\partial u_i}{\partial x_i} - \frac{u_i}{x_i} - \theta_{ii} x_i \right) \right\}_{(x_1, x_2) = (\bar{x}_1, \bar{x}_2)}$$

$$c = \left\{ \left(\frac{\partial u_1}{\partial x_1} - \frac{u_1}{x_1} - \theta_{11} x_1 \right) \left(\frac{\partial u_2}{\partial x_2} - \frac{u_2}{x_2} - \theta_{22} x_2 \right) - \left(\frac{\partial u_1}{\partial x_2} - a_{12} x_1 \right) \left(\frac{\partial u_2}{\partial x_1} - a_{21} x_2 \right) \right\}_{(x_1, x_2) = (\bar{x}_1, \bar{x}_2)}. \quad (2.8)$$

The roots of the characteristic equation are:

$$\lambda = b \pm \sqrt{b^2 - 4c}. \quad (2.9)$$

The roots of the characteristic equation will be complex numbers with negative real parts if the following conditions can be satisfied:

$$\sum_{i=1}^2 \left[\frac{\partial u_i}{\partial x_i} - \frac{u_i}{x_i} - \theta_{ii} x_i \right]_{(x_1, x_2) = (\bar{x}_1, \bar{x}_2)} < 0,$$

$$\left\{ \left(\frac{\partial u_1}{\partial x_1} - \frac{\partial u_2}{\partial x_2} - \frac{u_1}{x_1} + \frac{u_2}{x_2} - \theta_{11} x_1 + \theta_{22} x_2 \right)^2 - 4 \left(\frac{\partial u_1}{\partial x_2} - a_{12} x_1 \right) \left(\frac{\partial u_2}{\partial x_1} - a_{21} x_2 \right) \right\}_{(x_1, x_2) = (\bar{x}_1, \bar{x}_2)} < 0. \quad (2.10)$$

Table 3

Model parameters	Steady states (\bar{x}_1, \bar{x}_2)
$d_1 = 50, d_2 = 40, a_{12} = 75, a_{21} = 55, \theta_1 = 0.15, \theta_2 = 0.25, \theta_{11} = 0.2, \text{ and } \theta_{22} = 0.2$	$(0, 0)$, null state
	$(-2.4 + 1.72I, -0.786 + 0.04I)$, imaginary
	$(0.099, 2.9)$, positive

Table 4

Parameter	θ_{11}	θ_{22}	a_{12}	a_{21}	d_1	d_2	θ_1	θ_2	\bar{u}_1	\bar{u}_2	α	β
Value	0.12	0.15	6	5	6	8	0.02	0.04	100	50	8	5

Therefore the system (2.1) is stable in the linear sense if the conditions (2.10) are satisfied, otherwise this system is absolutely unstable. The absolutely stability of system (2.1) needs further complicated mathematical analysis.

The roots of the characteristic equation will be negative real numbers if the following conditions can be satisfied:

$$\sum_{i=1}^2 \left[\frac{\partial u_i}{\partial x_i} - \frac{u_i}{x_i} - \theta_{ii} x_i \right]_{(x_1, x_2) = (\bar{x}_1, \bar{x}_2)} < 0,$$

$$\left\{ \left(\frac{\partial u_1}{\partial x_1} - \frac{\partial u_2}{\partial x_2} - \frac{u_1}{x_1} + \frac{u_2}{x_2} - \theta_{11} x_1 + \theta_{22} x_2 \right)^2 + 4 \left(\frac{\partial u_1}{\partial x_2} - a_{12} x_1 \right) \left(\frac{\partial u_2}{\partial x_1} - a_{21} x_2 \right) \right\}_{(x_1, x_2) = (\bar{x}_1, \bar{x}_2)} > 0$$

$$\left\{ \left(\frac{\partial u_1}{\partial x_1} - \frac{u_1}{x_1} - \theta_{11} x_1 \right) \left(\frac{\partial u_2}{\partial x_2} - \frac{u_2}{x_2} - \theta_{22} x_2 \right) - \left(\frac{\partial u_1}{\partial x_2} - a_{12} x_1 \right) \left(\frac{\partial u_2}{\partial x_1} - a_{21} x_2 \right) \right\}_{(x_1, x_2) = (\bar{x}_1, \bar{x}_2)} > 0. \quad (2.11)$$

If the conditions (2.11) are satisfied, then the system (2.1) is stable in the linear sense, otherwise this system is absolutely unstable. The absolutely stability of system (2.1) needs further complicated mathematical analysis.

Next, we discuss some special cases in which the rates of supply take different functions of the inventory levels:

- (1) when the supply rates do not depend on the inventory levels, the linear stability conditions are reduced to

$$\frac{\bar{x}_1}{\bar{x}_2} < \frac{\bar{u}_1 + \theta_{11} \bar{x}_1^2}{\bar{u}_2 + \theta_{22} \bar{x}_2^2} \quad (2.12)$$

or

$$\frac{\bar{u}_1 + \theta_{11} \bar{x}_1^2}{a_{12} \bar{x}_1^2} < \frac{\bar{u}_2 + \theta_{22} \bar{x}_2^2}{a_{21} \bar{x}_2^2}, \quad (2.13)$$

Table 5

Parameter	θ_{11}	θ_{22}	a_{12}	a_{21}	d_1	d_2	θ_1	θ_2	α	β
Value	0.2	0.5	0.6	0.5	0.6	0.8	0.02	0.04	8	5

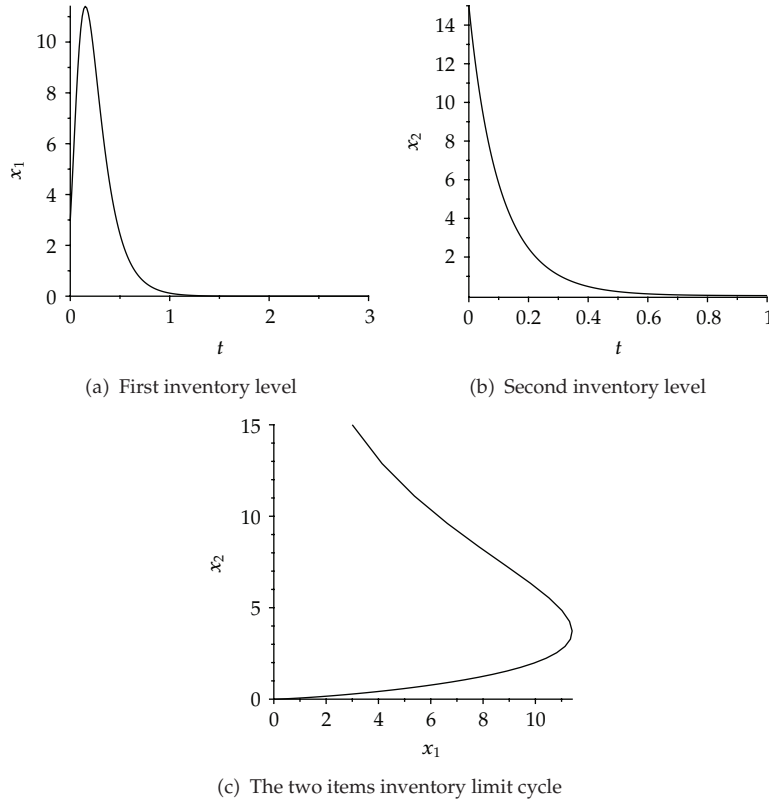


Figure 1: (a) and (b) are the first and the second inventory levels, respectively, of the uncontrolled system, with quadratic rates of supply. (c) is the trajectory of the inventory system in x_1x_2 -plane.

- (2) when the supply rates are linear function of the inventory levels, $u_i = \alpha_i x_i$, $i = 1, 2$, the linear stability conditions are reduced to

$$(\theta_{22}\bar{x}_2 - \theta_{11}\bar{x}_1)^2 < 4a_{12}a_{21}\bar{x}_1\bar{x}_2 \quad (2.14)$$

or

$$(\alpha_1 + \theta_{11}\bar{x}_1)(\alpha_2 + \theta_{22}\bar{x}_2) > a_{12}a_{21}\bar{x}_1\bar{x}_2; \quad (2.15)$$

- (3) when the supply rates are quadratic functions of the inventory levels, $u_i = \alpha_i x_1 x_2$, $i = 1, 2$, the linear stability conditions are reduced to

$$(\theta_{22}x_2 - \theta_{11}\bar{x}_1)^2 < 4(\alpha_1 - a_{12})(\alpha_2 - a_{21})\bar{x}_1\bar{x}_2 \quad (2.16)$$

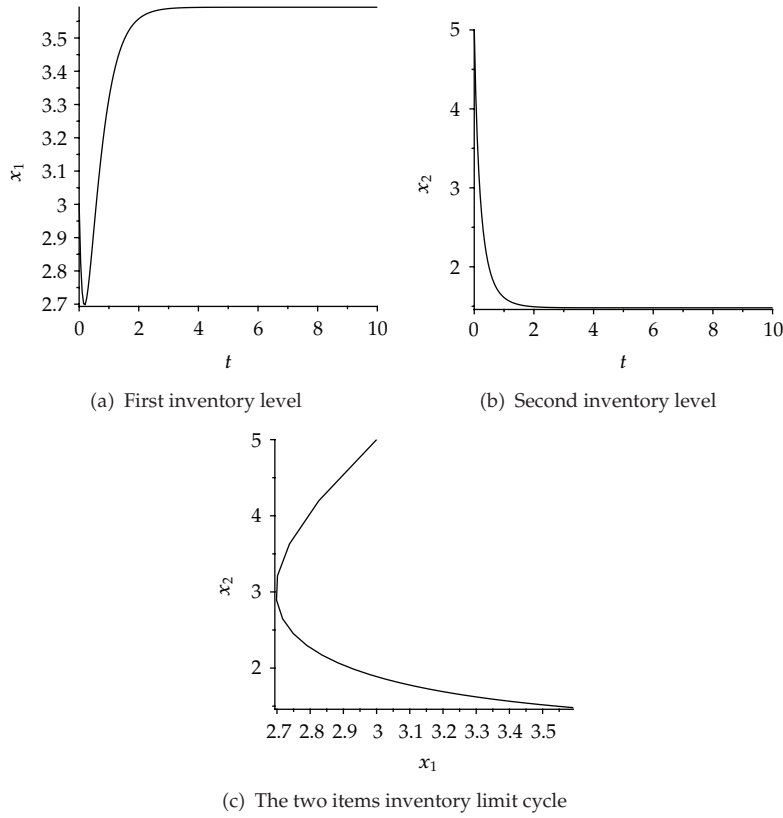


Figure 2: (a) and (b) are the first and the second inventory levels of the uncontrolled system, with constant rates of supply. (c) is the trajectory of the inventory system in x_1x_2 -plane.

or

$$\begin{aligned}
 (\theta_{22}x_2 - \theta_{11}\bar{x}_1)^2 + 4(\alpha_1 - a_{12})(\alpha_2 - a_{21})\bar{x}_1\bar{x}_2 &> 0, \\
 \theta_{11}\theta_{22} &> (\alpha_1 - a_{12})(\alpha_2 - a_{21}).
 \end{aligned}
 \tag{2.17}$$

In what follows, we study the problem of adaptive control. In order to study this problem, we start by obtaining the perturbed system of the two-item inventory model about its steady states (\bar{x}_1, \bar{x}_2) . To obtain this perturbed system, we introduce the following new variables:

$$\xi_i(t) = x_i(t) - \bar{x}_i, \quad v_i(t) = u_i(t) - \bar{u}_i, \quad (i = 1, 2), \tag{2.18}$$

Substituting from (2.18) into (2.1) and using (2.4), we get

$$\begin{aligned}
 \dot{\xi}_1(t) &= -\xi_1[d_1 + \theta_1 + a_{12}\xi_2 + a_{12}\bar{x}_2 + \theta_{11}(\xi_1 + 2\bar{x}_1)] - a_{12}\bar{x}_1\xi_2 + v_1, \\
 \dot{\xi}_2(t) &= -\xi_2[d_2 + \theta_2 + a_{21}\xi_1 + a_{21}\bar{x}_1 + \theta_{22}(\xi_2 + 2\bar{x}_2)] - a_{21}\bar{x}_2\xi_1 + v_2.
 \end{aligned}
 \tag{2.19}$$

The system (2.19) will be used to study the problem of adaptive control of the two-item inventory model with deteriorating item and unknown demand coefficients.

In adaptive control systems, we are concerned with changing the properties of dynamic systems so that they can exhibit acceptable behavior when perturbed from their operating point using a feedback approach.

3. The Adaptive Control Problem

The problem that we address in this section is the adaptive control of the two-item inventory system with different types of deterioration which are subjected to unknown demand rate coefficients. In such study, we assume that the demand coefficients a_{12} and a_{21} are unknown parameters. So we assume that the functions $\hat{a}_{12}(t)$ and $\hat{a}_{21}(t)$ represent their dynamic estimators. Using this assumption, we can rewrite the system (2.19) in the following form:

$$\begin{aligned}\dot{\xi}_1(t) &= -\xi_1[d_1 + \theta_1 + \hat{a}_{12}\xi_2 + \hat{a}_{12}\bar{x}_2 + \theta_{11}(\xi_1 + 2\bar{x}_1)] - \hat{a}_{12}\bar{x}_1\xi_2 + v_1, \\ \dot{\xi}_2(t) &= -\xi_2[d_2 + \theta_2 + \hat{a}_{21}\xi_1 + \hat{a}_{21}\bar{x}_1 + \theta_{22}(\xi_2 + 2\bar{x}_2)] - \hat{a}_{21}\bar{x}_2\xi_1 + v_2.\end{aligned}\quad (3.1)$$

The adaptive law is usually a differential equation whose state is designed using stability considerations or simple optimization techniques to minimize the difference between the state and its estimator with respect to the state at each time t .

In what follows, we discuss the asymptotic stability of the special solution of the system (3.1) which is given by

$$\xi_i(t) = 0, \quad v_i = 0, \quad (i = 1, 2), \quad \hat{a}_{12}(t) = a_{12}, \quad \hat{a}_{21}(t) = a_{21}. \quad (3.2)$$

This solution corresponds to the steady states solution of the system (2.1). So the asymptotic stability of this solution leads to the asymptotic stability of the (2.1) about its steady state.

The following theorem determines both of the perturbations of the continuous rates of supply v_i and the updating rules of $\hat{a}_{12}(t)$ and $\hat{a}_{21}(t)$ of demand rate coefficients from the conditions of the asymptotic stability of the solution (3.2).

Theorem 3.1. *If the perturbations of the continuous supply rates and the updating rules of the unknown parameters $\hat{a}_{12}(t)$ and $\hat{a}_{21}(t)$ are given by*

$$v_1(t) = a_{12}\bar{x}_1\xi_2 + a_{12}\xi_1\xi_2 + \theta_{11}\xi_1^3 - k_1\xi_1, \quad (3.3)$$

$$v_2(t) = a_{21}\bar{x}_2\xi_1 + a_{21}\xi_1\xi_2 + \theta_{22}\xi_2^3 - k_2\xi_2,$$

$$\dot{\hat{a}}_{12}(t) = \bar{x}_2\xi_1\xi_2 + \bar{x}_2\xi_1^2 + \xi_1^2\xi_2 - m_1(\hat{a}_{12} - a_{12}), \quad (3.4)$$

$$\dot{\hat{a}}_{21}(t) = \bar{x}_1\xi_1\xi_2 + \bar{x}_1\xi_2^2 + \xi_1\xi_2^2 - m_2(\hat{a}_{21} - a_{21}),$$

where k_i, m_i , and $(i = 1, 2)$ are positive real constant, then the solution (3.2) is asymptotically stable in the Liapunov sense.

Proof. The proof of this theorem can be reached by using the Liapunov technique. Assume that the Liapunov function of the system of equations (3.2) and (3.4) is

$$2V(\xi_i, \hat{a}_{12}, \hat{a}_{21}) = \sum_{i=1}^2 \xi_i^2 + (\hat{a}_{12} - a_{12})^2 + (\hat{a}_{21} - a_{21})^2. \quad (3.5)$$

Differentiating the function V in (3.5):

$$\dot{V} = \xi_1 \dot{\xi}_1 + \xi_2 \dot{\xi}_2 + (\hat{a}_{12} - a_{12})\dot{\hat{a}}_{12} + (\hat{a}_{21} - a_{21})\dot{\hat{a}}_{21}. \quad (3.6)$$

Substituting from (3.1) into (3.6), we get

$$\begin{aligned} \dot{V} = & \xi_1 \left[-(d_1 + \theta_1)\xi_1 - \bar{x}_1 \hat{a}_{12} \xi_2 - \bar{x}_2 \hat{a}_{12} \xi_1 - \hat{a}_{12} \xi_1 \xi_2 - \theta_{11} \xi_1^2 - 2\bar{x}_1 \xi_1 + v_1 \right] \\ & + \xi_2 \left[-(d_2 + \theta_2)\xi_2 - \bar{x}_2 \hat{a}_{21} \xi_1 - \bar{x}_1 \hat{a}_{21} \xi_2 - \hat{a}_{21} \xi_1 \xi_2 - \theta_{22} \xi_2^2 - 2\bar{x}_2 \xi_2 + v_2 \right] \\ & + (\hat{a}_{12} - a_{12})\dot{\hat{a}}_{12} + (\hat{a}_{21} - a_{21})\dot{\hat{a}}_{21}. \end{aligned} \quad (3.7)$$

Substituting from (3.1), (3.3), and (3.4) into (3.7), and after some simple calculations, we get

$$\begin{aligned} \dot{V} = & - \left[m_1 (\hat{a}_{12} - a_{12})^2 + m_2 (\hat{a}_{21} - a_{21})^2 + (d_1 + \theta_1 + a_{12} \bar{x}_2 + 2\theta_{11} \bar{x}_1) \xi_1^2 \right. \\ & \left. + (d_2 + \theta_2 + a_{21} \bar{x}_1 + 2\theta_{22} \bar{x}_2) \xi_2^2 \right], \end{aligned} \quad (3.8)$$

since the coefficients $d_1 + \theta_1 + a_{12} \bar{x}_2 + 2\theta_{11} \bar{x}_1$ and $d_2 + \theta_2 + a_{21} \bar{x}_1 + 2\theta_{22} \bar{x}_2$ are positive, then \dot{V} is a negative definite function of ξ_i , \hat{a}_{12} , and \hat{a}_{21} , so the solution (3.3) is asymptotically stable in the Liapunov sense, which completes the proof. \square

In Section 4, we will discuss the numerical solution of the controlled two-item inventory system with unknown demand rate coefficients for different values of the parameters and different initial states.

4. Numerical Solution and Examples

The objective of this section is to study the numerical solution of the problem of determining an adaptive control strategy for the two-item inventory system subjected to different types of deterioration and unknown demand rate coefficients. To illustrate the solution procedure, let us consider simple examples in which the system parameters and initial states take different values. In these examples, the numerical solutions of the controlled two-item inventory system with unknown demand rate coefficients are presented. The numerical solution algorithm is based on the numerical integration of the system using the Runge-Kutta method.

Table 6

Parameter	θ_{11}	θ_{22}	a_{12}	a_{21}	\bar{x}_1	\bar{x}_2	k_1	k_2	m_1	m_2	θ_1	θ_2	d_1	d_2
Value	0.18	0.26	6.0	8.0	0.4	0.2	0.25	0.5	1.0	0.5	0.1	0.25	5.0	6.0

Table 7

Parameter	θ_{11}	θ_{22}	a_{12}	a_{21}	\bar{x}_1	\bar{x}_2	k_1	k_2	m_1	m_2	θ_1	θ_2	d_1	d_2
Value	0.15	0.3	15.0	10.0	5.0	0.6	5	3	9.0	7.0	0.3	0.2	10.0	15.0

Substituting from (3.3) into (3.1) and adding the system (3.4), we can get the adaptive control system as follows:

$$\begin{aligned}
 \dot{\xi}_1(t) &= -(d_1 + \theta_1)\xi_1(t) - \hat{a}_{12}\xi_1(t)\xi_2(t) - a_{12}\bar{x}_2\xi_1(t) - 2\theta_{11}\bar{x}_1\xi_1(t) \\
 &\quad - \hat{a}_{12}\bar{x}_1\xi_2(t) + a_{12}\xi_1(t)\xi_2(t) + a_{12}\bar{x}_1\xi_2(t) + \theta_{11}\xi_1(t)^2 - k_1\xi_1(t), \\
 \dot{\xi}_2(t) &= -(d_2 + \theta_2)\xi_2(t) - \hat{a}_{21}\xi_1(t)\xi_2(t) - a_{21}\bar{x}_1\xi_2(t) - 2\theta_{22}\bar{x}_2\xi_2(t) \\
 &\quad - \hat{a}_{21}\bar{x}_2\xi_1(t) + a_{21}\xi_1(t)\xi_2(t) + a_{21}\bar{x}_2\xi_1(t) + \theta_{22}\xi_2(t)^2 - k_2\xi_2(t), \\
 \dot{\hat{a}}_{12}(t) &= \xi_1(t)^2\xi_2(t) + \bar{x}_2\xi_1(t)^2 + \bar{x}_1\xi_1(t)\xi_2(t) - m_1(\hat{a}_{12}(t) - a_{12}), \\
 \dot{\hat{a}}_{21}(t) &= \xi_2(t)^2\xi_1(t) + \bar{x}_1\xi_2(t)^2 + \bar{x}_2\xi_1(t)\xi_2(t) - m_2(\hat{a}_{21}(t) - a_{21}).
 \end{aligned} \tag{4.1}$$

Clearly, this system is non-linear and its general solution is not available, so we will discuss its solution numerically. Next, we solve this system numerically for some particular values of the parameters and initial states.

4.1. Example 1

In this example, a numerical solution of the adaptive controlled system (4.1) is displayed graphically assuming constant demand rates. The following set of parameter values is used in Table 6 with the following initial values of perturbations of inventory levels and estimators of demand rate coefficients: $\xi_1(0) = 2$; $\xi_2(0) = 10$; $\hat{a}_{12}(0) = 10$; $\hat{a}_{21}(0) = 13$.

The numerical results are illustrated in Figure 3. We conclude that both of the perturbations of inventory levels and the estimators of demand rate coefficients tend to zero and their real values, respectively. This means that both of the inventory levels and demand rate coefficients asymptotically tend to their values at the steady state.

4.2. Example 2

In this example, a numerical solution of the adaptive controlled system (4.1) is displayed graphically when the demand rate is a linear function of the inventory level. The following set of parameter values is used in Table 7 with the following initial values of perturbations of inventory levels and estimators of demand rate coefficients: $\xi_1(0) = 5$; $\xi_2(0) = 15$; $\hat{a}_{12}(0) = 5$; $\hat{a}_{21}(0) = 8$.

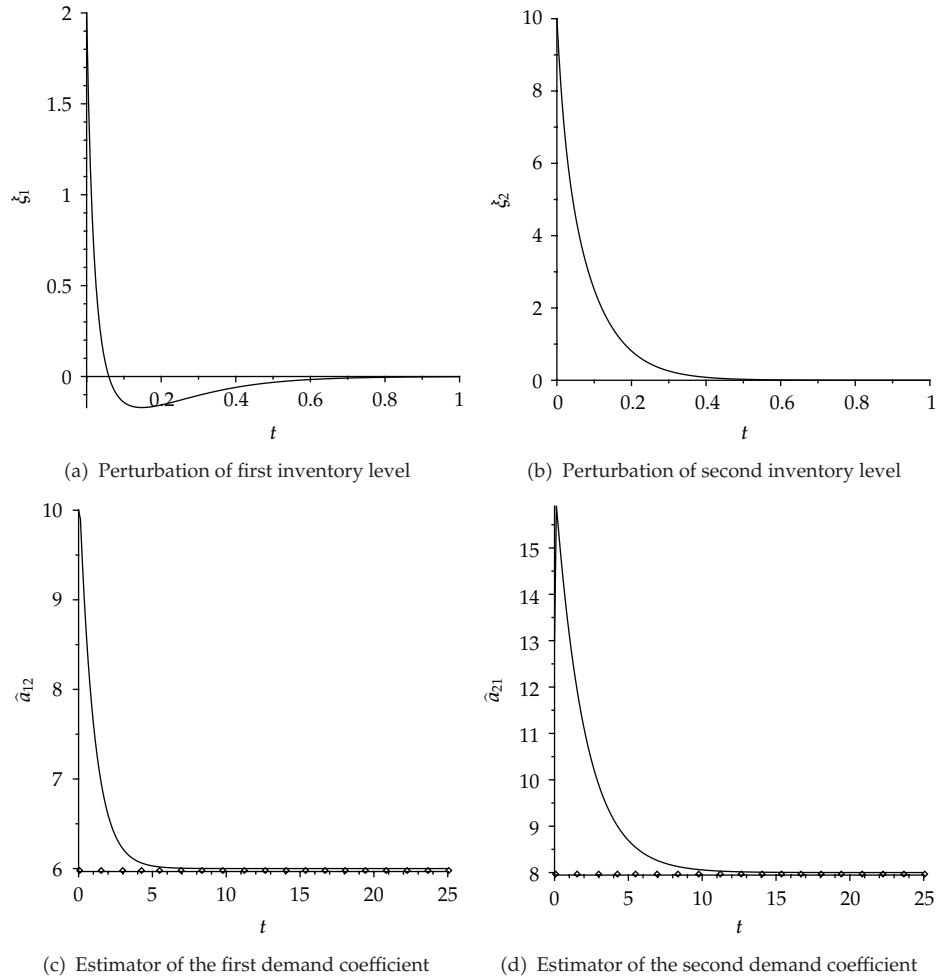


Figure 3: (a) and (b) are the perturbation of the first and second inventory levels about their inventory goal levels as the demand rate is a constant. (c) and (d) are the difference between dynamic estimators of the first and second demand rates and their real values.

The numerical results are illustrated in Figure 4. We conclude that both of the perturbations of inventory levels and the estimators of demand rate coefficients tend to zero and their real values, respectively. This means that both of the inventory levels and demand rate coefficients asymptotically tend to their values at the steady state. Also, we can easily notice that the estimators of the unknown demand rate coefficients are exponentially tend to the exact values.

4.3. Example 3

In this example, a numerical solution of the adaptive controlled system (4.1) is displayed graphically when the demand rate is an exponential function of time. The following set of parameter values is used in Table 8 with the following initial values of perturbations of

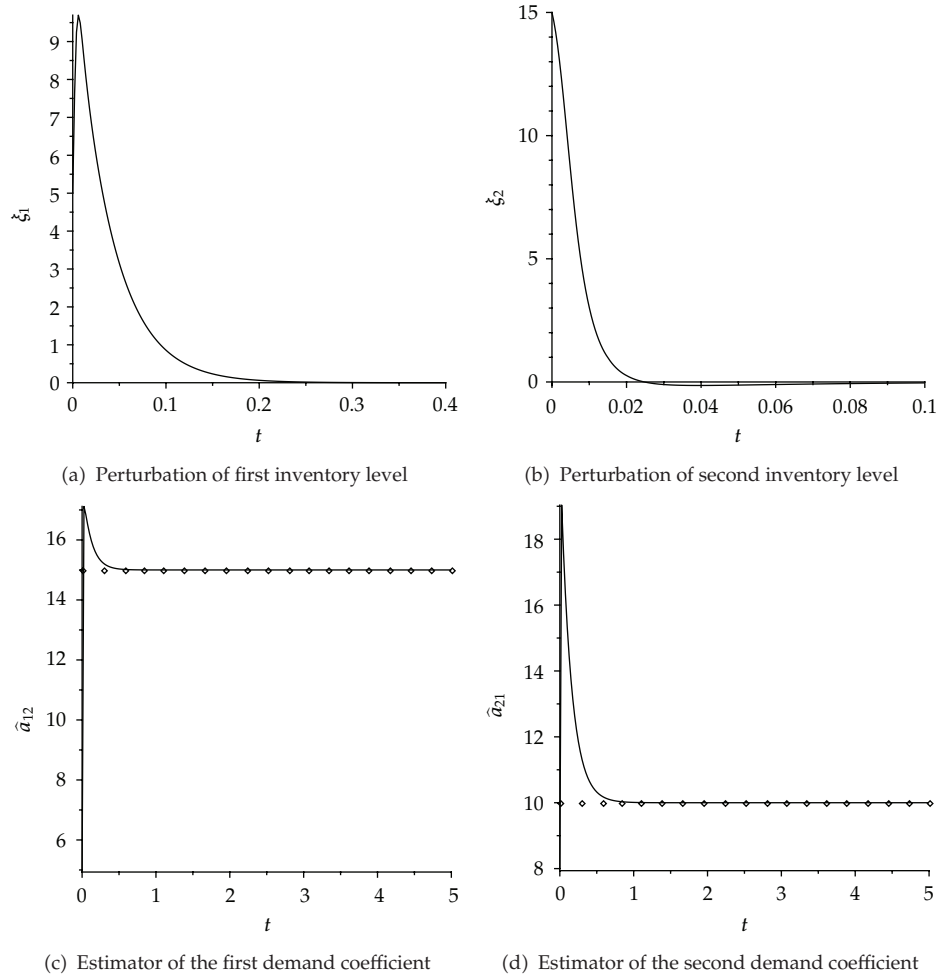


Figure 4: (a) and (b) are the perturbation of the first and second inventory levels about their inventory goal levels as the demand rate is a linear function of the inventory level. (c) and (d) are the difference between dynamic estimators of the first and second demand rates and their real values.

Table 8

Parameter	θ_{11}	θ_{22}	a_{12}	a_{21}	\bar{x}_1	\bar{x}_2	k_1	k_2	m_1	m_2	θ_1	θ_2	d_1	d_2
Value	0.08	0.06	20.0	15.0	0.2	1.0	0.02	0.015	0.1	0.2	0.15	0.1	20.0	30.0

inventory levels and estimators of demand rate coefficients: $\xi_1(0) = 25$; $\xi_2(0) = 0.2$; $\hat{a}_{12}(0) = 0.2$; $\hat{a}_{21}(0) = 10$.

The numerical results are illustrated in Figure 5. We conclude that both of the perturbations of inventory levels and the estimators of demand rate coefficients tend to zero and their real values, respectively. This means that both of the inventory levels and demand rate coefficients asymptotically tend to their values at the steady state.

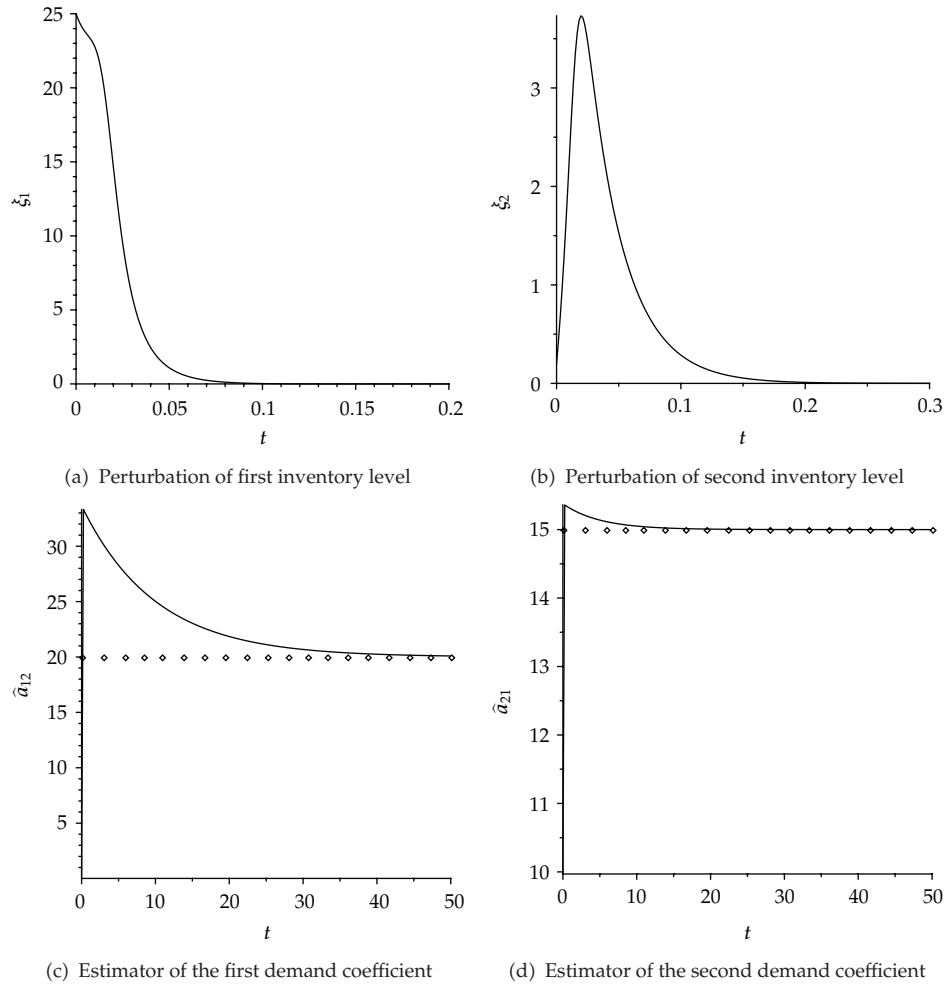


Figure 5: (a) and (b) are the perturbation of the first and second inventory levels about their inventory goal levels as the demand rate is an exponential function of the time. (c) and (d) are the difference between dynamic estimators of the first and second demand rates and their real values.

Table 9

Parameter	θ_{11}	θ_{22}	a_{12}	a_{21}	\bar{x}_1	\bar{x}_2	k_1	k_2	m_1	m_2	θ_1	θ_2	d_1	d_2
Value	0.08	0.06	20	15	2	1	20	15	25	30	0.15	0.1	20	30

4.4. Example 4

In this example, a numerical solution of the adaptive controlled system (4.1) is displayed graphically when the demand rate is an exponential function of time. The following set of parameter values is used in Table 9 with the following initial values of perturbations of inventory levels and estimators of demand rate coefficients: $\xi_1(0) = 1$; $\xi_2(0) = 2$; $\hat{a}_{12}(0) = 0.2$; $\hat{a}_{21}(0) = 10$.

The numerical results are illustrated in Figure 6. We conclude that both of the perturbations of inventory levels and the estimators of demand rate coefficients tend to zero

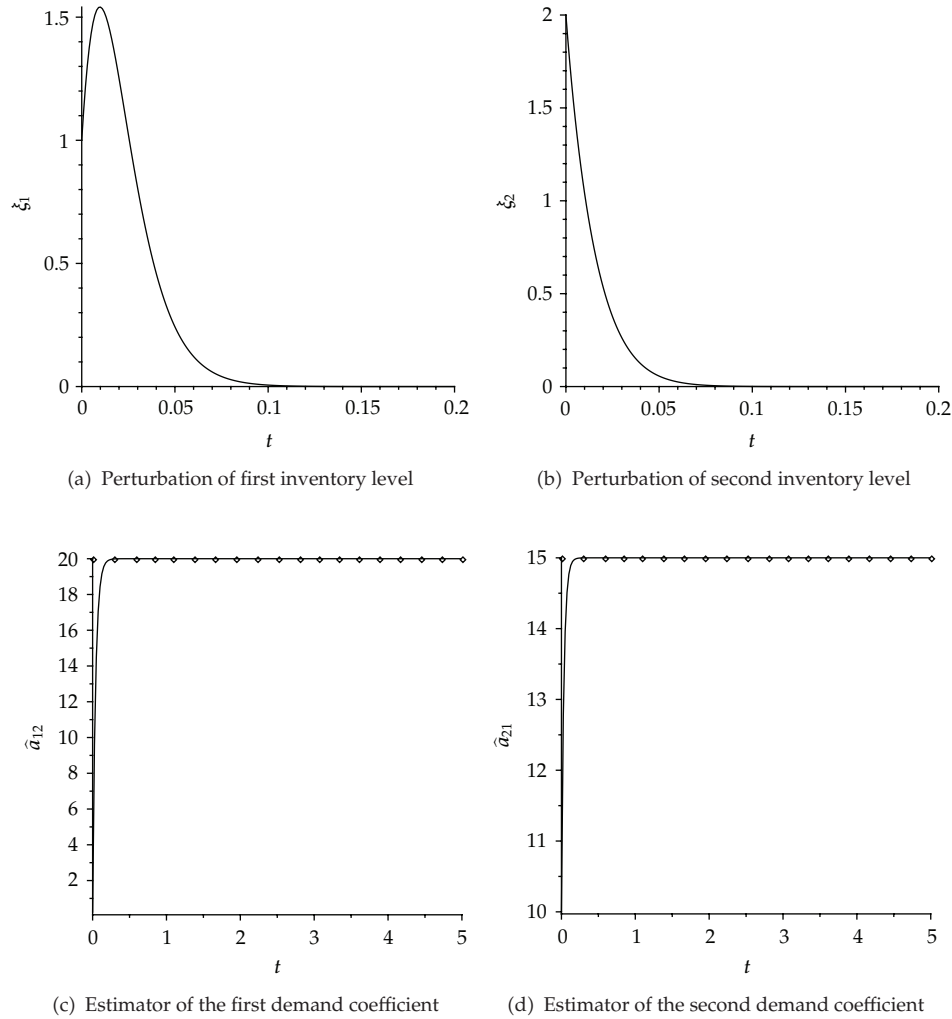


Figure 6: (a) and (b) are the perturbation of the first and second inventory levels about their inventory goal levels as the demand rate is an exponential function of the time. (c) and (d) are the difference between dynamic estimators of the first and second demand rates and their real values.

and their real values, respectively. This means that both of the inventory levels and demand rate coefficients asymptotically tend to their values at the steady state.

5. Conclusion

We have shown in this paper how to use an adaptive control approach to study the asymptotic stabilization of a two-item inventory model with unknown demand rate coefficients. A non-linear feedback approach is used to derive the continuous rate of supply. The Liapunov technique is used to prove the asymptotic stability of the adaptive controlled system. Also, the updating rules of the unknown demand rate coefficients have been derived

by using the conditions of the asymptotic stability of the perturbed system. Some numerical examples are presented to:

- (1) investigate the asymptotic behavior of both inventory levels and demand rate coefficient at the steady state;
- (2) estimate the unknown demand rate coefficients.

References

- [1] J. Aseltine, A. R. Mancini, and C. W. Sartune, "A survey of adaptive control systems," *IRE Transactions on Automatic Control*, vol. 3, no. 6, pp. 102–108, 1958.
- [2] W. Caldwell, "Control System with Automatic Response Adjustment," American Patent, 2517081, 1950.
- [3] I. D. Landau, *Adaptive Control: The Model Reference Approach*, Marcel Dekker, New York, NY, USA, 1979.
- [4] A. El-Gohary and R. Yassen, "Adaptive control and synchronization of a coupled dynamo system with uncertain parameters," *Chaos, Solitons & Fractals*, vol. 29, no. 5, pp. 1085–1094, 2006.
- [5] A. El-Gohary and A. S. Al-Ruzaiza, "Chaos and adaptive control in two prey, one predator system with nonlinear feedback," *Chaos, Solitons & Fractals*, vol. 34, no. 2, pp. 443–453, 2007.
- [6] L. Tadj, A. M. Sarhan, and A. El-Gohary, "Optimal control of an inventory system with ameliorating and deteriorating items," *Applied Sciences*, vol. 10, pp. 243–255, 2008.
- [7] A. Foul, L. Tadj, and R. Hedjar, "Adaptive control of inventory systems with unknown deterioration rate," *Journal of King Saud University—Science*, vol. 24, no. 3, pp. 215–220, 2012.
- [8] A. Alshamrani and A. El-Gohary, "Optimal control of two-item inventory model with different types of deterioration," *Economic Quality Control*, vol. 26, no. 2, pp. 201–213, 2011.
- [9] A. El-Gohary and A. El-Syed, "Optimal control of multi-item inventory models," *International Mathematical Forum*, vol. 3, no. 27, pp. 1295–1312, 2008.
- [10] K. S. Narendra and A. M. Annaswamy, *Stable Adaptive Systems*, Prentice Hall, Englewood Cliffs, NJ, USA, 1989.
- [11] D. K. Bhattacharya, "Production, manufacturing and logistics on multi-item inventory," *European Journal of Operational Research*, vol. 162, no. 3, pp. 786–791, 2005.
- [12] G. Hadley and J. Whitin, *Analysis of Inventory Systems*, Prentice Hall, Englewood Cliffs, NJ, USA, 1963.
- [13] J. D. Lenard and B. Roy, "Multi-item inventory control: a multicriteria view," *European Journal of Operational Research*, vol. 87, no. 3, pp. 685–692, 1995.
- [14] M. J. Rosenblatt, "Multi-item inventory system with budgetary constraint: a comparison between the lagrangian and the fixed cycle approach," *International Journal of Production Research*, vol. 19, no. 4, pp. 331–339, 1981.
- [15] A. M. Lyapunov, *The General Problem of Motion Stability*, vol. 17 of *Annals of Mathematics Studies*, Princeton University Press, Princeton, NJ, USA, 1892.

Research Article

Optimal Control of a Spatio-Temporal Model for Malaria: Synergy Treatment and Prevention

**Malicki Zorom,^{1,2} Pascal Zongo,^{2,3} Bruno Barbier,^{1,3}
and Blaise Somé²**

¹ Laboratoire d'Hydrologie et Ressources en Eau, Institut International d'Ingénierie de l'Eau et de l'Environnement (2iE), 01 Rue de la Science, BP 594, Ouagadougou, Burkina Faso

² Laboratoire CEREGMIA, Université des Antilles et de la Guyane, 2091 Route de Baduel, 97157 Pointe-à-Pitre, France

³ CIRAD, UMR G-EAU, 34398 Montpellier, France

Correspondence should be addressed to Bruno Barbier, bbarbier@cirad.fr

Received 29 January 2012; Revised 10 May 2012; Accepted 10 May 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 Malicki Zorom et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose a metapopulation model for malaria with two control variables, treatment and prevention, distributed between n different patches (localities). Malaria spreads between these localities through human travel. We used the theory of optimal control and applied a mathematical model for three connected patches. From previous studies with the same data, two patches were identified as reservoirs of malaria infection, namely, the patches that sustain malaria epidemic in the other patches. We argue that to reduce the number of infections and semi-immunes (i.e., asymptomatic carriers of parasites) in overall population, two considerations are needed, (a) For the reservoir patches, we need to apply both treatment and prevention to reduce the number of infections and to reduce the number of semi-immunes; neither the treatment nor prevention were specified at the beginning of the control application, except prevention that seems to be effective at the end. (b) For unreservoir patches, we should apply the treatment to reduce the number of infections, and the same strategy should be applied to semi-immune as in (a).

1. Introduction

Malaria is a mosquito-borne infection caused by protozoa of the genus *plasmodium*. Parasites are transmitted indirectly from humans to humans by the bite of infected female mosquitoes of the genus *Anopheles*. Malaria is a public health problem for tropical countries, which has negative impacts on development. The fight against mosquitoes passes through the draining of marshes or conversion to running water and elimination of stagnant water

especially around houses. These measures are difficult to apply where health facilities are inadequate [1].

Mathematical models coupled to microeconomic concepts can be applied to malaria control using the theory of optimal control [2–5]. The latter has already been used to discuss strategies to reduce or eradicate other diseases such as chronic myeloid leukemia [6], AIDS [7, 8], tuberculosis [9, 10], smoking [11], West Nile virus [12], and Chikungunya disease [13].

Human movements play a key role in the spatiotemporal spread of malaria [14, 15]. Models have already been proposed to study malaria spread, but neither control variables such as treatment and prevention nor human mobility were considered [16, 17]. In [18], a model with control variables (synergy prevention and treatment) was developed for malaria spread. Because human population was assumed to be motionless, their model is only applicable within a small geographical region. A model that takes into account human mobility was developed in [19] to analyze the impact of human migration within n geographical patches (localities) on the malaria spread.

Our model is not an extension of the model developed in [18] but rather the one developed in [19].

First, in [19], the authors have considered two infectious classes in the human population: infectious and semi-immune individuals (i.e., asymptomatic carriers). This consideration is very important because an experimental evidence showed that 60–90% of humans in endemic area are semi-immune [1, 16, 17]. Introduction of semi-immunes in their model presents the difficulty to introduce the control variable, mainly the treatment variable because treated individuals may become either susceptible or semi-immune depending on the type of the used drugs. In this paper, we will introduce a parameter which has a role of regulation denoted by θ to derive a biological meaningful model (see Figure 1).

Second, the model developed in [19] is of the metapopulation type that considers the explicit movement of humans between many patches. Because the mathematical analysis of their model has provided a methodology to identify the spatial reservoirs of malaria infection (i.e., the patches that sustain malaria epidemic in the other patches) based on the theory of the type *reproductive number*, our main objective in this paper is to extend their results by introducing control variables (treatment and prevention) within each patch. With these innovations, the simulations identify the best strategy of control and answer the following question: what control should be used when the patch is (or is not) a reservoir? This question is not trivial because the infectious individuals can migrate in all the other patches.

The paper is structured as follows: in Section 2, we summarize the main points of the metapopulation model from [19] by introducing prevention and treatment controls. Furthermore, we show that it is mathematically well posed. Section 3 includes the formulation of the objective function with the discount rate, and properties of optimal control existence follow its characterization. In Section 4, we present the results of simulations and discussion for three connected patches by migration according to the type of reservoir of infections. The last section includes the conclusion and perspectives.

2. Mathematical Modelling

2.1. Model Description

In this section, a metapopulation model with control variables (prevention and treatment) is developed. In the sequel, we use even and odd index to represent the human and mosquito

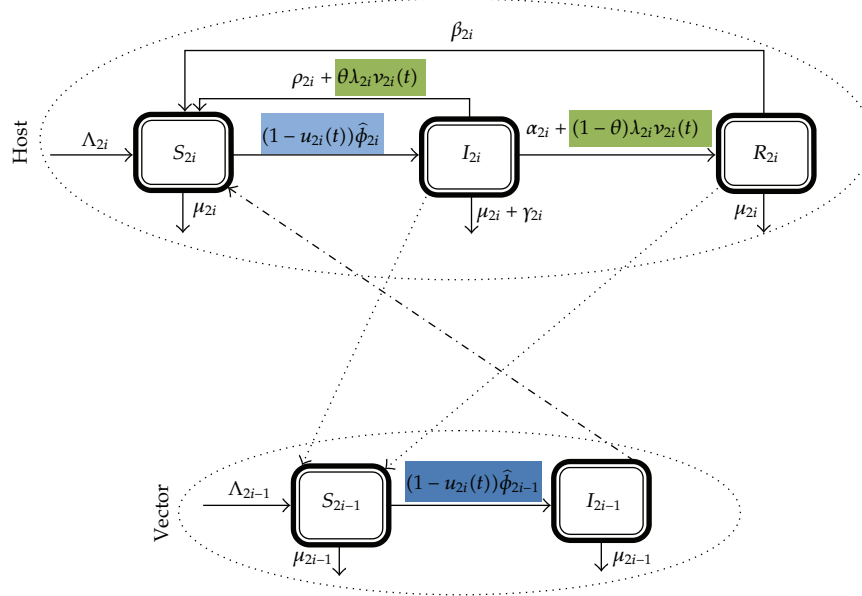


Figure 1: A conceptual mathematical model for malaria transmission involving human hosts and vector mosquitoes in each patch i , $i = 1, \dots, n$. The dotted arrow shows the direction of the transmission from humans to mosquitoes (through infectious humans to susceptible mosquitoes) or from mosquitoes to humans (through infectious mosquitoes to susceptible humans); u_{2i} and v_{2i} represent the prevention and treatment control over time, respectively. $\Phi_{2i} = (1 - u_{2i})\hat{\Phi}_{2i}$ and $\Phi_{2i-1} = (1 - u_{2i})\hat{\Phi}_{2i-1}$ represent the force of infection from mosquitoes to humans and from humans to mosquitoes, respectively, where $\hat{\Phi}_{2i}$ and $\hat{\Phi}_{2i-1}$ are defined in (2.2). The other parameters are described in Table 1.

variables, respectively, as in [1]. Within each small patch, the human hosts are split into three subclasses: susceptible S_{2i} , infectious I_{2i} , and semi-immune R_{2i} . $N_{2i}(t) = S_{2i}(t) + I_{2i}(t) + R_{2i}(t)$ denotes the total size of the human population in the patch i at time t . The mosquito population is split into two subclasses: susceptible S_{2i-1} and infectious I_{2i-1} in the patch i , $i = 1, \dots, n$. The total size of the mosquito population is denoted by $N_{2i-1}(t) = S_{2i-1}(t) + I_{2i-1}(t)$ at time t . $N_h(t) = \sum_{i=1}^n N_{2i}(t)$ and $N_v(t) = \sum_{i=1}^n N_{2i-1}(t)$ denote the total size of the human and mosquito population for the complete system, respectively, at any time t (See Figure 1). The model with control reads as follows: for all $i = 1, \dots, n$,

$$\frac{dS_{2i}}{dt} = \Lambda_{2i} + \beta_{2i}R_{2i} + \rho_{2i}I_{2i} - (\mu_{2i} + \Phi_{2i})S_{2i} + \theta\lambda_{2i}v_{2i}I_{2i} + \sum_{j=1}^n (m_{ij}^S S_{2j} - m_{ji}^S S_{2i}), \quad (2.1a)$$

$$\frac{dI_{2i}}{dt} = \Phi_{2i}S_{2i} - \epsilon_{2i}I_{2i} - \lambda_{2i}v_{2i}I_{2i} + \sum_{j=1}^n (m_{ij}^I I_{2j} - m_{ji}^I I_{2i}), \quad (2.1b)$$

$$\frac{dR_{2i}}{dt} = \alpha_{2i}I_{2i} - \delta_{2i}R_{2i} + (1 - \theta)\lambda_{2i}v_{2i}I_{2i} + \sum_{j=1}^n (m_{ij}^R R_{2j} - m_{ji}^R R_{2i}), \quad (2.1c)$$

$$\frac{dS_{2i-1}}{dt} = \Lambda_{2i-1} - \mu_{2i-1}S_{2i-1} - \Phi_{2i-1}S_{2i-1}, \quad (2.1d)$$

$$\frac{dI_{2i-1}}{dt} = \Phi_{2i-1}S_{2i-1} - \mu_{2i-1}I_{2i-1}, \quad (2.1e)$$

where $e_{2i} = \alpha_{2i} + \gamma_{2i} + \rho_{2i} + \mu_{2i}$ and $\delta_{2i} = \beta_{2i} + \mu_{2i}$. Initial conditions are assumed to satisfy $S_{2i}(0) > 0$, $S_{2i-1}(0) > 0$, $I_{2i}(0) \geq 0$, $R_{2i}(0) \geq 0$, and $I_{2i-1}(0) \geq 0$ for $i = 1, \dots, n$.

In the above model, Φ_{2i} and Φ_{2i-1} denote the force of infection from mosquitoes to humans and from humans to mosquitoes, respectively. Therefore, infection only involves vectors or hosts present in the patch; there is no between-patch infection. These forces of infection are modeled to take into account the prevention as in [18] as follows: $\Phi_{2i} = (1 - u_{2i})\hat{\Phi}_{2i}$ and $\Phi_{2i-1} = (1 - u_{2i})\hat{\Phi}_{2i-1}$ where

$$\begin{aligned}\hat{\Phi}_{2i} &= \frac{\tilde{a}_{2i-1}\tilde{a}_{2i}N_{2i-1}}{\tilde{a}_{2i-1}N_{2i-1} + \tilde{a}_{2i}N_{2i}}\sigma_{2i-1,2i}\frac{I_{2i-1}}{N_{2i-1}}, \\ \hat{\Phi}_{2i-1} &= \frac{\tilde{a}_{2i-1}\tilde{a}_{2i}N_{2i}}{\tilde{a}_{2i-1}N_{2i-1} + \tilde{a}_{2i}N_{2i}}\left(\sigma_{2i,2i-1}\frac{I_{2i}}{N_{2i}} + \hat{\sigma}_{2i,2i-1}\frac{R_{2i}}{N_{2i}}\right)\end{aligned}\quad (2.2)$$

are defined in [17] for one patch, \tilde{a}_{2i} , \tilde{a}_{2i-1} , $\sigma_{2i-1,2i}$, $\sigma_{2i,2i-1} \in \mathbb{R}_+$, $\hat{\sigma}_{2i,2i-1} \in \mathbb{R}_+^*$. The infection force Φ_{2i} and Φ_{2i-1} depends on the individuals within patch i and not in another patch $j \neq i$: infection that only involves those individuals (vectors or hosts) present in the patch is no between-patch infection.

u_{2i} is the prevention effort for humans which reduces the infection rate with a failure probability $1 - u_{2i}$ if prevention controls are introduced. The control function u_{2i} represents time-dependent efforts of prevention on human and practiced on a time interval $[0, T]$. Prevention could come from surveillance, treating vector-breeding ground, and reducing vector-host contacts. Note that when $u_{2i} = 0$, then Φ_{2i} and Φ_{2i-1} correspond to those used in [19].

$\lambda_{2i}v_{2i}$ is the per capita recovery rate of humans. $0 \leq \lambda_{2i} \leq 1$ is the proportion of effective treatment of humans. The control function v_{2i} represents the measure of the rate at which infected humans are cured by drugs or vaccination on a time interval $[0, T]$.

The difference between the effects of drugs is related to the fact that they act at different stages of the parasite-cell mutation in the human body. There are drugs that act against preerythrocytic stages, against the asexual blood stages, and against antigens of sexual stages that prevent fertilization in the stomach of the mosquito [20]. We thought that each drug has its own effect on the mode of healing. Therefore, there are drugs that favor the individuals who immunize quickly, while there are others that favor the total healing without being immunized. We then introduced the model of parameter θ , which regulates the process. θ is the probability that the treated infectious humans pass the sensitive compartment, and $1 - \theta$ is the probability to pass the semi-immune compartment. When using treatments that immunize the majority of patients, θ tends to 1, and these patients will go into the compartment of the semi-immune. Otherwise, θ tends to 0, and the patients will move into the susceptible compartment.

We also provide an insight into the major assumptions made in the original model in [19] as follows: disease-induced death rate of semi-immune was assumed to be negligible because this host acquires some immunity. Human mobility from one patch to another was considered, while immigration of mosquitoes was neglected because they can explore only a few kilometers during their lives. During the travel, humans do not change status. m_{ij}^π , $\pi = S, I, R$ denote the constant rate of travel of humans from an area j to an area i for all $i \neq j$ with $M^\pi = [m_{ij}^\pi]$, and $\pi = S, I, R$ denote the travel rate matrices. The matrices M^S were assumed to be irreducible and $m_{ii}^\pi = 0$, $\pi = S, I, R$; $i = 1, \dots, n$.

Table 1: Parameters for the model described in any patch i , $i = 1, \dots, n$.

Parameters and biological description		
Λ_{2i} : recruitment into the susceptible human		$\Lambda_{2i} > 0$
α_{2i} : rate of progression from the infectious human class to the semi-immune class		$\alpha_{2i} > 0$
ρ_{2i} : rate of progression from the infectious human class to the susceptible human class		$\rho_{2i} > 0$
β_{2i} : rate of progression from the semi-immune class to the susceptible human class		$\beta_{2i} > 0$
γ_{2i} : disease-induced death rate		$\gamma_{2i} \geq 0$
μ_{2i} : naturally induced death rate of the human population		$\mu_{2i} > 0$
μ_{2i-1} : naturally induced death rate of the mosquitoes		$\mu_{2i-1} > 0$
Λ_{2i-1} : recruitment into susceptible mosquitoes class		$\Lambda_{2i-1} > 0$
$\hat{\sigma}_{2i,2i-1}$: probability of transmission of the infection from a semi-immune human to a susceptible mosquito		$\hat{\sigma}_{2i,2i-1} > 0$
$\sigma_{2i-1,2i}$: probability of transmission of infection from an infectious mosquito to a susceptible human		$\sigma_{2i-1,2i} \in [0; 1]$
$\sigma_{2i,2i-1}$: probability of transmission of infection from an infectious human to a susceptible mosquito		$\sigma_{2i,2i-1} \in [0; 1]$
\tilde{a}_{2i} : maximum number of mosquito bites a human can receive per time unit		$\tilde{a}_{2i} \geq 0$
\tilde{a}_{2i-1} : number of time one mosquito would "want to" bite humans per time unit		$\tilde{a}_{2i-1} \geq 0$

Table 1 summarizes the parameters and their biological description that will be used in the metapopulation model.

By adding up (2.1a)–(2.1c) and (2.1d)–(2.1e), we obtain expressions for the total human and mosquito populations, respectively, in patch $i = 1, \dots, n$:

$$\begin{aligned} \frac{dN_{2i}}{dt} &= \Lambda_{2i} - \mu_{2i}N_{2i} - \gamma_{2i}I_{2i} + \sum_{\pi=S,I,R} \left(\sum_{j=1}^n m_{ij}^{\pi} \pi_{2j} - \sum_{j=1}^n m_{ji}^{\pi} \pi_{2i} \right), \\ \frac{dN_{2i-1}}{dt} &= \Lambda_{2i-1} - \mu_{2i-1}N_{2i-1}. \end{aligned} \quad (2.3)$$

Let $\Omega = (\mathbb{R}_+ \setminus \{0\})^{2n} \times \mathbb{R}_+^{3n}$, and denote the points in Ω by $(S, I)^T$, where $S = (S_2, S_1, \dots, S_{2n}, S_{2n-1})$ and $I = (I_2, R_2, I_1, \dots, I_{2n}, R_{2n}, I_{2n-1})$. Then we rewrite the system (2.1a)–(2.1e) in compact form

$$\begin{aligned} \frac{dS}{dt} &= \Psi_1(S, I), \\ \frac{dI}{dt} &= \Psi_2(S, I). \end{aligned} \quad (2.4)$$

For any initial condition $(S(0), I(0))$ in Ω , system (2.1a)–(2.1e) has a unique globally defined solution $(S(t), I(t))$ which remains in Ω . Moreover, the total human population, $N_h(t)$, and mosquitoes, $N_v(t)$, are bounded for all $t \geq 0$. This latter result was proved in [19].

2.2. Formulation of the Objective Functional

In this section, we formulate the optimal control problem with the following functional objective (cost):

$$J(u_2, v_2, \dots, u_{2n}, v_{2n}) = \sum_{i=1}^n \left[\int_0^T e^{-rt} \left(I_{2i} + R_{2i} + \frac{A_{2i}}{2} (u_{2i})^2 + \frac{B_{2i}}{2} (v_{2i})^2 \right) dt - \Upsilon^i(S_{2i}(T)) \right]. \quad (2.5)$$

$\sum_{i=1}^n I_{2i}$ and $\sum_{i=1}^n R_{2i}$ are the number of infected and semi-immune of n patches, respectively. The term $(A_{2i}/2)(u_{2i})^2 + (B_{2i}/2)(v_{2i})^2$ is the cost of prevention and treatment where $A_{2i}, B_{2i} > 0$ are the weight factor in the cost of control. $\Upsilon^i(S_{2i}(T))$ is the fitness of the susceptibles at the end of the process as a result of the treatment and prevention efforts for the patch $i = 1, \dots, n$. We also take the same form of the $\Upsilon^i(S_{2i}(T)) = W_{2i}^S S_{2i}(T)$, $W_{2i}^S \geq 0$ as in [18]. r is the discount rate. The discount rate is included to allow for long-term changes, thus giving greater emphasis to control in the short rather than the long term [21]. In the above formulation, one can note that the time $t = 0$ is the time when treatment and prevention are initiated, and the time $t = T$ is the time when treatment and prevention are stopped.

Additionally to the above assumptions, we assume that finance for treatment and prevention is not transferable through time, so that money which is not spent immediately cannot be saved for the future purchase of treatment and prevention.

Basically, we assume that the costs are proportional to the square of the corresponding control function due to some mathematics properties (positivity, convexity...).

2.3. Existence of an Optimal Control

The basic framework of this section is to characterize the optimal control and to prove the existence and uniqueness of the optimal control. We begin to simplify the writing by noting $(u_2, \dots, u_{2n}, v_2, \dots, v_{2n}) = (u, v)$ and $(u_2^*, \dots, u_{2n}^*, v_2^*, \dots, v_{2n}^*) = (u^*, v^*)$. Because the model is linear with respect to the control variables and bounded by a linear system with respect to the state variables, the conditions for the existence of an optimal control are satisfied. While applying the Fleming and Rishel theorem [22], the existence of the $2n$ -upplet optimal control can be obtained in our case.

Given

$$\mathcal{U} = \{(u, v), u, v, \text{ measurable } 0 \leq a_{2i} \leq u_{2i} \leq b_{2i} \leq 1, 0 \leq c_{2i} \leq v_{2i} \leq d_{2i} \leq 1\}, \quad (2.6)$$

therefore, one can state the following theorem.

Theorem 2.1. *Given the objective functional $J(u, v)$ defined by*

$$J(u, v) := \sum_{i=1}^n \left[\int_0^T e^{-rt} \left(I_{2i} + R_{2i} + \frac{A_{2i}}{2} (u_{2i})^2 + \frac{B_{2i}}{2} (v_{2i})^2 \right) dt - \Upsilon^i(S_{2i}(T)) \right], \quad (2.7)$$

for all $t \in [0; T]$ subject to the equations of system (2.1a)–(2.1e) with $S_{2i}(0) > 0$, $S_{2i-1}(0) > 0$, $R_{2i}(0) \geq 0$, $I_{2i}(0) \geq 0$, and $I_{2i-1}(0) \geq 0$ for $i = 1, \dots, n$, then there exists $2n$ -upplet optimal control (u^*, v^*) such that

$$J(u^*, v^*) = \min_{(u,v) \in \mathcal{U}} J(u, v), \quad (2.8)$$

when the following conditions are satisfied:

- (i) the class of all initial conditions with the $2n$ -upplet optimal control in the admissible control set and corresponding state variables is nonempty,
- (ii) the admissible control set \mathcal{U} is convex and closed,
- (iii) the right-hand side of the state system is bounded by a linear function in the state and control,
- (iv) the integrand of the objective functional is convex on \mathcal{U} and is bounded below by $\sum_{i=1}^n (h_{2i}(|u_{2i}|^2 + |v_{2i}|^2)^{\varrho_{2i}/2} - k_{2i})$, where $h_{2i}, k_{2i} > 0$, and $\varrho_{2i} > 1$,
- (v) the function $\sum_{i=1}^n Y^i(S_{2i}(T))$ is continuous with respect to the variable S_{2i} .

Proof.

- (i) It is obtained by definition.
- (ii) By definition, the admissible control set \mathcal{U} is convex and closed.
- (iii) The right-hand side of the state system (2.1a)–(2.1e) is bounded by a linear function in the state (refer to Theorem 1 of [19]). Our state system is bilinear in the control variable.
- (iv) To show that the integrand of the objective functional is convex on \mathcal{U} , we must prove that

$$F_{2i} \left(t, I_{2i}, R_{2i}, \sum_{j=1}^2 \eta_{2j} X_{2j} \right) \leq \sum_{j=1}^2 \eta_{2j} F_{2i}(t, I_{2i}, R_{2i}, X_{2j}), \quad (2.9)$$

where $\sum_{j=1}^2 \eta_{2j} = 1$, $X_{2j} = (u_{2j}, v_{2j})$ and

$$\begin{aligned} J(u, v) + \sum_{i=1}^n S_{2i}(T) &= \sum_{i=1}^n \int_0^T e^{-rt} \left(I_{2i} + R_{2i} + \frac{A_{2i}}{2} (u_{2i})^2 + \frac{B_{2i}}{2} (v_{2i})^2 \right) dt \\ &= \sum_{i=1}^n \int_0^T F_{2i}(t, I_{2i}, R_{2i}, X_{2i}), \end{aligned} \quad (2.10)$$

where

$$F_{2i}(t, I_{2i}, R_{2i}, X_{2i}) = e^{-rt} \left(I_{2i} + R_{2i} + \frac{A_i}{2} (u_{2i})^2 + \frac{B_{2i}}{2} (v_{2i})^2 \right), \quad (2.11)$$

$$\begin{aligned} F_{2i} \left(t, I_{2i}, R_{2i}, \sum_{j=1}^2 \eta_{2j} X_{2j} \right) &= I_{2i} + R_{2i} + \frac{A_{2i}}{2} \left(\sum_{j=1}^2 \eta_{2j} u_{2j} \right)^2 + \frac{B_{2i}}{2} \left(\sum_{j=1}^2 \eta_{2j} v_{2j} \right)^2 \\ &\leq I_{2i} + R_{2i} + \frac{A_{2i}}{2} \sum_{j=1}^2 \eta_{2j} (u_{2j})^2 + \frac{B_{2i}}{2} \sum_{j=1}^2 \eta_{2j} (v_{2j})^2 \\ &\leq \sum_{j=1}^2 \eta_{2j} \left(I_{2i} + R_{2i} + \frac{A_{2i}}{2} (u_{2j})^2 + \frac{B_{2i}}{2} (v_{2j})^2 \right) \\ &= \sum_{j=1}^2 \eta_{2j} F_{2i}(t, I_{2i}, R_{2i}, X_{2j}). \end{aligned} \quad (2.12)$$

Since the sum of convex functions in the domain convex is convex, then there exists $h_{2i}, k_{2i}, q_{2i} > 1$ satisfying

$$e^{-rt} \left(I_{2i} + R_{2i} + \frac{A_{2i}}{2} (u_{2i})^2 + \frac{B_{2i}}{2} (v_{2i})^2 \right) \geq \left(h_{2i} (|u_{2i}|^2 + |v_{2i}|^2)^{q_{2i}/2} - k_{2i} \right), \quad (2.13)$$

because the state variable is bounded. So summing member to member, one obtains the result.

(v) The function $Y^i(S_{2i}(T))$ is continuous so that $\sum_{i=1}^n Y^i(S_{2i}(T))$ is also continuous. □

2.4. Characterization of the 2n-Upplet Optimal Control

Since there exists 2n-upplet optimal control for minimizing the functional, (2.7), subject to system (2.1a)–(2.1e), we derive the necessary conditions on the optimal control. We discuss the theorem that relates to the characterization of the optimal control. In order to derive the necessary conditions for this optimal control, we use Pontryagin's maximum principle [23]. The Lagrangian, sometimes called the Hamiltonian, augmented with penalty terms for control constraints is defined as

$$\begin{aligned} L &= \sum_{i=1}^n \left(I_{2i} + R_{2i} + \frac{A_{2i}}{2} (u_{2i})^2 + \frac{B_{2i}}{2} (v_{2i})^2 \right) \\ &\quad + \sum_{i=1}^n \lambda_{S_{2i}} \left(\Lambda_{2i} + \beta_{2i} R_{2i} + \rho_{2i} I_{2i} - (\mu_{2i} + \Phi_{2i}) S_{2i} + \theta \lambda_{2i} v_{2i} I_{2i} + \sum_{j=1}^n (m_{ij}^S S_{2j} - m_{ji}^S S_{2i}) \right) \end{aligned}$$

$$\begin{aligned}
& + \sum_{i=1}^n \lambda_{I_{2i}} \left(\Phi_{2i} S_{2i} - \epsilon_{2i} I_{2i} - \lambda_{2i} v_{2i} I_{2i} + \sum_{j=1}^n (m_{ij}^I I_{2j} - m_{ji}^I I_{2i}) \right) \\
& + \sum_{i=1}^n \lambda_{R_{2i}} \left(\alpha_{2i} I_{2i} - \delta_{2i} R_{2i} + (1 - \theta) \lambda_{2i} v_{2i} I_{2i} + \sum_{j=1}^n (m_{ij}^R R_{2j} - m_{ji}^R R_{2i}) \right) \\
& + \sum_{i=1}^n \lambda_{S_{2i-1}} (\Lambda_{2i-1} - \mu_{2i-1} S_{2i-1} - \Phi_{2i-1} S_{2i-1}) + \sum_{i=1}^n \lambda_{I_{2i-1}} (\Phi_{2i-1} S_{2i-1} - \mu_{2i-1} I_{2i-1}) \\
& - \sum_{i=1}^n \omega_{2i} (u_{2i} - a_{2i}) - \sum_{i=1}^n \varpi_{2i} (b_{2i} - u_{2i}) - \sum_{i=1}^n \zeta_{2i} (v_{2i} - c_{2i}) - \sum_{i=1}^n \xi_{2i} (d_{2i} - v_{2i}),
\end{aligned} \tag{2.14}$$

where λ_{π} , $\pi = S_{2i}, I_{2i}, R_{2i}, S_{2i-1}, I_{2i-1}$ is the costate variable to the state variable (S, I) , respectively, for patch $i = 1, \dots, n$. We can interpret $\lambda_{\pi}(t)$ as the marginal value or shadow price of the last unit of $S_{2i}, S_{2i-1}, I_{2i}, I_{2i-1}$, and R_{2i} was evaluated at time t [3]. For example, $\lambda_{S_{2i}}$ is the increase in welfare if the number of susceptible is exogenously increased at time t . λ_{π} can be negative. The parameters $\omega_{2i}, \varpi_{2i}, \zeta_{2i}, \xi_{2i}$ with $i = 1, \dots, n$ are the penalty multipliers satisfying these conditions:

$$\begin{aligned}
\omega_{2i} &\geq 0, \quad u_{2i} - a_{2i} \geq 0, \quad \omega_{2i}(u_{2i} - a_{2i}) = 0, \quad i = 1, \dots, n, \\
\varpi_{2i} &\geq 0, \quad b_{2i} - u_{2i} \geq 0, \quad \varpi_{2i}(b_{2i} - u_{2i}) = 0, \quad i = 1, \dots, n, \\
\zeta_{2i} &\geq 0, \quad v_{2i} - c_{2i} \geq 0, \quad \zeta_{2i}(v_{2i} - c_{2i}) = 0, \quad i = 1, \dots, n, \\
\xi_{2i} &\geq 0, \quad d_{2i} - v_{2i} \geq 0, \quad \xi_{2i}(d_{2i} - v_{2i}) = 0, \quad i = 1, \dots, n.
\end{aligned} \tag{2.15}$$

The supplementary condition at the first and second line of the system (2.15) realized at optimal control u_{2i}^* and the last two lines of this system is realized at the optimal control v_{2i}^* .

Theorem 2.2. *Given $2n$ -upplet optimal controls (u^*, v^*) and solutions (S, I) of the corresponding state system (2.1a)–(2.1e), there exists adjoint variables λ_{π} , with $\pi = S_{2i}, S_{2i-1}, R_{2i}, I_{2i}, I_{2i-1}$ where $i = 1, \dots, n$ satisfying the following canonical equations:*

$$\begin{aligned}
\frac{d\lambda_{S_{2i}}}{dt} &= r\lambda_{S_{2i}} - \frac{\partial L}{\partial S_{2i}} = r\lambda_{S_{2i}} + \lambda_{S_{2i}} \left(\mu_{2i} + \Phi_{2i} \left(1 - \frac{\tilde{a}_{2i} S_{2i}}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} \right) + \sum_{j=1}^n m_{ji}^S \right) \\
&\quad - \lambda_{I_{2i}} \Phi_{2i} \left(1 - \frac{\tilde{a}_{2i} S_{2i}}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} \right) + (\lambda_{I_{2i-1}} - \lambda_{S_{2i-1}}) \frac{\tilde{a}_{2i} S_{2i-1} \Phi_{2i-1}}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}}, \\
\frac{d\lambda_{I_{2i}}}{dt} &= r\lambda_{I_{2i}} - \frac{\partial L}{\partial I_{2i}} = r\lambda_{I_{2i}} - 1 - \lambda_{S_{2i}} \left(\rho_{2i} + \frac{\tilde{a}_{2i} S_{2i} \Phi_{2i}}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} + \theta \lambda_{2i} v_{2i} \right) \\
&\quad + \lambda_{I_{2i}} \left(\frac{\tilde{a}_{2i} S_{2i} \Phi_{2i}}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} + \epsilon_{2i} + \lambda_{2i} v_{2i} + \sum_{j=1}^n m_{ji}^I \right) - \lambda_{R_{2i}} (\alpha_{2i} + (1 - \theta) \alpha_{2i} v_{2i}) \\
&\quad + (\lambda_{S_{2i-1}} - \lambda_{I_{2i-1}}) \frac{\tilde{a}_{2i} (\tilde{a}_{2i-1} (1 - u_{2i}) \sigma_{2i,2i-1} - \Phi_{2i-1}) S_{2i-1}}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}},
\end{aligned}$$

$$\begin{aligned}
\frac{d\lambda_{R_{2i}}}{dt} &= r\lambda_{R_{2i}} - \frac{\partial L}{\partial R_{2i}} = r\lambda_{R_{2i}} - 1 - \lambda_{S_{2i}} \left(\beta_{2i} + \frac{\tilde{a}_{2i}S_{2i}\Phi_{2i}}{\tilde{a}_{2i-1}N_{2i-1} + \tilde{a}_{2i}N_{2i}} \right) + \lambda_{I_{2i}} \frac{\tilde{a}_{2i}S_{2i}\Phi_{2i}}{\tilde{a}_{2i-1}N_{2i-1} + \tilde{a}_{2i}N_{2i}} \\
&\quad + \lambda_{R_{2i}} \left(\delta_{2i} + \sum_{j=1}^n m_{ji}^R \right) + (\lambda_{S_{2i-1}} - \lambda_{I_{2i-1}}) \frac{\tilde{a}_{2i}(\tilde{a}_{2i-1}(1-u_{2i})\hat{\sigma}_{2i,2i-1} - \Phi_{2i-1})S_{2i-1}}{\tilde{a}_{2i-1}N_{2i-1} + \tilde{a}_{2i}N_{2i}}, \\
\frac{d\lambda_{S_{2i-1}}}{dt} &= r\lambda_{S_{2i-1}} - \frac{\partial L}{\partial S_{2i-1}} = r\lambda_{S_{2i-1}} + (\lambda_{I_{2i}} - \lambda_{S_{2i}}) \frac{\tilde{a}_{2i-1}S_{2i}\Phi_{2i}}{\tilde{a}_{2i-1}N_{2i-1} + \tilde{a}_{2i}N_{2i}} \\
&\quad + \lambda_{S_{2i-1}}\mu_{2i-1} + (\lambda_{I_{2i-1}} - \lambda_{S_{2i-1}})\Phi_{2i-1} \left(\frac{\tilde{a}_{2i-1}S_{2i-1}}{\tilde{a}_{2i-1}N_{2i-1} + \tilde{a}_{2i}N_{2i}} - 1 \right), \\
\frac{d\lambda_{I_{2i-1}}}{dt} &= r\lambda_{I_{2i-1}} - \frac{\partial L}{\partial I_{2i-1}} = r\lambda_{I_{2i-1}} + (\lambda_{S_{2i}} - \lambda_{I_{2i}}) \frac{\tilde{a}_{2i-1}(\tilde{a}_{2i}(1-u_{2i})\sigma_{2i-1,2i} - \Phi_{2i})S_{2i}}{\tilde{a}_{2i-1}N_{2i-1} + \tilde{a}_{2i}N_{2i}} \\
&\quad + \lambda_{I_{2i-1}}\mu_{2i-1} - (\lambda_{S_{2i-1}} - \lambda_{I_{2i-1}}) \frac{\tilde{a}_{2i-1}S_{2i-1}\Phi_{2i-1}}{\tilde{a}_{2i-1}N_{2i-1} + \tilde{a}_{2i}N_{2i}},
\end{aligned} \tag{2.16}$$

with the transversality conditions (terminal conditions):

$$\lambda_{S_{2i}}(T) = \left. \frac{\partial Y^i}{\partial S_{2i}} \right|_{t=T}, \quad \lambda_{\pi}(T) = 0, \quad i = 1, \dots, n \quad \text{for } \pi = I_{2i}, I_{2i-1}, R_{2i}, S_{2i-1}. \tag{2.17}$$

Furthermore, the following characterization of optimal control holds:

$$\begin{aligned}
u_{2i}^* &= \max \left(a_{2i}, \min \left(b_{2i}, \frac{\tilde{a}_{2i-1}\tilde{a}_{2i}}{A_{2i}} \left[\frac{\lambda_{S_{I_{2i}}}\sigma_{2i-1,2i}I_{2i-1}S_{2i}}{\tilde{a}_{2i-1}N_{2i-1} + \tilde{a}_{2i}N_{2i}} + \frac{\lambda_{S_{I_{2i-1}}}(\sigma_{2i,2i-1}I_{2i} + \hat{\sigma}_{2i,2i-1}R_{2i})S_{2i-1}}{\tilde{a}_{2i-1}N_{2i-1} + \tilde{a}_{2i}N_{2i}} \right] \right) \right), \\
v_{2i}^* &= \max \left(c_{2i}, \min \left(d_{2i}, \frac{-\lambda_{2i}I_{2i}(\theta\lambda_{S_{2i}} - \lambda_{I_{2i}} + (1-\theta)\lambda_{R_{2i}})}{B_{2i}} \right) \right),
\end{aligned} \tag{2.18}$$

where $\lambda_{S_j} - \lambda_{I_j} = -\lambda_{S_{I_j}}$.

Proof. The adjoint equations and transversality conditions are standard results from Pontryagin's maximum principle. Also, solutions to the adjoint system exist and are bounded. To determine the interior optimum of our Lagrangian, we take the partial derivatives of Lagrangian L with respect to u_{2i}^* and v_{2i}^* and set it equal to zero:

$$\begin{aligned}
\frac{\partial L}{\partial u_{2i}} &= -\frac{\tilde{a}_{2i-1}\tilde{a}_{2i}\lambda_{S_{I_{2i}}}\sigma_{2i-1,2i}I_{2i-1}S_{2i}}{\tilde{a}_{2i-1}N_{2i-1} + \tilde{a}_{2i}N_{2i}} - \frac{\tilde{a}_{2i-1}\tilde{a}_{2i}S_{2i-1}\lambda_{S_{I_{2i-1}}}(\sigma_{2i,2i-1}I_{2i} + \hat{\sigma}_{2i,2i-1}R_{2i})}{\tilde{a}_{2i-1}N_{2i-1} + \tilde{a}_{2i}N_{2i}} \\
&\quad + \varpi_{2i} - \omega_{2i} + A_{2i}u_{2i} = 0, \\
\frac{\partial L}{\partial v_{2i}} &= \lambda_{2i}I_{2i}(\theta\lambda_{S_{2i}} - \lambda_{I_{2i}} + (1-\theta)\lambda_{R_{2i}}) + \xi_{2i} - \zeta_{2i} + B_{2i}v_{2i} = 0.
\end{aligned} \tag{2.19}$$

Solving for optimal control, we have

$$u_{2i}^* = \frac{1}{A_{2i}} \left[\frac{\tilde{a}_{2i-1} \tilde{a}_{2i} \lambda_{SI_{2i}} \sigma_{2i-1,2i} I_{2i-1} S_{2i}}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} + \frac{\tilde{a}_{2i-1} \tilde{a}_{2i} S_{2i-1} \lambda_{SI_{2i-1}} (\sigma_{2i,2i-1} I_{2i} + \hat{\sigma}_{2i,2i-1} R_{2i})}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} - \varpi_{2i} + \omega_{2i} \right],$$

$$v_{2i}^* = \frac{1}{B_{2i}} [-\lambda_{2i} I_{2i} (\theta \lambda_{S_{2i}} - \lambda_{I_{2i}} + (1 - \theta) \lambda_{R_{2i}}) - \xi_{2i} + \zeta_{2i}]. \quad (2.20)$$

To determine an explicit expression for the optimal control without the penalty multipliers ω_{2i} , ϖ_{2i} , ξ_{2i} , ζ_{2i} , a standard optimality technique is used. We consider the following cases to discuss the control: case of the prevention or case of the treatment.

(i) Case of prevention:

(1) on the set

$$\{ta_{2i} < u_{2i}^* < b_{2i}, i = 1, \dots, n\}, \quad (2.21)$$

we have $\omega_{2i} = \varpi_{2i} = 0$. Hence, the optimal control is

$$u_{2i}^* = \frac{1}{A_{2i}} \left[\frac{\tilde{a}_{2i-1} \tilde{a}_{2i} \lambda_{SI_{2i}} \sigma_{2i-1,2i} I_{2i-1} S_{2i}}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} + \frac{\tilde{a}_{2i-1} \tilde{a}_{2i} S_{2i-1} \lambda_{SI_{2i-1}} (\sigma_{2i,2i-1} I_{2i} + \hat{\sigma}_{2i,2i-1} R_{2i})}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} \right], \quad (2.22)$$

(2) on the set

$$\{tu_{2i}^* = b_{2i}, i = 1, \dots, n\}, \quad (2.23)$$

we have $\omega_{2i}(t) = 0$. Hence,

$$b_{2i} = u_{2i}^* = \frac{1}{A_{2i}} \left[\frac{\tilde{a}_{2i-1} \tilde{a}_{2i} \lambda_{SI_{2i}} \sigma_{2i-1,2i} I_{2i-1} S_{2i}}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} + \frac{\tilde{a}_{2i-1} \tilde{a}_{2i} S_{2i-1} \lambda_{SI_{2i-1}} (\sigma_{2i,2i-1} I_{2i} + \hat{\sigma}_{2i,2i-1} R_{2i})}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} - \varpi_{2i} \right]. \quad (2.24)$$

This implies that

$$\frac{1}{A_{2i}} \left[\frac{\tilde{a}_{2i-1} \tilde{a}_{2i} \lambda_{SI_{2i}} \sigma_{2i-1,2i} I_{2i-1} S_{2i}}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} + \frac{\tilde{a}_{2i-1} \tilde{a}_{2i} S_{2i-1} \lambda_{SI_{2i-1}} (\sigma_{2i,2i-1} I_{2i} + \hat{\sigma}_{2i,2i-1} R_{2i})}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} \right] \geq b_{2i}, \quad (2.25)$$

since $\varpi_{2i}(t) \geq 0$,

(3) on the set

$$\{tu_{2i}^* = a_{2i}, i = 1, \dots, n\}, \quad (2.26)$$

we have $\varpi_{2i}(t) = 0$. Hence,

$$u_{2i}^* = \frac{1}{A_{2i}} \left[\frac{\tilde{a}_{2i-1} \tilde{a}_{2i} \lambda_{S_{I_{2i}}} \sigma_{2i-1,2i} I_{2i-1} S_{2i}}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} + \frac{\tilde{a}_{2i-1} \tilde{a}_{2i} S_{2i-1} \lambda_{S_{I_{2i-1}}} (\sigma_{2i,2i-1} I_{2i} + \hat{\sigma}_{2i,2i-1} R_{2i})}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} + \omega_{2i} \right]. \quad (2.27)$$

This implies that

$$\frac{1}{A_{2i}} \left[\frac{\tilde{a}_{2i-1} \tilde{a}_{2i} \lambda_{S_{I_{2i}}} \sigma_{2i-1,2i} I_{2i-1} S_{2i}}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} + \frac{\tilde{a}_{2i-1} \tilde{a}_{2i} S_{2i-1} \lambda_{S_{I_{2i-1}}} (\sigma_{2i,2i-1} I_{2i} + \hat{\sigma}_{2i,2i-1} R_{2i})}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} \right] \leq a_{2i}. \quad (2.28)$$

Combining these cases, the optimal control u_{2i}^* for $i = 1, \dots, n$ is characterized as

$$u_{2i}^* = \max \left(a_{2i}, \min \left(b_{2i}, \frac{\tilde{a}_{2i-1} \tilde{a}_{2i}}{A_{2i}} \left[\frac{\lambda_{S_{I_{2i}}} \sigma_{2i-1,2i} I_{2i-1} S_{2i}}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} + \frac{\lambda_{S_{I_{2i-1}}} (\sigma_{2i,2i-1} I_{2i} + \hat{\sigma}_{2i,2i-1} R_{2i}) S_{2i-1}}{\tilde{a}_{2i-1} N_{2i-1} + \tilde{a}_{2i} N_{2i}} \right] \right) \right). \quad (2.29)$$

(ii) Case of treatment:

using similar arguments as in the case of prevention, we also obtain the second optimal v_{2i}^* with $i = 1, \dots, n$ control function is characterized by

$$v_{2i}^* = \max \left(c_{2i}, \min \left(d_{2i}, \frac{-\lambda_{2i} I_{2i} (\theta \lambda_{S_{2i}} - \lambda_{I_{2i}} + (1 - \theta) \lambda_{R_{2i}})}{B_{2i}} \right) \right). \quad (2.30) \quad \square$$

3. Numerical Results and Discussion

3.1. Parameters

We fix the probability for treatment of infectious humans for patches 2 and 3 at $\theta = 0.5$. Also we take $\lambda_{2j} = 0.5$, the weights of prevention and treatment $A_{2j} = B_{2j} = 50$, and the bounds of all control $a_{2j} = c_{2j} = 0$, $b_{2j} = d_{2j} = 1$ with $j = 2, 3$. We fix the coefficient of fitness $W_{2i}^S = 1$.

We take a very small discount rate $r = 0.0001$ because the daily discounting of the cost decreases very slowly. The other parameters of the model were obtained from [1] as well as the following value of the migration matrix: data for migration matrix for the semi-immune, M^R ,

$$M^R = \begin{bmatrix} 0 & 0.7 \times 10^{-1} & 0.8 \times 10^{-1} \\ 0.1 \times 10^{-1} & 0 & 0.1 \times 10^{-4} \\ 0.2 \times 10^{-1} & 0.1 \times 10^{-4} & 0 \end{bmatrix}, \quad (3.1)$$

M^S , data for migration matrix for susceptible, and M^I , data for migration matrix for the infectious

$$M^S = M^I = \begin{bmatrix} 0 & 0.7 \times 10^{-3} & 0.8 \times 10^{-3} \\ 0.1 \times 10^{-3} & 0 & 0.1 \times 10^{-6} \\ 0.2 \times 10^{-3} & 0.1 \times 10^{-6} & 0 \end{bmatrix}. \quad (3.2)$$

3.2. Implementation

To solve our problem of optimal control, we used the program MATLAB dynamic optimisation code (*DYNOPT*), which is a set of MATLAB functions for the determination of optimal control trajectory by describing the process, the cost to be minimized, subject to equality and inequality constraints, and using orthogonal collocation on the finite elements method [24]. For more information about this algorithm, we can see the user's guide in [24].

We implemented the model with the initial condition: $S_2(0) = 15000$; $S_4(0) = 50$; $S_6(0) = 1000$; $I_2(0) = 1000$; $I_4(0) = 50$; $I_6(0) = 100$; $R_2(0) = 100$; $R_4(0) = 250$; $R_6(0) = 10$; $S_1(0) = 5000$; $S_3(0) = 8000$; $S_5(0) = 5000$; $I_1(0) = 50$; $I_3(0) = 6000$; $I_5(0) = 4000$.

The weight assigned to the controls is much higher than the weight assigned to the state variables because the two functions are not expressed in the same scale. The controls are expressed in terms of cost, while infections and semi-immune are expressed in term of number of individuals. We chose the time at $T = 300$ days for our simulation.

3.3. Results and Discussions

3.3.1. Basic Reproductive Number and Reservoir of Infection

The basic reproduction number generally denoted by \mathcal{R}_0 is the expected number of secondary cases produced by a typical infective individual introduced into a completely susceptible population, in the absence of any control measure [25, 26]. Using the data on Table 2 which were compiled in [1] without control variable, \mathcal{R}_0 was equal to 3.864. Therefore, there is a persistence of the disease in the whole population (patches 1, 2, and 3). In [1], it was shown that only patches 2 and 3 constitute a reservoir of infection. Indeed, a subgroup of patches is said to be a reservoir when only targeting a control on the reservoir is sufficient to eliminate the malaria in the whole population (all the three patches). As such the patch 1 cannot sustain an epidemic by itself.

3.3.2. Evolution over Time of the Optimal Control in the Three Patches

We seek the optimal solution by minimizing the number of infectious hosts and semi-immune, in all patches by considering four cases: the first case where we seek the optimal solution when we consider simultaneously prevention and treatment in the two patches (see Figure 2(a)), the second case where we seek the optimal solution only with prevention without treatment in two patches (see Figure 2(b)), the third case where we seek the optimal solution only with the treatment without prevention in two patches (see Figure 2(c)), and finally the fourth case where no strategy of prevention and treatment is applied. Figure 2 shows a strong preventive action early in the process of elimination of the disease and a high processing action at the end of the process. Between these two strategies, prevention and

Table 2: Value compiled in [1]: patches 2 and 3 correspond to rural areas, while patch 1 corresponds to urban area.

Patch 1	Patch 2	Patch 3	Dimension
$\beta_2 = 2.7 \times 10^{-3}$	$\beta_4 = 5.5 \times 10^{-4}$	$\beta_6 = 5.5 \times 10^{-4}$	Days ⁻¹
$\gamma_2 = 0,9.0 \times 10^{-4}$	$\gamma_4 = 9.0 \times 10^{-5}$	$\gamma_6 = 7.3 \times 10^{-5}$	Days ⁻¹
$\mu_2 = 4.5 \times 10^{-5}$	$\mu_4 = 6.08 \times 10^{-5}$	$\mu_6 = 6.08 \times 10^{-5}$	Humans ⁻¹ \times days ⁻¹
$\alpha_2 = 0.0035$	$\alpha_4 = 0.0035$	$\alpha_6 = 0.0035$	Days ⁻¹
$\rho_2 = 0.0083$	$\rho_4 = 0.035$	$\rho_6 = 0.0335$	Days ⁻¹
$\Lambda_2 = 4.0$	$\Lambda_4 = 0.5$	$\Lambda_6 = 0.3$	Humans \times days ⁻¹
$\Lambda_1 = 700$	$\Lambda_3 = 500$	$\Lambda_5 = 600$	Mosquitoes \times days ⁻¹
$\mu_1 = 0.04$	$\mu_3 = 0.04$	$\mu_5 = 0.04$	Mosquitoes ⁻¹ \times days ⁻¹
$\tilde{a}_1 = 0.6$	$\tilde{a}_3 = 0.70$	$\tilde{a}_5 = 0.50$	1
$\tilde{a}_2 = 6.0$	$\tilde{a}_4 = 19.0$	$\tilde{a}_6 = 19.0$	1
$\sigma_{12} = 0.022$	$\sigma_{34} = 0.022$	$\sigma_{56} = 0.022$	1
$\sigma_{21} = 0.24$	$\sigma_{43} = 0.48$	$\sigma_{65} = 0.48$	1
$\hat{\sigma}_{21} = 0.024$	$\hat{\sigma}_{43} = 0.048$	$\hat{\sigma}_{65} = 0.048$	1

treatment are preferred to reduce the number of infections and semi-immunes in all patches. Interestingly, these results show that the dynamic of controls depends on the bounds that we choose for the controls.

3.3.3. Dynamics of Human Infection in the Three Patches

Figures 3 and 4 show that the increase of susceptible hosts involves a decrease of infectious hosts. Figure 4(a) shows that no action should be taken during the half time interval in a patch which is our urban area. The second half time should be considered the treatment of infections from two other patches. This is because it takes time $(T/2)$ for production of sick people from the reservoir area of infection, and after this time $(T/2)$, we realize that all the patches contain enough sick people. However, we must now apply a treatment in the area that is not a reservoir of infection initially. This treatment will be done by setting up at the entrances to the urban area by the distribution of drugs to fight malaria infection before accessing it. These measures of treatment become necessary to prevent the urban area, constitutes a reservoir of infection.

Figure 4(b) shows that the treatment is effective for the infectious $50(T/6)$ first days because patch 2 is a reservoir of infection. Between days $50(T/6)$ and $250(5T/6)$, we must apply simultaneously prevention and treatment, and after $250(5T/6)$ days, only prevention can be applied in this patch.

Figure 4(c) shows that during the first $200(2T/3)$ days we must apply simultaneously the prevention and the treatment, and after this time, only treatment should be applied to reduce the number of infections in patch 3.

3.3.4. Dynamics of Semi-Immune in the Three Patches

Figure 5(a) shows that no strategies must be applied during the $T/4$ first days for the semi-immune in patch 1. Between $T/4$ and $3T/4$ days, we must apply simultaneously prevention and treatment, and during the remaining period, only prevention should be applied.

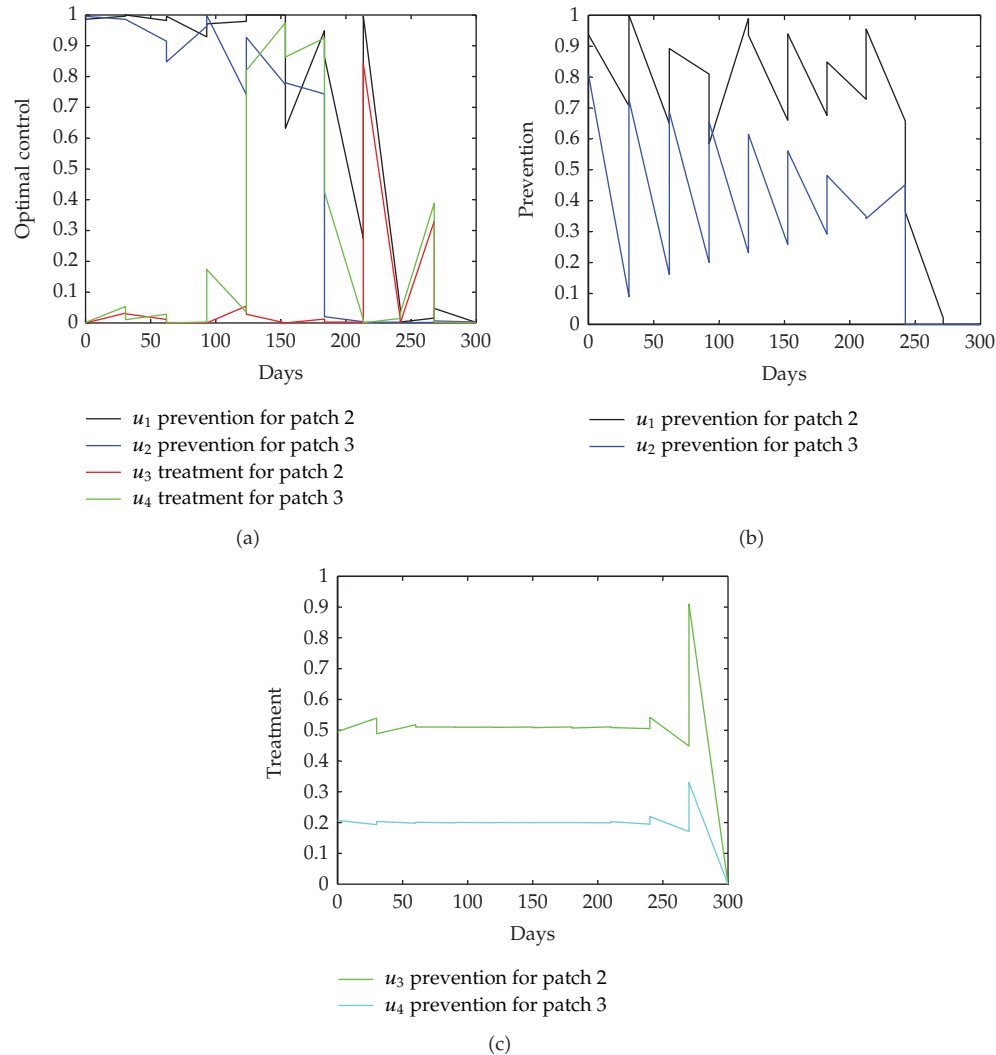


Figure 2: Results of the simulations achieved using data from Table 2 showing the evolution over time of the optimal control in the two reservoirs of infection. (a) Optimal control for prevention and treatment; (b) Optimal control for prevention control with no treatment control; (c) optimal control for treatment control with no prevention control.

Figures 5(b) and 5(c) show that the same strategies of control should be considered during the same period in patch 1 to reduce, respectively, the number of semi-immunes in patches 2 and 3.

To summarize, we used a recent technique of identification of the spatial infection of connected patches, to design a strategy based on the status of infection of the reservoirs of infection. We show that it is better to treat people only in areas that do not constitute a reservoir of infection and use simultaneously the prevention and the treatment to reduce the number of infections in all patches constituting a reservoir of infection. While reducing the number of semi-immunes, no differences in control strategies is made based on the type of infection reservoir. Whatever the level of infection of the reservoir of infection, the strategy

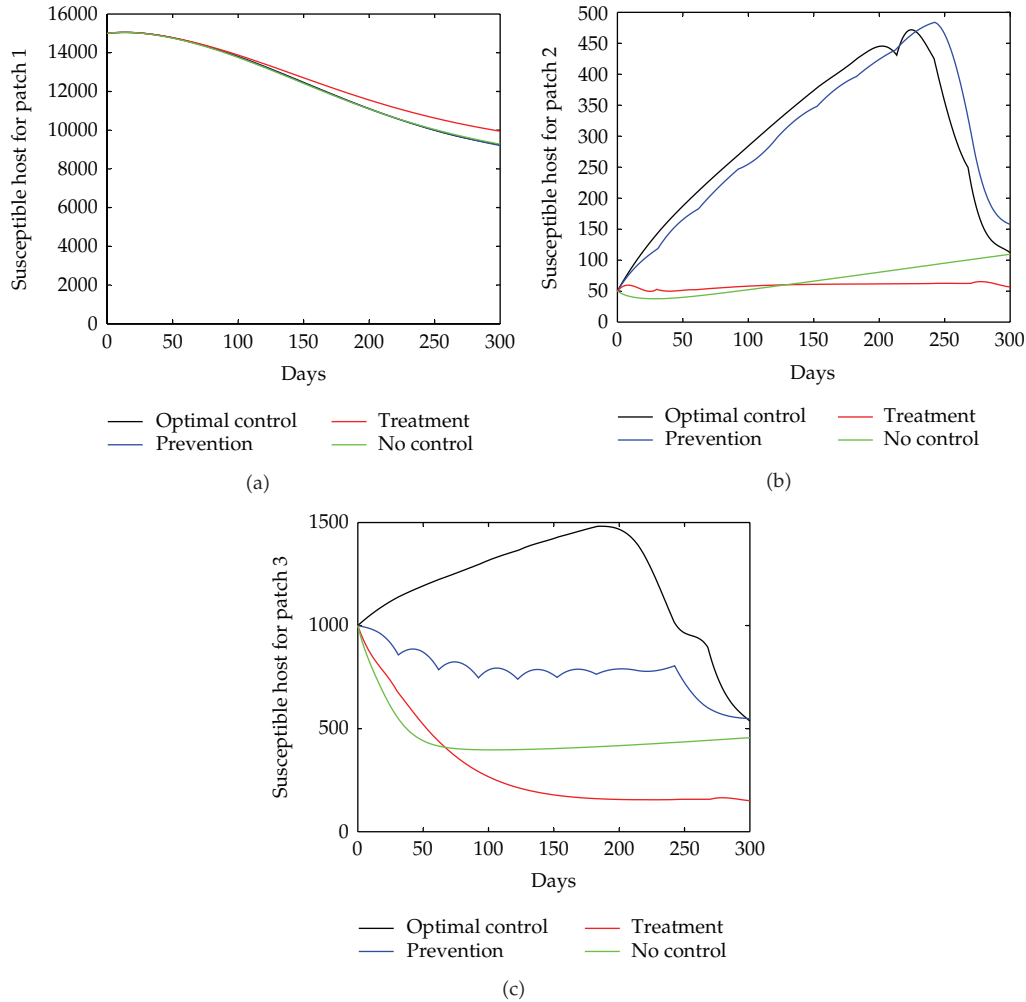


Figure 3: Results of the simulations achieved using data from Table 2 showing the evolution over time of susceptible hosts for the three patches. We show the four cases: *black line* optimal solution solved with prevention and treatment in the 2 patches; *blue line* optimal solution solved with prevention control in the 2 patches; *red line* optimal solution solved with treatment control in the 2 patches; *green line* solution without control. (a) Evolution of the susceptible hosts for patch 1; (b) evolution of the susceptible hosts for patch 2; (c) evolution of the susceptible hosts for patch 3.

to reduce the number of semi-immunes remains the same: no strategy is adopted in the early stage of malaria control, then both treatment and prevention are implemented, and in the last period, only prevention is implemented.

4. Conclusion

A mathematical model has been developed for malaria using the theory of optimal control. The formulation of the optimal control includes n control variables for prevention and n variables for treatment. The mathematical analysis proved the existence of an optimal control

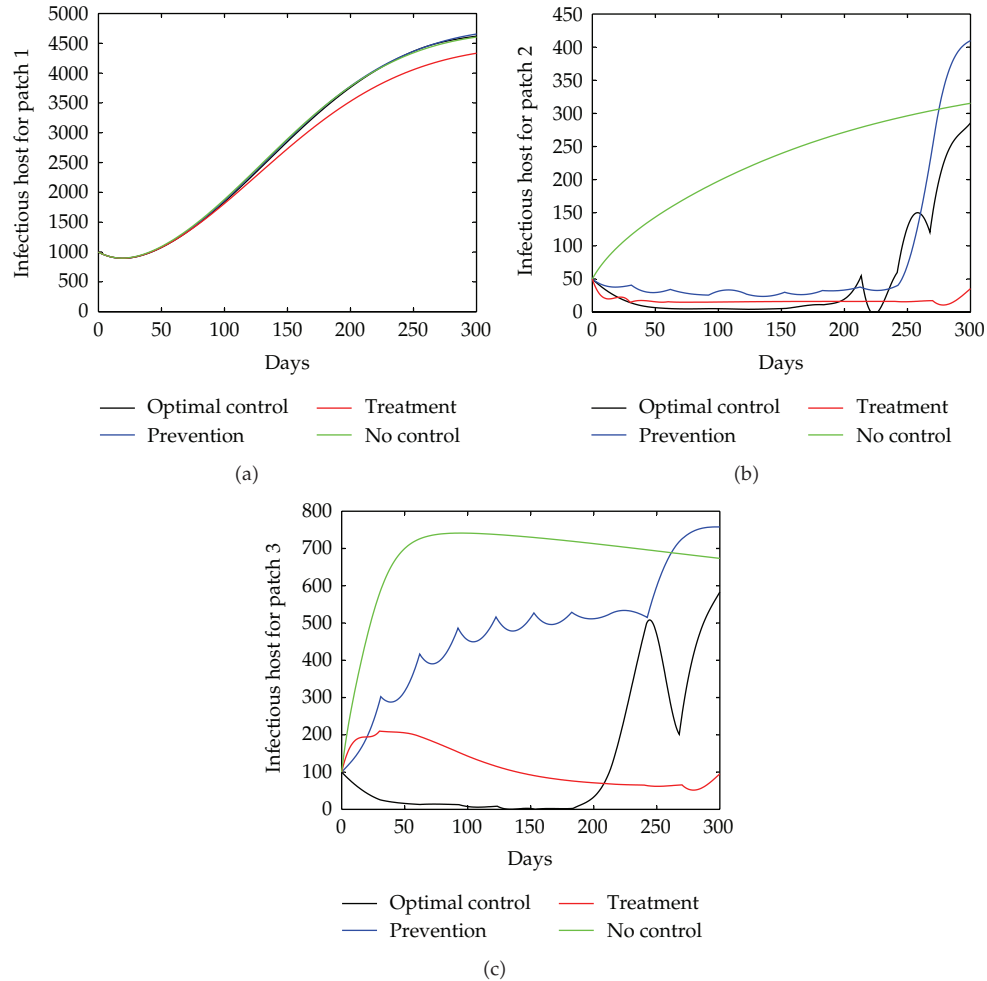


Figure 4: Results of the simulations achieved using data from Table 2 showing the evolution over time of infectious host for the three patches. We show the four cases: *black line* optimal solution solved with prevention and treatment in the 2 patches; *blue line* optimal solution solved with prevention control in the 2 patches; *red line* optimal solution solved with treatment control in the 2 patches; *green line* solution without control. (a) Evolution of the infectious hosts for patch 1; (b) evolution of the infectious hosts for patch 2; (c) evolution of the infectious hosts for patch 3.

for n connected patches under suitable conditions using the Fleming and Rishel theorem. Furthermore, using Pontryagin's maximum principle, a characterization of the optimal control was given. Numerical simulations were also performed showing the evolution over time of the optimal control as well as the different health status of humans and mosquitoes within each patch. These results underline the usefulness of a synergy control rather than only the prevention or the treatment. The results of our simulation show that we must choose a strategy based on the infectious status of the reservoir of infection. We show that it is better to treat people only in areas that do not constitute a reservoir of infection and use simultaneously the prevention and the treatment to reduce the number of infections in all patches constituting a reservoir of infection. While reducing the number of semi-immunes,

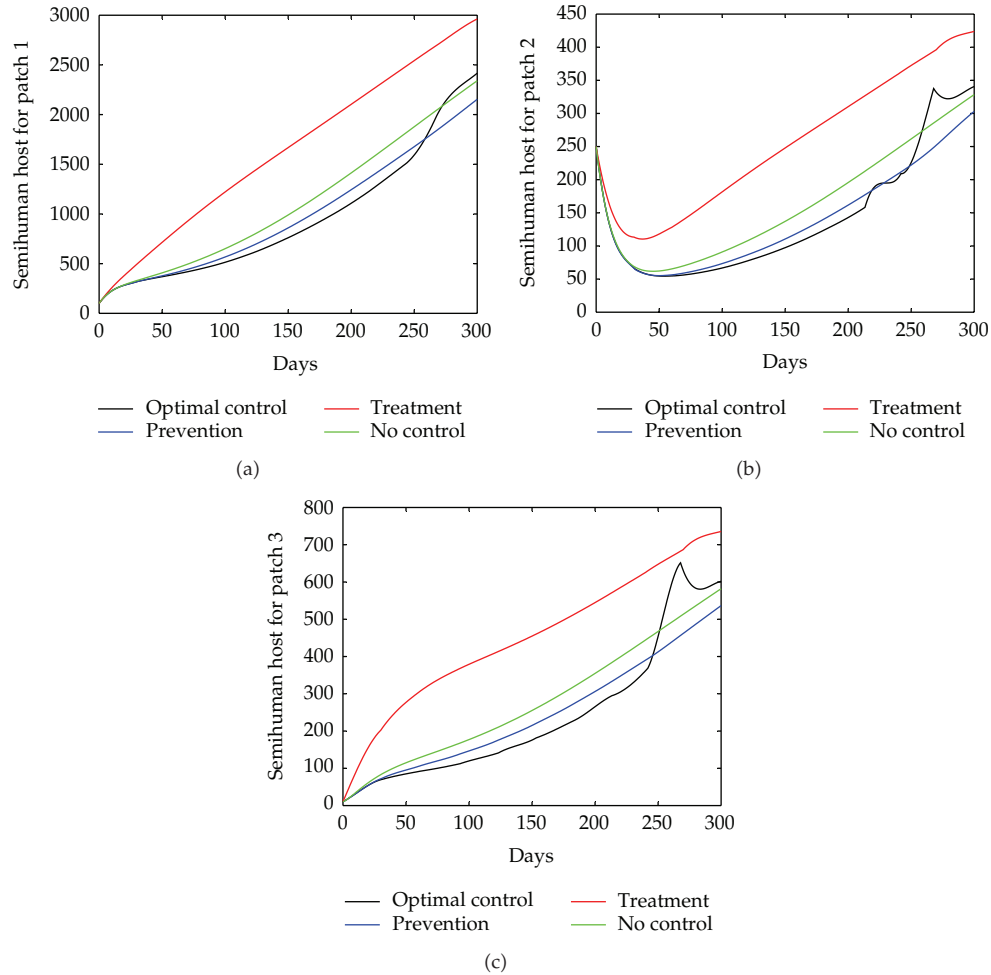


Figure 5: Results of the simulations achieved using data from Table 2 showing the evolution over time of semi-immune host for the three patches. We show the four cases: *black line* optimal solution solved with prevention and treatment in the 2 patches; *blue line* optimal solution solved with prevention control in the 2 patches; *red line* optimal solution solved with treatment control in the 2 patches; *green line* solution without control. (a) Evolution of the semi-immune hosts for patch 1; (b) evolution of the semi-immune hosts for patch 2; (c) evolution of the semi-immune hosts for patch 3.

no differences in control strategies is made based on the type of infection reservoir. Whatever the level of infection of the reservoir of infection, the strategy to reduce the number of semi-immunes remains the same: no strategy is adopted in the early stage of malaria control, then both treatment and prevention are implemented, and in the last period, only prevention is implemented.

To state our main perspectives, we will include a budget constraint in our optimal problem. Before characterizing the optimal prevention and treatment, two cases may arise under budget constraints: (i) when the budget allocation for the prevention and treatment is sufficient; (ii) when the budget is insufficient. Moreover, it would be interesting to apply a sensitivity analysis for some key parameters of the model.

References

- [1] P. Zongo, *Modélisation mathématique de la dynamique de transmission du paludisme* [Ph.D. thesis in Applied Mathematics], University of Ouagadougou, Ouagadougou, Burkina Faso, 2009, <http://tel.archives-ouvertes.fr/tel-00419519/en/>.
- [2] S. M. Goldman and J. Lightwood, "Cost optimization in the SIS model of infectious Disease with treatment," *The Berkley Electronic Journal of Economic Analysis and Policy*, vol. 2, no. 1, 2002.
- [3] E. Naevdal, "Fighting transient epidemics: optimal vaccination schedules before and after an outbreak," Working Paper, Health Economics Research Programme at the University of Oslo HERO, 2006.
- [4] A. Seierstad and K. Sydsæter, *Optimal Control Theory with Economic Applications*, vol. 24 of *Advanced Textbooks in Economics*, North-Holland, Amsterdam, The Netherlands, 1987.
- [5] S. M. Rowthorn, R. Laxminarayan, and C. A. Gilligan, "Optimal control of epidemics in metapopulations," *Journal of the Royal Society Interface*, vol. 6, no. 41, pp. 1135–1144, 2009.
- [6] S. Nanda, H. Moore, and S. Lenhart, "Optimal control of treatment in a mathematical model of chronic myelogenous leukemia," *Mathematical Biosciences*, vol. 210, no. 1, pp. 143–156, 2007.
- [7] J. M. Orellana, "Application du contrôle optimal à l'amélioration des trithérapies," *Comptes Rendus Mathématique*, vol. 348, no. 21–22, pp. 1179–1183, 2010.
- [8] W. Garira, S. D. Musekwa, and T. Shiri, "Optimal control of combined therapy in a single strain HIV-1 model," *Electronic Journal of Differential Equations*, vol. 2005, no. 52, pp. 1–22, 2005.
- [9] E. Jung, S. Lenhart, and Z. Feng, "Optimal control of treatments in a two-strain tuberculosis model," *Discrete and Continuous Dynamical Systems. Series B*, vol. 2, no. 4, pp. 473–482, 2002.
- [10] S. Bowong, "Optimal control of the transmission dynamics of tuberculosis," *Nonlinear Dynamics*, vol. 61, no. 4, pp. 729–748, 2010.
- [11] G. Zaman, "Optimal campaign in the smoking dynamics," *Computational and Mathematical Methods in Medicine*, vol. 2011, Article ID 163834, 9 pages, 2011.
- [12] K. W. Blayneh, A. B. Gumel, S. Lenhart, and T. Clayton, "Backward bifurcation and optimal control in transmission dynamics of West Nile virus," *Bulletin of Mathematical Biology*, vol. 72, no. 4, pp. 1006–1028, 2010.
- [13] D. Moulay, M. A. Aziz-Alaoui, and H.-D. Kwon, "Optimal control of chikungunya disease: larvae reduction, treatment and prevention," *Mathematical Biosciences and Engineering*, vol. 9, no. 2, pp. 369–392, 2012.
- [14] S. T. Stoddard, A. C. Morrison, G. M. Vazquez-Prokopec et al., "The role of human movement in the transmission of vector-borne pathogens," *PLoS Neglected Tropical Diseases*, vol. 3, no. 7, article e481, 2009.
- [15] N. M. Wayant, *Spatio-temporal analysis of malaria in Paraguay* [Thesis and Dissertations in Geography], University of Nebraska-Lincoln, Lincoln, Neb, USA, 2011.
- [16] A. Ducrot, S. B. Sirima, B. Somé, and P. Zongo, "A mathematical model for malaria involving differential susceptibility, exposedness and infectivity of human host," *Journal of Biological Dynamics*, vol. 3, no. 6, pp. 574–598, 2009.
- [17] N. Chitnis, J. M. Cushing, and J. M. Hyman, "Bifurcation analysis of a mathematical model for malaria transmission," *SIAM Journal on Applied Mathematics*, vol. 67, no. 1, pp. 24–45, 2006.
- [18] K. Blayneh, Y. Cao, and H.-D. Kwon, "Optimal control of vector-borne diseases: treatment and prevention," *Discrete and Continuous Dynamical Systems. Series B*, vol. 11, no. 3, pp. 587–611, 2009.
- [19] J. Arino, A. Ducrot, and P. Zongo, "A metapopulation model for malaria with transmission-blocking partial immunity in hosts," *Journal of Mathematical Biology*, vol. 64, no. 3, pp. 423–448, 2012.
- [20] Life Cycle of the Malaria Parasite with vaccines, 2008, <http://www.malariavaccine.org/malvac-lifecycle.php>.
- [21] G. A. Forster and C. A. Gilligan, "Optimizing the control of disease infestations at the landscape scale," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 12, pp. 4984–4989, 2007.
- [22] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer, BerlinM Germany, 1975.
- [23] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, Translated from the Russian by K. N. Trirogoff; edited by L. W. Neustadt, John Wiley & Sons, New York, NY, USA, 1962.
- [24] M. Čizniar, M. Fikar, and M. A. Latifi, "MATLAB dynamic optimisation code DYNOPT. User's guide," Tech. Rep., KIRP FCHPT STU, Bratislava, Slovakia, 2006.

- [25] O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz, "On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations," *Journal of Mathematical Biology*, vol. 28, no. 4, pp. 365–382, 1990.
- [26] P. van den Driessche and J. Watmough, "Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission," *Mathematical Biosciences*, vol. 180, pp. 29–48, 2002.

Research Article

Dynamic Analysis of a Predator-Prey (Pest) Model with Disease in Prey and Involving an Impulsive Control Strategy

Min Zhao,¹ Yanzhen Wang,² and Lansun Chen³

¹ School of Life and Environmental Science, Wenzhou University, Zhejiang, Wenzhou 325035, China

² School of Mathematics and Information Science, Wenzhou University, Zhejiang, Wenzhou 325035, China

³ Institute of Mathematics, Academia Sinica, Beijing 100080, China

Correspondence should be addressed to Min Zhao, zmcn@tom.com

Received 9 March 2012; Revised 1 May 2012; Accepted 1 May 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 Min Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The dynamic behaviors of a predator-prey (pest) model with disease in prey and involving an impulsive control strategy to release infected prey at fixed times are investigated for the purpose of integrated pest management. Mathematical theoretical works have been pursuing the investigation of the local asymptotical stability and global attractivity for the semitrivial periodic solution and population persistent, which depicts the threshold expression of some critical parameters for carrying out integrated pest management. Numerical analysis indicates that the impulsive control strategy has a strong effect on the dynamical complexity and population persistent using bifurcation diagrams and power spectra diagrams. These results show that if the release amount of infective prey can satisfy some critical conditions, then all biological populations will coexist. All these results are expected to be of use in the study of the dynamic complexity of ecosystems.

1. Introduction

Predator-prey models with disease are a major concern and are now becoming a new field of study known as ecoepidemiology. The disease factor in predator-prey systems has been firstly considered by Anderson and May [1]. In subsequent years, many authors studied the dynamics of ecological models with infected prey, and their papers mainly focused on this issue [2–8]. The infection rate and the predation rate are the two primary factors, which can control the chaotic dynamics of an ecoepidemiological system [9]. Das et al. [10] studied

the HP model [11] by introducing disease in prey populations, which can be described as follows:

$$\begin{aligned}
 \frac{ds}{dt} &= rs \left(1 - \frac{s}{k} \right) - ais - c_1 a_1 \frac{p_1 s}{b_1 + s}, \\
 \frac{di}{dt} &= ais - a_2 i p_1 - d_1 i, \\
 \frac{dp_1}{dt} &= a_1 \frac{p_1 s}{b_1 + s} + c_2 i p_1 - a_3 \frac{p_1 p_2}{b_2 + p_1} - d_2 p_1, \\
 \frac{dp_2}{dt} &= c_3 a_3 \frac{p_1 p_2}{b_2 + p_1} - d_3 p_2,
 \end{aligned} \tag{1.1}$$

where s , i , p_1 , and p_2 are respectively the susceptible prey population, infected prey population, the intermediate predator population, and the top-predator population, a_1 and a_2 are the maximal predation rate of intermediate predator for susceptible and infected prey, respectively, a_3 is the maximal predation rate of top-predator for intermediate predator, b_1 and b_2 are the half saturation constant for functional response of intermediate and the top-predator respectively, c_1 is the conversion rate of susceptible prey to intermediate predator, c_2 is the conversion rate of infected prey to intermediate predator, and c_3 is the conversion rate of intermediate predator to top predator.

Through the dimensionless transformation (seeing [10]), the system can change into the following form:

$$\begin{aligned}
 \frac{dx}{dt} &= x(1-x) - ais - b \frac{p_1 s}{1 + cs}, \\
 \frac{di}{dt} &= ais - dip_1 - ei, \\
 \frac{dp_1}{dt} &= f \frac{p_1 s}{1 + cs} + gip_1 - h \frac{p_1 p_2}{1 + mp_1} - jp_1, \\
 \frac{dp_2}{dt} &= k \frac{p_1 p_2}{1 + mp_1} - lp_2.
 \end{aligned} \tag{1.2}$$

In recent decades, technological revolutions have recently hit the industrial world; thus, infected population can now be controlled by many methods such as spraying pesticides and vaccination. It is well known that pest management involves using pesticides and releasing natural enemies, which have been focused by many researchers [12–14]. Control of an infected population can be achieved by chemical or biological control or both, which is called an impulsive control strategy in biomathematics. Systems with impulsive control strategies to describe time-varying processes are characterized by the fact that at certain moments, their states undergo abrupt change. Recently, impulsive control strategies have been recently introduced into population ecology [15–18], chemotherapeutic approaches to treat disease [19], and food webs [20–25].

Based on the two aspects discussed, the authors constructed a predator-prey model with disease in prey (a pest) and involving an impulsive control strategy for the purpose of

integrated pest management. The impulsive control strategy was used to introduce infected prey (a pest) at a fixed time on the basis of system (1.2). The predator-prey model with disease in prey and involving an impulsive control strategy can be described by the following differential equations:

$$\begin{aligned}
 \frac{dx(t)}{dt} &= x(t)(1 - x(t)) - ax(t)y(t) - b\frac{x(t)z(t)}{1 + cx(t)}, \\
 \frac{dy(t)}{dt} &= ax(t)y(t) - dy(t)z(t) - ey(t), \\
 \frac{dz(t)}{dt} &= f\frac{x(t)z(t)}{1 + cx(t)} + gy(t)z(t) - h\frac{q(t)z(t)}{1 + mz(t)} - jz(t), \\
 \frac{dq(t)}{dt} &= K\frac{q(t)z(t)}{1 + mz(t)} - lq(t),
 \end{aligned} \quad t \neq nT,$$

$$\begin{aligned}
 \Delta x(t) &= 0, \\
 \Delta y(t) &= p, \\
 \Delta z(t) &= 0, \\
 \Delta q(t) &= 0,
 \end{aligned} \quad t = nT,$$
(1.3)

where $x(t)$, $y(t)$, $z(t)$, and $q(t)$ are respectively the densities of susceptible prey (a pest), infected prey (a pest), the intermediate predator (natural enemy), and the top predator at time t . Then, $\Delta x(t) = x(t^+) - x(t)$, $\Delta y(t) = y(t^+) - y(t)$, $\Delta z(t) = z(t^+) - z(t)$, and $\Delta q(t) = q(t^+) - q(t)$. We have

$$\begin{aligned}
 a &= \frac{ak}{r}, & b &= \frac{c_1 a_1 k}{r b_1}, & c &= \frac{k}{b_1}, & d &= \frac{a_2 k}{r}, & e &= \frac{d_1}{r}, & f &= \frac{a_1 k}{b_1 r}, \\
 g &= \frac{c_2 a_2 k}{r}, & h &= \frac{a_3 k}{b_2 r}, & m &= \frac{k}{b_1}, & j &= \frac{d_2}{r}, & K &= \frac{c_3 a_3 k}{b_2 r}, & l &= \frac{d_3}{r},
 \end{aligned} \quad (1.4)$$

where a_1 and a_2 are the maximal predation rates of the intermediate predator on susceptible and infected prey respectively; a_3 is the maximal predation rate of the top predator on the intermediate predator; b_1 and b_2 are the half-saturation constants for functional response of the intermediate prey and the top predator respectively; c_1 is the conversion rate of susceptible prey to intermediate predators; c_2 and c_3 are, respectively, the conversion rate of infected prey to intermediate predators and the conversion rate of the intermediate predator to the top predator; d_1 , d_2 , d_3 are the death rates of infected prey, the intermediate predator, and the top predator, respectively; α is the incidence rate; r is the intrinsic growth rate; k is the carrying capacity (see [10]); $p > 0$ is the introduced amount of infective prey population at $t = nT$, $n \in N$, $N = \{0, 1, 2, \dots\}$, where T is the period of the impulsive control. It is known that pest outbreak will cause some serious ecological and economic problems, and we can directly gather infected prey to increase the amount of infected prey and indirectly carry out integrated pest management.

The paper is organized as follows: in the next section, a mathematical analysis of the model is carried out. Section 3 describes some numerical simulations, and the last section contains a brief discussion.

2. Mathematical Analysis

Some important notations, lemmas, and definitions will be provided, which are frequently used in subsequent proofs.

Let $R_+ = [0, +\infty)$, $R_+^4 = \{X = (x(t), y(t), z(t), q(t)) \in R^4 | X \geq 0\}$. Denote $f = (f_1, f_2, f_3, f_4)$ as the map defined by the right-hand side of the first, second, third, and fourth equations of system (1.3). Let $V_0 = \{V : R_+ \times R_+^4 \rightarrow R_+\}$, then V is said to belong to class V_0 if

(1) V is continuous on $(nT, (n+1)T] \times R_+^4$, $n \in N$, and for each $X \in R^4$ $\lim_{(t, \mu) \rightarrow (nT^+, X)} V(t, \mu) = V(nT^+, X)$ exists;

(2) V is locally Lipschitzian in X .

Definition 2.1 (see [26]). Let $V \in V_0$, and then, for $(nT, (n+1)T] \times R_+^4$, the upper right derivative of $V(t, X)$ with respect to the impulsive differential system (1.3) can be defined as

$$D^+V(t, X) = \lim_{h \rightarrow 0^+} \sup \frac{1}{h} [V(t+h, X+hf(t, X)) - V(t, X)]. \quad (2.1)$$

The solution of system (1.3) is a piecewise continuous function $X : R_+ \times R_+^4$, where $X(t)$ is continuous on $(nT, (n+1)T]$, $n \in N$, and $X(nT^+) = \lim_{t \rightarrow nT^+} X(t)$ exists. Obviously the smoothness properties of f can guarantee the global existence and uniqueness of the solution of system (1.3); for details see [26–28].

Definition 2.2 (see [21]). system (1.3) is said to be uniformly persistent if there is an $\omega > 0$ (independent of the initial conditions) such that every solution $(x(t), y(t), z(t), q(t))$ of system (1.3) satisfies the following:

$$\liminf_{t \rightarrow \infty} x(t) \geq \omega, \quad \liminf_{t \rightarrow \infty} y(t) \geq \omega, \quad \liminf_{t \rightarrow \infty} z(t) \geq \omega, \quad \liminf_{t \rightarrow \infty} q(t) \geq \omega. \quad (2.2)$$

Definition 2.3 (see [24]). System (1.3) is said to be permanent if there exists a compact region $\Omega_0 \subset \text{int } R_+^4$ such that every solution $(x(t), y(t), z(t), q(t))$ of system (1.3) will eventually enter and remain in the region Ω_0 .

Lemma 2.4 (see [24]). Suppose that $X(t)$ is a solution of system (1.3) with $X(0^+) \geq 0$; then $X(t) \geq 0$ for all $t \geq 0$. Furthermore, $X(t) > 0$, $t > 0$ if $X(0^+) > 0$.

Lemma 2.5. There exists a constant M such that $x(t) \leq M$, $y(t) \leq M$, $z(t) \leq M$, and $q(t) \leq M$ for each solution $X = (x(t), y(t), z(t), q(t))$ of system (1.3) for all sufficiently large t . Details can be found in Theorem 2.2 of [29].

Lemma 2.6 (see [26]). Let $V \in V_0$, and assume that

$$\begin{aligned} D^+V(t, X) &\leq g(t, V(t, X)), \quad t \neq nT, \\ V(t, X(t^+)) &\leq \Phi_n(V(t, X(t))), \quad t = nT, \end{aligned} \quad (2.3)$$

where $g : R_+ \times R_+ \rightarrow R$ is continuous in $(nT, (n+1)T]$ for $u \in R_+^2, n \in N, \lim_{(t,y) \rightarrow (nT^+)} g(t, v) = g(nT^+, u)$ existing, and $\phi_n^i (i = 1, 2) : R_+ \rightarrow R_+$ nondecreasing. Let $r(t)$ be a maximal solution of the scalar impulsive differential equation as follows:

$$\begin{aligned} \frac{du(t)}{dt} &= g(t, u(t)), \quad t \neq nT, \\ u(t^+) &= \Phi_n(u(t)), \quad t = nT, \\ u(0^+) &= u_0, \end{aligned} \quad (2.4)$$

existing on $(0, +\infty]$. Then $V(0^+, X_0) \leq u_0$, implying that $V(t, X(t)) \leq r(t), t \geq 0$, where $X(t)$ is any solution of system (1.3). Note that if certain smoothness conditions on g exist to guarantee the existence and uniqueness of solutions for (2.4), then $r(t)$ is the unique solution of (2.4).

For convenience, some basic properties of certain subsystems of system (1.3) are now provided as follows:

$$\begin{aligned} \frac{dy(t)}{dt} &= -ey(t), \quad t \neq nT, \\ y(t^+) &= y(t) + p, \quad t = nT, \\ y(0^+) &= y_0. \end{aligned} \quad (2.5)$$

Therefore, the following lemma holds.

Lemma 2.7 (see [26]). For a positive periodic solution $y^*(t)$ of system (2.5) and the solution $y(t)$ of system (2.5) with initial value $y_0 = y(0^+) \geq 0, |y(t) - y^*(t)| \rightarrow 0, t \rightarrow \infty$, where

$$\begin{aligned} y^*(t) &= \left(\frac{p \exp(-e(t - nT))}{1 - \exp(-eT)} \right), \quad t \in (nT, (n+1)T], n \in N, \\ y^*(0^+) &= \left(\frac{p}{1 - \exp(-eT)} \right), \\ y(t) &= \left(y(0^+) - \left(\frac{p}{1 - \exp(-eT)} \right) \right) \exp(-eT) + y^*(t). \end{aligned} \quad (2.6)$$

Next, the stability of susceptible prey and of predator-eradication periodic solutions will be studied.

Theorem 2.8. The solution $(0, y^*(t), 0, 0)$ is said to be locally asymptotically stable if $T < (p/e)$.

Proof. The local stability of periodic solution $(0, y^*(t), 0, 0)$ may be determined by considering the behavior of small-amplitude perturbations of the solution. Define

$$x(t) = u(t), \quad y(t) = v(t) + y^*(t), \quad z(t) = w(t), \quad q(t) = h(t). \quad (2.7)$$

Substituting (2.7) into (1.3), a linearization of the system can be obtained as follows:

$$\begin{aligned}
 \frac{du(t)}{dt} &= (1 - ay^*(t))u(t), \\
 \frac{dv(t)}{dt} &= ay^*(t)u(t) - ev(t) - dy^*(t)w(t), \\
 \frac{dw(t)}{dt} &= (gy^*(t) - j)w(t), \\
 \frac{dh(t)}{dt} &= -lh(t), \\
 \Delta u(t) &= 0, \\
 \Delta v(t) &= p, \\
 \Delta w(t) &= 0, \\
 \Delta h(t) &= 0,
 \end{aligned} \quad t \neq nT, \quad t = nT. \quad (2.8)$$

This can be rewritten as

$$\begin{pmatrix} u(t) \\ v(t) \\ w(t) \\ h(t) \end{pmatrix} = \phi(t) \begin{pmatrix} u(0) \\ v(0) \\ w(0) \\ h(0) \end{pmatrix}, \quad 0 \leq t \leq T, \quad (2.9)$$

where $\phi(t)$ satisfies

$$\frac{d\phi(t)}{dt} = \begin{pmatrix} 1 - ay^*(t) & 0 & 0 & 0 \\ ay^*(t) & -e & -dy^*(t) & 0 \\ 0 & 0 & gy^*(t) - j & 0 \\ 0 & 0 & 0 & -l \end{pmatrix}, \quad (2.10)$$

with $\phi(0) = I$, where I is the identity matrix, and

$$\begin{pmatrix} u(nT^+) \\ v(nT^+) \\ w(nT^+) \\ h(nT^+) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u(nT) \\ v(nT) \\ w(nT) \\ h(nT) \end{pmatrix}. \quad (2.11)$$

Hence, the stability of the periodic solution $(0, y^*(t), 0, 0)$ is determined by the eigenvalues of

$$\theta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \phi(t). \quad (2.12)$$

If the absolute values of all eigenvalues are less than one, the periodic solution $(0, y^*(t), 0, 0)$ is locally stable. Then all eigenvalues of ϕ can be denoted by $\lambda_1, \lambda_2, \lambda_3$, and λ_4 , where $\lambda_1 = \exp \int_0^T (1 - ay^*(t))dt$, $\lambda_2 = \exp(-eT) < 1$, $\lambda_3 = \exp \int_0^T (gy^*(t) - j)dt$, $\lambda_4 = \exp(-lT) < 1$.

Clearly, $|\lambda_3| = \exp(-gp) < 1$ with $|\lambda_1| < 1$ only if $T < (p/e)$ according to the Floquet theory of impulsive differential equations, and the periodic solution $(0, y^*(t), 0, 0)$ is locally stable. This completes the proof. \square

Theorem 2.9. *The solution $(0, y^*(t), 0, 0)$ is said to be globally attractive if $gM < j$ and*

$$\frac{1}{a} < \frac{p \exp(-(dM + e)T)}{1 - \exp(-(dM + e)T)}. \quad (2.13)$$

Proof. Let $V(t) = fKx(t) + bKz(t) + bhlq(t)$; then

$$V'|_{(1.1)} = fK(1 - ay(t))x(t) + bK(gy(t) - j)z(t) - fKx^2(t) - bhlq(t). \quad (2.14)$$

By Lemma 2.5, there exists a constant $M > 0$ such that $x(t) \leq M$, $y(t) \leq M$, $z(t) \leq M$, $q(t) \leq M$ for each solution $X = (x(t), y(t), z(t), q(t))$ of system (1.3) with sufficiently large t .

Then,

$$\frac{dy(t)}{dt} = ay(t)x(t) - dy(t)z(t) - ey(t) \geq -(dM + e)y(t), \quad t \neq nT \quad (2.15)$$

$$\Delta y = p, \quad t = nT$$

$$\begin{aligned} V|_{(1.1)} &= fK(1 - ay(t))x(t) + bK(gy(t) - j)z(t) - fKx^2(t) - bhlq(t) \\ &\leq fK(1 - ay(t))x(t) + bK(gM - j)z(t) - fKx(t) - bhlq(t). \end{aligned} \quad (2.16)$$

By Lemmas 2.6 and 2.7, there exists a $t_1 > 0$, and an $\varepsilon > 0$ can be selected to be small enough so that $y(t) \geq y_1^*(t) - \varepsilon$ for all $t \geq t_1$. By (2.15),

$$\begin{aligned} y(t) &\geq y_1^*(t) - \varepsilon = \frac{p \exp(-(dM + e)T)}{1 - \exp(-(dM + e)T)} - \varepsilon, \\ \lambda &\triangleq \frac{p \exp(-(dM + e)T)}{1 - \exp(-(dM + e)T)} - \varepsilon. \end{aligned} \quad (2.17)$$

Let $1 - a\lambda < 0$ and $gM - j < 0$. Therefore, when $t \geq t_1$, by (2.16), $V'|_{(1.1)} < 0$. So $V(t) \rightarrow 0$ and $x(t) \rightarrow 0$, $z(t) \rightarrow 0$, $q(t) \rightarrow 0$ as $t \rightarrow \infty$. It is known from the fact that the limiting state of system (1.3) is exactly system (2.5) and from Lemma 2.7 that $(0, y^*(t), 0, 0)$ is globally attractive. This completes the proof. \square

Theorem 2.10. *System (1.3) is permanent if $T > p/e$, $gM > j$,*

$$\frac{1}{a} > \frac{p \exp(-(dM + e)T)}{1 - \exp(-(dM + e)T)}, \quad bhlM > \frac{p \exp((aM - e)T)}{1 - \exp((aM - e)T)}. \quad (2.18)$$

Proof. From Lemma 2.5, there exists a constant $M > 0$ such that $x(t) \leq M$, $y(t) \leq M$, $z(t) \leq M$, $q(t) \leq M$ for each solution $X = (x(t), y(t), z(t), q(t))$ of system (1.3) with t sufficiently large.

From (2.15), it is known that $y(t) \geq y_1^*(t) - \varepsilon = (p \exp(-(dM+e)T)) / (1 - \exp(-(dM+e)T)) - \varepsilon \triangleq \delta_1$ for large enough t .

Therefore, it is only necessary to find a δ_2 that satisfies $x(t) > \delta_2$, $z(t) > \delta_2$, $q(t) > \delta_2$. This will be achieved in the following two steps.

Let $\delta_3 > 0, \delta_4 > 0, \gamma = e - a\delta_3$, and $V(t) = fKx(t) + bKz(t) + bhlq(t)$.

Then

$$\begin{aligned} V|_{(1.1)} &= fK(1 - ay(t))x(t) + bK(gy(t) - j)z(t) - fKx^2(t) - bhlq(t) \\ &\geq fK(1 - ay(t) - M)x(t) + bK(gy(t) - j)z(t) - fKx(t) - bhlq(t). \end{aligned} \quad (2.19)$$

First, it will be proved that there exists a $t_2 \in (0, +\infty)$ such that $x(t_2) > \delta_4$, $z(t_2) > \delta_4$, and $q(t_2) > \delta_4$ because $V(t)$ is ultimately bounded.

Next, it will be proved that $x(t) < \delta_3$, $z(t) < \delta_3$, $q(t) < \delta_3$ cannot hold for all $t \in (0, +\infty)$. Otherwise,

$$\begin{aligned} \frac{dy(t)}{dt} &= ay(t)x(t) - dy(t)z(t) - ey(t) \leq (a\delta_3 - e)y(t), \quad t \neq nT \\ \Delta y &= p, \quad t = nT. \end{aligned} \quad (2.20)$$

Then let $v_1(t)$ be the solution of

$$\begin{aligned} \frac{dv_1(t)}{dt} &= (a\delta_3 - e)v_1(t), \quad t \neq nT \\ \Delta v_1(t) &= p, \quad t = nT. \end{aligned} \quad (2.21)$$

It follows that $y(t) < v_1(t)$ and $v_1(t) \rightarrow v_1^*(t) (t \rightarrow \infty)$ where $v_1^*(t) = (p \exp(-\gamma(t-n)T)) / (1 - \exp(-\gamma T))$.

So there exists a $t_3 > 0$ such that

$$y(t) < v_1(t) < v_1^*(t) + \varepsilon_1 = \frac{p \exp(-\gamma(t-n)T)}{1 - \exp(-\gamma T)} + \varepsilon_1 < \frac{p}{1 - \exp(-\gamma T)} + \varepsilon_1. \quad (2.22)$$

Then

$$\begin{aligned} V|_{(1.1)} &= fK(1 - ay(t))x(t) + bK(gy(t) - j)z(t) - fKx^2(t) - bhlq(t) \\ &\geq fK(1 - ay(t) - M)x(t) + bK(gy(t) - j)z(t) - fKx(t) - bhlq(t) \\ &\geq fK \left(1 - a \frac{p}{1 - \exp(-\gamma T)} + \varepsilon_1 - M \right) x(t) + bK \left(g \frac{p}{1 - \exp(-\gamma T)} + \varepsilon_1 - j \right) z(t). \end{aligned} \quad (2.23)$$

According to the above conditions, $V|_{(1.1)} > 0$; then $V(t) \rightarrow \infty$ and $x(t) \rightarrow \infty$, $z(t) \rightarrow \infty$, $q(t) \rightarrow \infty$ as $t \rightarrow \infty$; however, this is a contradiction. Therefore, $V(t)$ is ultimately bounded.

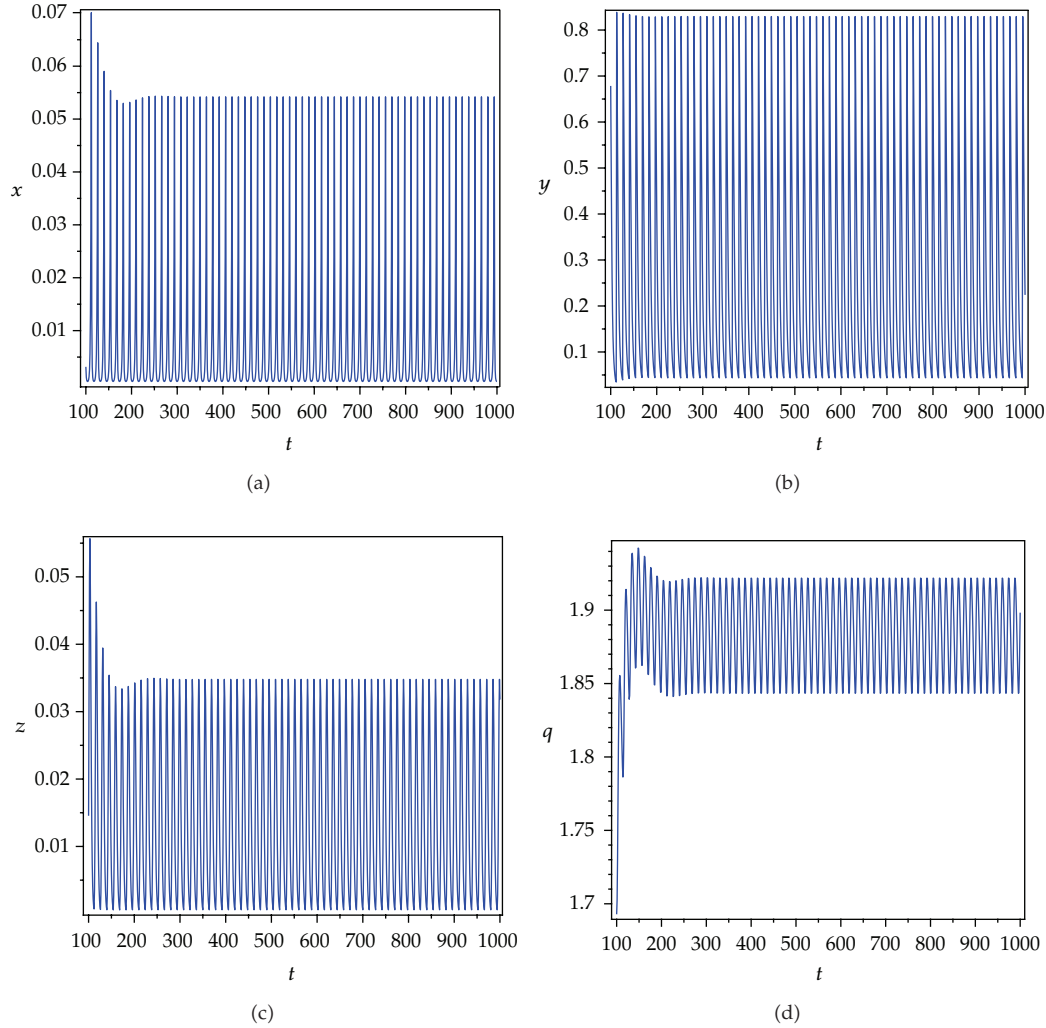


Figure 1: Dynamics of system (1.3) with $b = 8, d = 3, c = 3, e = 0.2, f = 3, g = 2.5, h = 0.3, m = 1.2, j = 0.2, k = 0.6, l := 0.008, a := 3, p = 1.03, T = 14$. Time series of (a) susceptible prey, (b) infected prey, (c) the intermediate predator, and (d) the top predator.

Second, if $x(t) > \delta_3, z(t) > \delta_3, q(t) > \delta_3$ for all $t \geq t_2$, then the objective has been attained. To show this, let $t^* = \inf_{t \geq t_2} \{V(t) < \delta_5\}$, and it follows that $V(t) \geq \delta_5$ for $t \in [t_2, t^*)$ and that $V(t^*) = \delta_5$. Suppose that $t^* \in (n_1 T, (n_1 + 1)T], n_1 \in N$. Select $n_2, n_3 \in N$ such that $n_2 T > (\ln(\varepsilon_1 / (M + p)) / -\gamma), \exp(n_3 \alpha_1 T) \exp(\alpha_2 (n_2 + 1)T) > 1$, where

$$\begin{aligned} \alpha_1 &= fK \left(1 - a \frac{p}{1 - \exp(-\gamma T)} + \varepsilon_1 - M \right) x(t) + bK \left(g - \frac{p}{1 - \exp(-\gamma T)} + \varepsilon_1 - j \right) z(t) > 0, \\ \alpha_2 &= -\frac{ap \exp((aM - e)T)}{1 - \exp((aM - e)T)} x(t) - jz(t) < 0. \end{aligned} \quad (2.24)$$

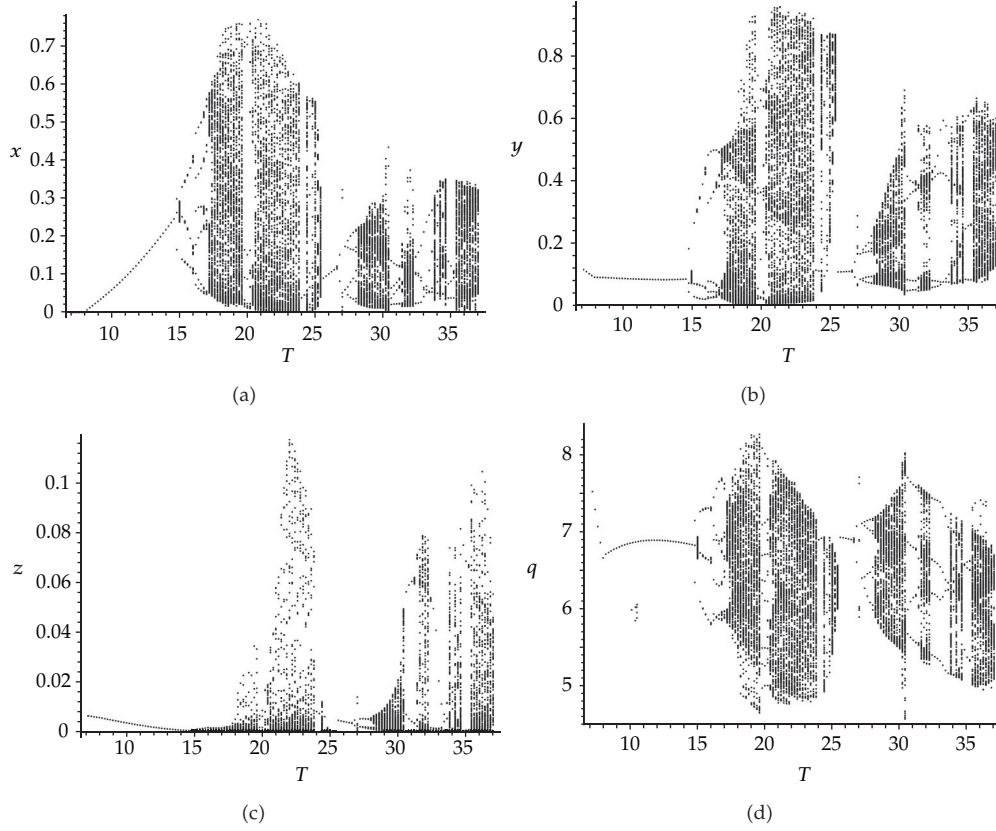


Figure 2: Bifurcation diagram of system (1.3) with initial conditions $x(0) = 0.1, y(0) = 0.2, z(0) = 0.3, q(0) = 0.4, 7 \leq T \leq 37, b = 6, d = 3, c = 3, e = 0.2, f = 2, g = 2.5, h = 0.1, m = 2; j = 0.1, k = 0.6, l = 0.01, a = 3$, and $p = 0.6$.

Let $T_2 = n_2T + n_3T$. It is claimed that there must be a $t_3 \in ((n_1 + 1)T, (n_1 + 1)T + T_2]$ such that $V(t) \geq \delta_5$. Otherwise, consider (2.21) with $v_1(t^{**}) = y(t^{**})$. Then

$$v_1(t) = \left(v_1((n_1 + 1)T^+) - \frac{p}{1 - \exp(-\gamma T)} \right) \exp(-\gamma(t - (n_1 + 1)T)) + v_1^*(t). \quad (2.25)$$

For $t \in ((n + 1)T, (n + 1)T]$, $n_1 + 1 < n < n_1 + n_2 + n_3 + 1$, it can be shown that $|v_1(t) - v_1^*(t)| < (M + p) \exp(-\gamma n_1 T) < \varepsilon_1$, $y(t) \leq v_1(t) \leq v_1^*(t) + \varepsilon_1$ for $t \in ((n_1 + n_2 + 1)T, (n_1 + 1)T + T_2]$,

$$\begin{aligned} V'|_{(1,1)} &\geq fK \left(1 - a \frac{p}{1 - \exp(-\gamma T)} + \varepsilon_1 - M \right) x(t) + bK \left(g \frac{p}{1 - \exp(-\gamma T)} + \varepsilon_1 - j \right) z(t) \\ &= \alpha_1 > 0, \end{aligned} \quad (2.26)$$

and $V((n_1 + 1)T + T_2) \geq V((n_1 + n_2 + 1)T) \exp(\alpha_1 n_3 T)$. For $t \in [t^*, (n_1 + n_2 + 1)T]$, it can be shown that $V|_{(1,1)} \geq -\alpha_2 V(t) > 0$; then, $V((n_1 + n_2 + 1)T) \geq V^*(t) \exp(-\alpha_2(n_2 + 1)T)$, so

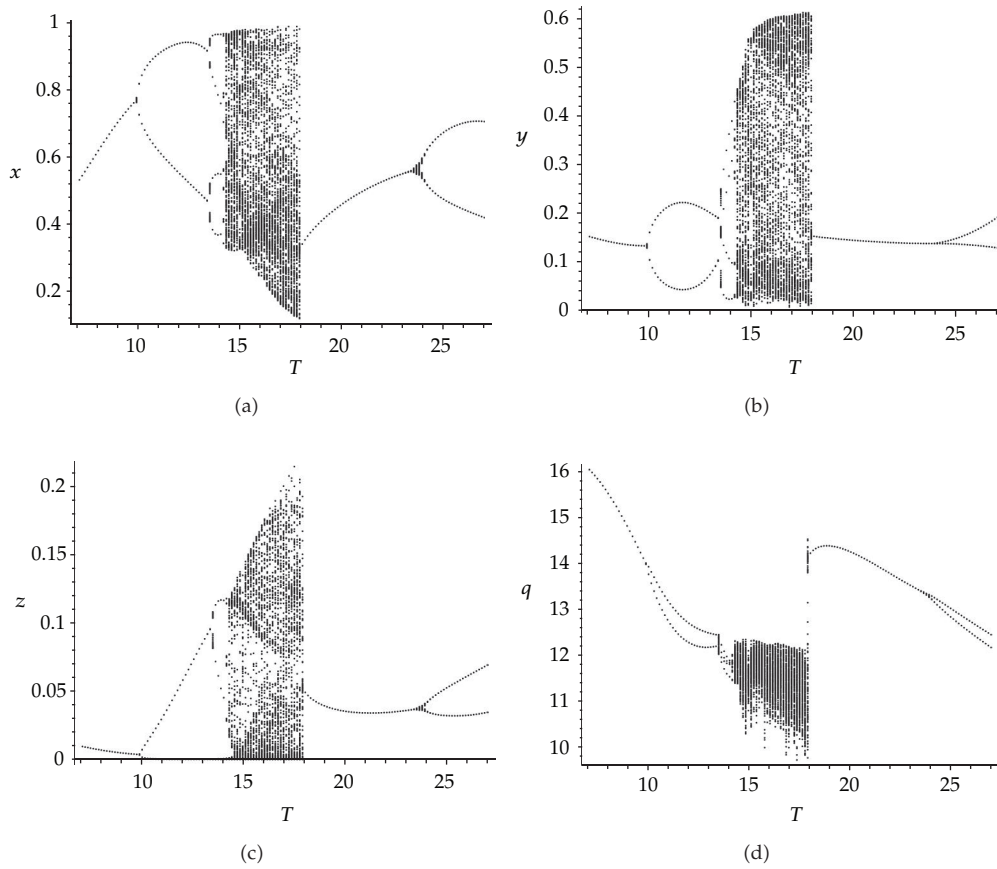


Figure 3: Bifurcation diagram of system (1.3) with initial conditions $x(0) = 0.1, y(0) = 0.2, z(0) = 0.3, q(0) = 0.4, 7 \leq T \leq 27, b = 10, d = 3, c = 3, e = 0.5, f = 5, g = 2.5, h = 0.1, m = 2; j = 0.2, k = 0.3, l = 0.01, a = 1.3$, and $p = 0.6$.

$V((n_1 + 1)T + T_2) \geq V^*(t) \exp(-\alpha_2(n_2 + 1)T + \alpha_1 n_3 T) > \delta_5$, which is a contradiction. Therefore, there exists a $t_3 \in ((n_1 + 1)T, (n_1 + 1)T + T_2]$ such that $V(t) \geq \delta_5$, resulting in $V(t) \geq V^*(t) \exp(-\alpha_2(n_1 + n_2 + n_3 + 1)T) \triangleq \delta_6$.

When $t \geq t_3$, the same procedure can be performed. According to the above discussion, if $\Omega_0 = \{(x(t), y(t), z(t), q(t)) : V(t) = fKx(t) + bKz(t) + bhlq(t), \delta \leq V(t) \leq M_1\} \subset \text{int } \mathbb{R}_+^3$, every solution of system (1.3) will eventually enter and remain in the region Ω_0 . This completes the proof. \square

3. Numerical Analysis

3.1. Bifurcation

To study the dynamics of system (1.3), the period T and the impulsive control parameter p are used as the bifurcation parameter. The bifurcation diagram provides a summary of the basic dynamic behavior of the system [30, 31].

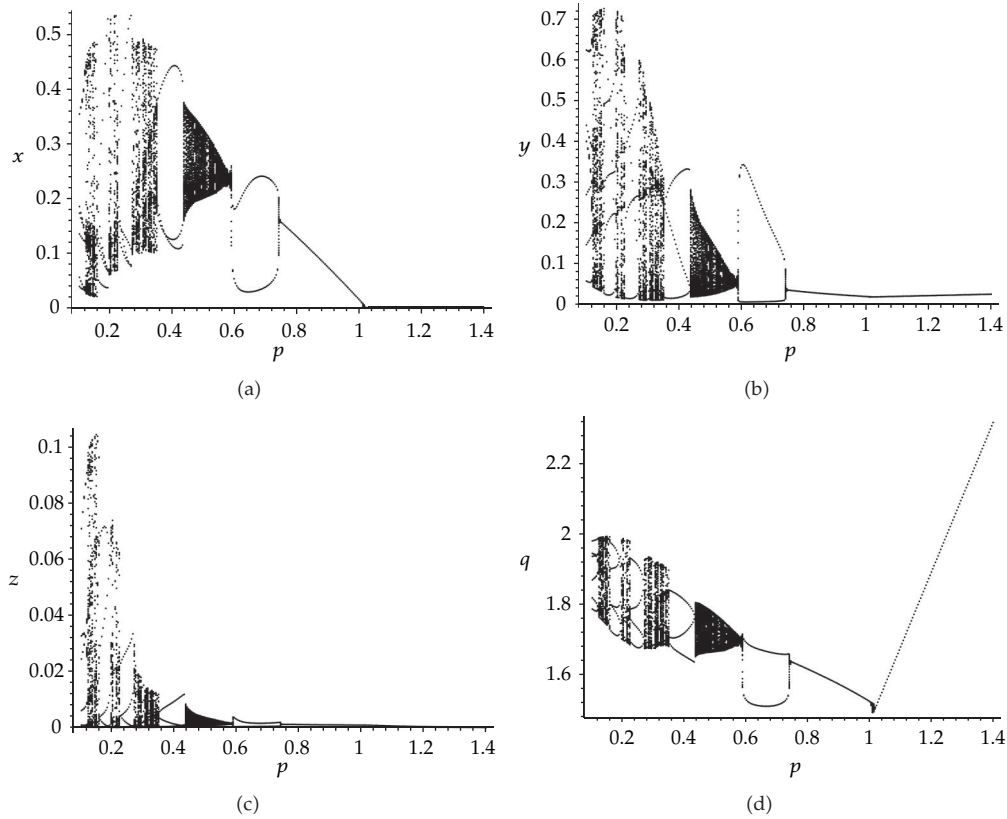


Figure 4: Bifurcation diagram of system (1.3) with initial conditions $x(0) = 0.1$, $y(0) = 0.2$, $z(0) = 0.3$, $q(0) = 0.4$, $0.1 \leq p \leq 2.4$, $b = 8$, $d = 3$, $c = 3$, $e = 0.2$, and $f = 3s$, $g = 2.5$, $h = 0.1$, $m = 0.2$, $j = 0.1$, $k = 0.6$, $l = 0.008$, $a = 3$, $p = 0.6$.

First, the influence of the period T is studied using the time series shown in Figure 1. The bifurcation diagrams are shown in Figures 2 and 3. Next, the influence of the impulsive control parameter p is investigated. The bifurcation diagrams for this are shown in Figure 4.

To clearly see the dynamics of system (1.3), it is necessary to examine the phase diagrams at different value of the period T and parameter p corresponding to the bifurcation diagrams in Figures 2 and 4; the results of this analysis are shown in Figures 5 and 6.

Figures 2, 3, and 4 reveal the complex dynamics of system (1.3), including period-doubling cascades, symmetry-breaking pitchfork bifurcation, chaos, and nonunique dynamics. Because every bifurcation diagram is similar, only one needs to be explained. Take Figure 4(a) as an example. When $p \in [0, 0.124]$, the dynamics of the system are not obvious, but with increasing p , the dynamics become more obvious. The system enters into a chaotic band with periodic windows. When p is between 0.124 and 0.153, the chaotic behavior is intense, as can be seen in Figure 6(a). When p moves beyond 0.153, the chaotic behavior disappears. When $p \in [0.203, 0.219]$, the chaotic attractor gains in strength, and the chaotic behavior appears again. When p becomes greater than 0.219, periodic windows appear, as can be seen in Figures 6(b) and 6(c). When p is in the interval between 0.328 and 0.35, chaotic behavior ensues, as can be seen in Figure 6(d). As the value of p increases further, the system enters a stable state, as is shown in Figures 6(e), 6(f), and 6(g). When p moves

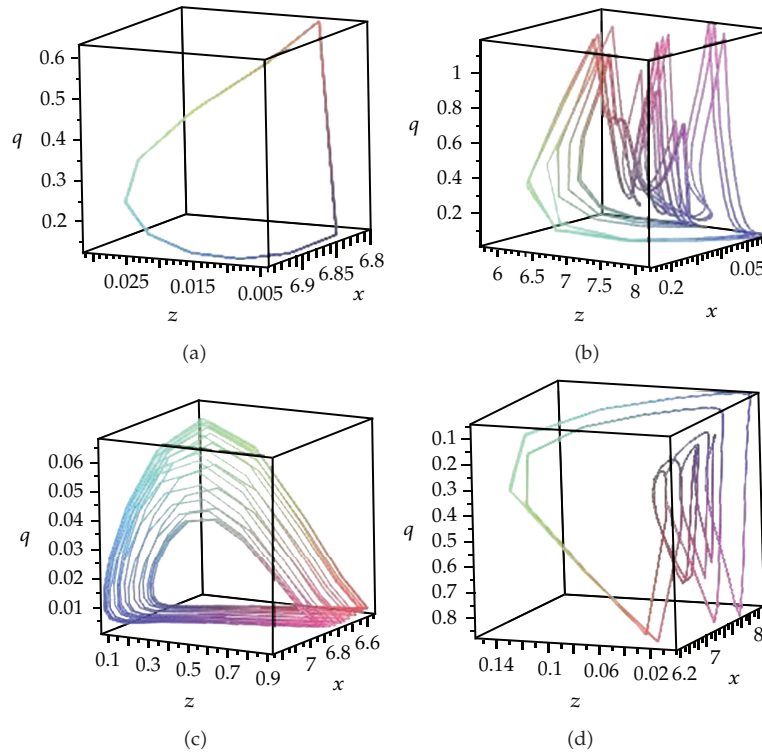


Figure 5: Periodic and chaotic behavior corresponding to Figure 2 as shown by phase diagrams: (a) $T = 10$, (b) $T = 17$, (c) $T = 15$, and (d) $T = 27.7$.

beyond 1.001, an unexpectedly chaotic phase appears, as is shown in Figure 6(h). It is clear that seasonal disturbances have little effect on the maximum density of all species; however, serious periodic oscillations are generated, and weak periodic solutions lose their stability and move into chaos. In summary, the key factor in the long-term dynamic behavior of system (1.3) is impulse perturbations, but seasonal disturbances can aggravate periodic oscillations and promote the emergence of chaos. Based on the above numerical simulation analysis, it is clear that impulsive control strategy has an important effect on the dynamical behaviors of the system, and weak periodic solutions lose their stability and move into chaos. In summary, the key factor in the long-term dynamic behavior of system (1.3) is impulsive control strategy, but disease disturbances can aggravate periodic oscillations and promote the emergence of chaos.

3.2. The Largest Lyapunov Exponent

To detect whether the system exhibits chaotic behavior, one of the commonest methods is to calculate the largest Lyapunov exponent. The largest Lyapunov exponent takes into account the average exponential rates of divergence or convergence of nearby orbits in phase space [32]. A positive largest Lyapunov exponent indicates that the system is chaotic. If the largest Lyapunov exponent is negative, there must be periodic windows or a stable state. Through the largest Lyapunov exponent, it is possible to judge that at what time the system is chaotic,

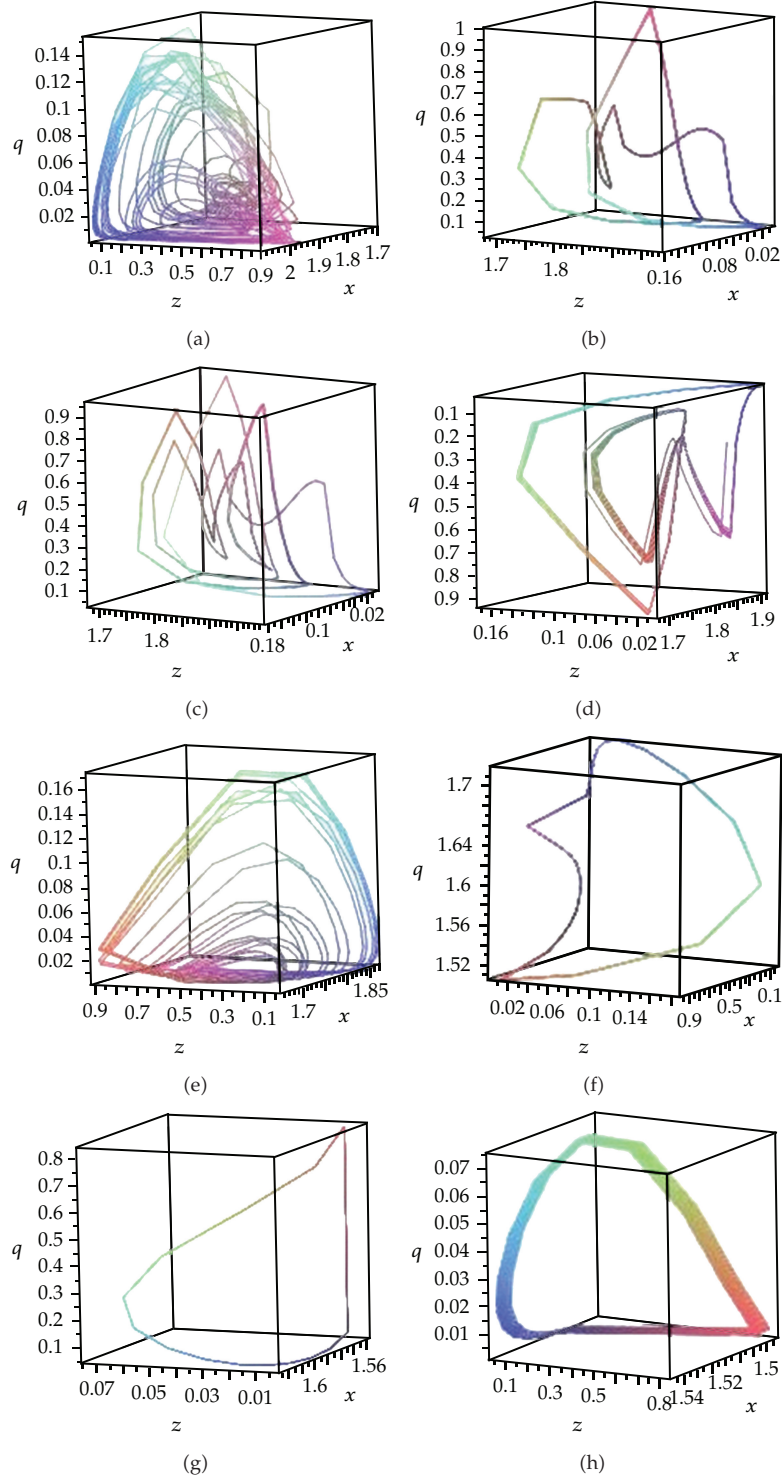


Figure 6: Periodic and chaotic behavior corresponding to Figure 3, as shown in phase diagrams: (a) $p = 0.13$, (b) $p = 0.25$, (c) $p = 0.3$, (d) $p = 0.35$, (e) $p = 0.45$, (f) $p = 0.7$, (g) $p = 0.9$, and (h) $p = 1.0009$.

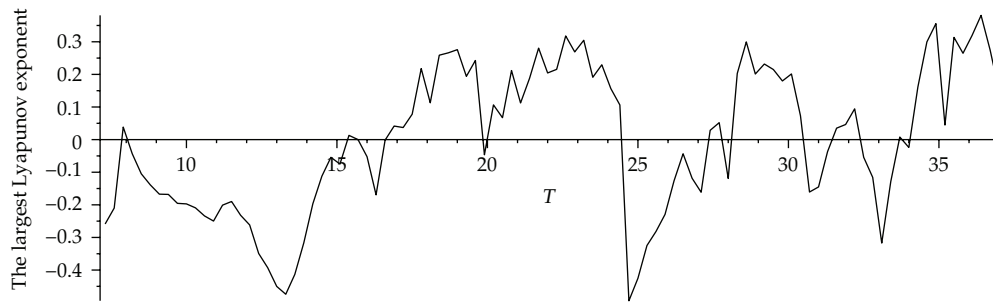


Figure 7: The largest Lyapunov exponents (LLE) corresponding to Figure 2.

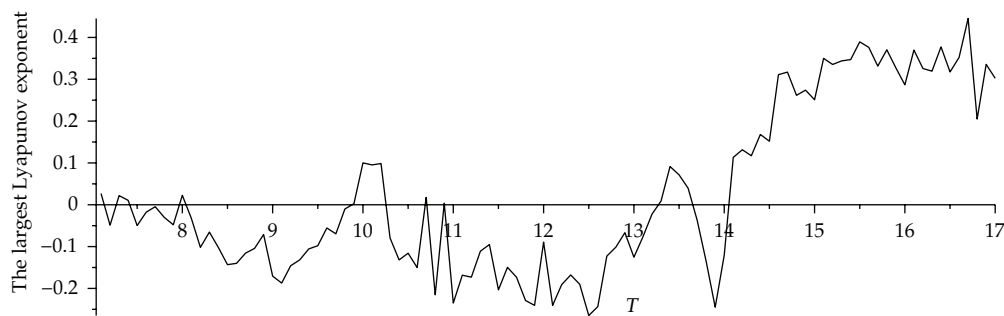


Figure 8: The largest Lyapunov exponents (LLE) corresponding to Figure 3.

and at what time the system is stable. The largest Lyapunov exponents corresponding to Figures 2, 3, and 4 can be calculated and are shown in Figures 7, 8, and 9, which shows the accuracy and effectiveness of numerical simulation. Moreover, using the simulation of the largest Lyapunov exponents, the existence of chaotic behavior in system (1.3) can be further confirmed.

3.3. Strange Attractors and Power Spectra

To understand the qualitative nature of strange attractors, power spectra are used [33]. From Section 3.2, it is known that the largest Lyapunov exponent for strange attractor (a) is 0.0413, and for strange attractor (b) is 0.124. Therefore, they are both chaotic attractors, and the exponent of (b) is larger than that of (a), which means that the chaotic dynamics of (b) are more extreme than those of (a). The power spectrum of strange attractor (a) is composed of strong broadband components and sharp peaks, as are shown in Figure 10(c). On the contrary, in the spectrum of strong chaotic attractor (b), it is difficult to distinguish any sharp peaks, as can be seen in Figure 10(d). These power spectra can be interpreted as meaning that (a) comes from a strong limit cycle, but that (b) experiences some weak limit cycles. Hence, it is obvious that the impulsive control strategy has a strong effect on the dynamical behaviors of system (1.3) with t the period of the impulsive control T varying but that (b) experiences some weak limit cycles.

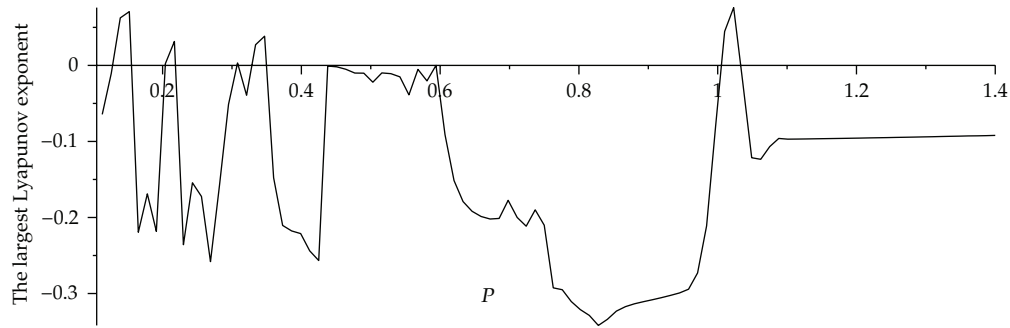


Figure 9: The largest Lyapunov exponents (LLE) corresponding to Figure 4.

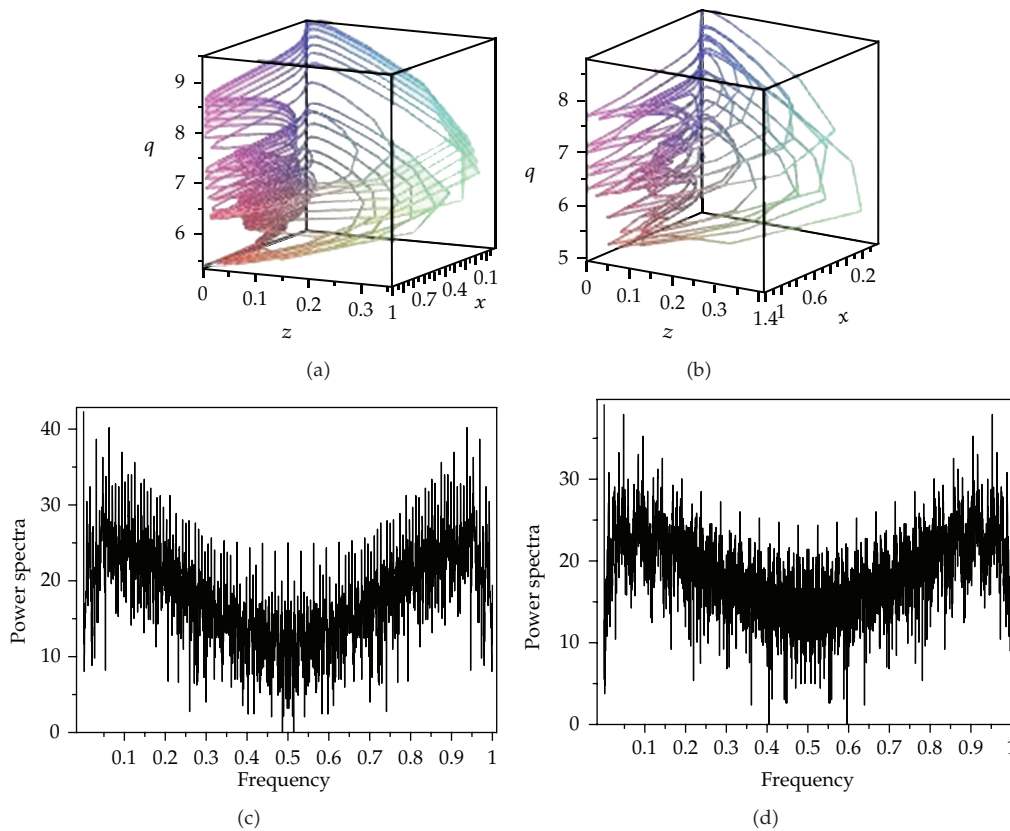


Figure 10: Strange attractors and power spectra: (a) strange attractor when $T = 32$, (b) strange attractor when $T = 21$, (c) power spectrum of attractor (a), and (d) power spectrum of attractor (b).

4. Conclusions and Remarks

In the paper, the dynamic behaviors of a predator-prey (pest) model with disease in prey and involving an impulsive control strategy are presented analytically and numerically. The critical conditions are obtained to ensure the local asymptotical stability and global attractivity of semitrivial periodic solution as well as population permanence. Numerical

analysis indicates that the impulsive control strategy has a strong effect on the dynamical complexity and population persistent using bifurcation diagrams and power spectra diagrams. In addition, the largest Lyapunov exponents are computed. This computation further confirms the existence of chaotic behavior and the accuracy of numerical simulation. These results revealed that the introduction of disease and the use of an impulsive control strategy can change the dynamic behaviors of the system. The same results also have been observed in continuous-time models of predator-prey or three-species food-chain models [34–37] and other systems [38]. In a word, it should be stressed that the impulsive control strategy is an effective method to control complex dynamics of predator-prey (pest) model.

Acknowledgment

The authors would like to thank the editor and the anonymous referees for their valuable comments and suggestions on this paper. This work was supported by the National Natural Science Foundation of China (NSFC no. 31170338 and no. 30970305) and also by the Key Program of Zhejiang Provincial Natural Science Foundation of China (Grant no. LZ12C03001).

References

- [1] R. M. Anderson and R. M. May, "The invasion, persistence and spread of infectious diseases within animal and plant communities," *Philosophical transactions of the Royal Society of London. Series B*, vol. 314, no. 1167, pp. 533–570, 1986.
- [2] K. P. Hadeler and H. I. Freedman, "Predator-prey populations with parasitic infection," *Journal of Mathematical Biology*, vol. 27, no. 6, pp. 609–631, 1989.
- [3] G.-P. Hu and X.-L. Li, "Stability and Hopf bifurcation for a delayed predator-prey model with disease in the prey," *Chaos, Solitons and Fractals*, vol. 45, no. 3, pp. 229–237, 2012.
- [4] W. Bob, A.K. George, V Voornb, and D. Krishna, "Stabilization and complex dynamics in a predator-prey model with predator suffering from an infectious disease," *Ecological Complexity*, vol. 8, no. 1, pp. 113–122, 2011.
- [5] X. Niu, T. Zhang, and Z. Teng, "The asymptotic behavior of a nonautonomous eco-epidemic model with disease in the prey," *Applied Mathematical Modelling*, vol. 35, no. 1, pp. 457–470, 2011.
- [6] E. Venturino, "Epidemics in predator-prey models: disease in the predators," *IMA Journal of Mathematics Applied in Medicine and Biology*, vol. 19, no. 3, pp. 185–205, 2002.
- [7] Y. Xiao and L. Chen, "Modeling and analysis of a predator-prey model with disease in the prey," *Mathematical Biosciences*, vol. 171, no. 1, pp. 59–82, 2001.
- [8] Y. Pei, S. Li, and C. Li, "Effect of delay on a predator-prey model with parasitic infection," *Nonlinear Dynamics*, vol. 63, no. 3, pp. 311–321, 2011.
- [9] S. Chatterjee, M. Bandyopadhyay, and J. Chattopadhyay, "Proper predation makes the system disease free - Conclusion drawn from an eco-epidemiological model," *Journal of Biological Systems*, vol. 14, no. 4, pp. 599–616, 2006.
- [10] K. P. Das, S. Chatterjee, and J. Chattopadhyay, "Disease in prey population and body size of intermediate predator reduce the prevalence of chaos-conclusion drawn from Hastings-Powell model," *Ecological Complexity*, vol. 6, no. 3, pp. 363–374, 2009.
- [11] A. Hastings and T. Powell, "Chaos in a three-species food chain," *Ecology*, vol. 72, no. 3, pp. 896–903, 1991.
- [12] X. Liu and L. Chen, "Complex dynamics of Holling type II Lotka-Volterra predator-prey system with impulsive perturbations on the predator," *Chaos, Solitons and Fractals*, vol. 16, no. 2, pp. 311–320, 2003.
- [13] R. Shi, X. Jiang, and L. Chen, "A predator-prey model with disease in the prey and two impulses for integrated pest management," *Applied Mathematical Modelling*, vol. 33, no. 5, pp. 2248–2256, 2009.
- [14] H. Liu, H. Xu, J. Yu, and G. Zhu, "Stability on coupling SIR epidemic model with vaccination," *Journal of Applied Mathematics*, no. 4, pp. 301–319, 2005.

- [15] G. Ballinger and X. Liu, "Permanence of population growth models with impulsive effects," *Mathematical and Computer Modelling*, vol. 26, no. 12, pp. 59–72, 1997.
- [16] L. Dong, L. Chen, and L. Sun, "Optimal harvesting policies for periodic Gompertz systems," *Nonlinear Analysis*, vol. 8, no. 2, pp. 572–578, 2007.
- [17] X. Song and Y. Li, "Dynamic behaviors of the periodic predator-prey model with modified Leslie-Gower Holling-type II schemes and impulsive effect," *Nonlinear Analysis*, vol. 9, no. 1, pp. 64–79, 2008.
- [18] X. Liu and L. Chen, "Complex dynamics of Holling type II Lotka-Volterra predator-prey system with impulsive perturbations on the predator," *Chaos, Solitons and Fractals*, vol. 16, no. 2, pp. 311–320, 2003.
- [19] A. Lakmeche and O. Arino, "Bifurcation of non trivial periodic solutions of impulsive differential equations arising chemotherapeutic treatment," *Dynamics of Continuous, Discrete and Impulsive Systems*, vol. 7, no. 2, pp. 265–287, 2000.
- [20] H. Yu, S. Zhong, R. P. Agarwal, and S. K. Sen, "Three-species food web model with impulsive control strategy and chaos," *Communications in Nonlinear Science and Numerical Simulation*, vol. 16, no. 2, pp. 1002–1013, 2011.
- [21] H. Yu and M. Zhao, "Seasonally perturbed prey-predator ecological system with the beddington-deangelis functional response," *Discrete Dynamics in Nature and Society*, vol. 2012, Article ID 150359, 12 pages, 2012.
- [22] M. Zhao, X. Wang, H. Yu, and J. Zhu, "Dynamics of an ecological model with impulsive control strategy and distributed time delay," *Mathematics and Computers in Simulation*, vol. 82, no. 8, pp. 1432–1444, 2012.
- [23] Y. Wang and M. Zhao, "Dynamic analysis of an impulsively controlled predator-prey model with holling type iv functional response," *Discrete Dynamics in Nature and Society*, vol. 2012, Article ID 141272, 18 pages, 2012.
- [24] J. Yang and M. Zhao, "Complex behavior in a fish algae consumption model with impulsive control strategy," *Discrete Dynamics in Nature and Society. An International Multidisciplinary Research and Review Journal*, vol. 2011, Article ID 163541, 17 pages, 2011.
- [25] J. Yang and M. Zhao, "A mathematical model for the dynamics of a fish algae consumption model with impulsive control strategy," *Journal of Applied Mathematics*, vol. 2012, Article ID 452789, 17 pages, 2012.
- [26] V. Lakshmikantham, D. D. Bařnov, and P. S. Simeonov, *Theory of Impulsive Differential Equations*, vol. 6 of *Series in Modern Applied Mathematics*, World Scientific Publishing, Singapore, 1989.
- [27] V. Lakshmikantham, D.D. Bainov, and P.S. Simeonov, *Impulsive Differential Equations: Periodic Solutions and Applications*, Pitman Monographs and Surveys in Pure and Applied Mathematics, 1993.
- [28] K. E. Brennan, S. L. Campbell, and L. R. Petzold, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, vol. 14 of *Classics in Applied Mathematics*, SIAM, Philadelphia, Pa, USA, 1996.
- [29] H. Yu, S. Zhong, and R. P. Agarwal, "Mathematics analysis and chaos in an ecological model with an impulsive control strategy," *Communications in Nonlinear Science and Numerical Simulation*, vol. 16, no. 2, pp. 776–786, 2011.
- [30] Z. W. Gao and D. W. C. Ho, "On state-space realization of Bezout factorizations in singular systems," *Dynamics of Continuous, Discrete & Impulsive Systems*, vol. 13, no. 3-4, pp. 387–410, 2006.
- [31] H. Yu, M. Zhao, S. Lv, and L. Zhu, "Dynamic complexities in a parasitoid-host-parasitoid ecological model," *Chaos, Solitons and Fractals*, vol. 39, no. 1, pp. 39–48, 2009.
- [32] F. Grond, H. H. Diebner, S. Sahle, A. Mathias, S. Fischer, and O. E. Rossler, "A robust, locally interpretable algorithm for Lyapunov exponents," *Chaos, Solitons and Fractals*, vol. 16, no. 5, pp. 841–852, 2003.
- [33] C. Masoller, A. C. S. Schifino, and L. Romanelli, "Characterization of strange attractors of lorenz model of general circulation of the atmosphere," *Chaos, Solitons and Fractals*, vol. 6, no. C, pp. 357–366, 1995.
- [34] S. Lv and M. Zhao, "The dynamic complexity of a three species food chain model," *Chaos, Solitons and Fractals*, vol. 37, no. 5, pp. 1469–1480, 2008.
- [35] M. Zhao and S. Lv, "Chaos in a three-species food chain model with a Beddington-DeAngelis functional response," *Chaos, Solitons and Fractals*, vol. 40, no. 5, pp. 2305–2316, 2009.
- [36] S. Lv and M. Zhao, "The dynamic complexity of a host-parasitoid model with a lower bound for the host," *Chaos, Solitons and Fractals*, vol. 36, no. 4, pp. 911–919, 2008.
- [37] M. Zhao, H. Yu, and J. Zhu, "Effects of a population floor on the persistence of chaos in a mutual interference host-parasitoid model," *Chaos, Solitons and Fractals*, vol. 42, no. 2, pp. 1245–1250, 2009.
- [38] Z. Gao and S. X. Ding, "Actuator fault robust estimation and fault-tolerant control for a class of nonlinear descriptor systems," *Automatica*, vol. 43, no. 5, pp. 912–920, 2007.

Research Article

Pareto Design of Decoupled Sliding-Mode Controllers for Nonlinear Systems Based on a Multiobjective Genetic Algorithm

**M. J. Mahmoodabadi,¹ A. Bagheri,¹ N. Nariman-Zadeh,^{1,2}
A. Jamali,¹ and R. Abedzadeh Maafi³**

¹ Department of Mechanical Engineering, Faculty of Engineering, University of Guilan,
P.O. Box 3756, Rasht, Iran

² Intelligent-Based Experimental Mechanics Center of Excellence, School of Mechanical Engineering,
Faculty of Engineering, University of Tehran, Tehran, Iran

³ Department of Mechanical Engineering, Takestan Branch, Islamic Azad University, Takestan, Iran

Correspondence should be addressed to M. J. Mahmoodabadi, mahmoodabadi@guilan.ac.ir

Received 11 January 2012; Revised 4 April 2012; Accepted 8 April 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 M. J. Mahmoodabadi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents Pareto design of decoupled sliding-mode controllers based on a multiobjective genetic algorithm for several fourth-order coupled nonlinear systems. In order to achieve an optimum controller, at first, the decoupled sliding mode controller is applied to stabilize the fourth-order coupled nonlinear systems at the equilibrium point. Then, the multiobjective genetic algorithm is applied to search the optimal coefficients of the decoupled sliding-mode control to improve the performance of the control system. Considered objective functions are the angle and distance errors. Finally, the simulation results implemented in the MATLAB software environment are presented for the inverted pendulum, ball and beam, and seesaw systems to assure the effectiveness of this technique.

1. Introduction

There are many control techniques that have been used to investigate the control behavior of the nonlinear systems [1–4]. A variable structure control with sliding mode, which is commonly known as sliding-mode control, is a nonlinear control strategy that is well-known for its guaranteed stability, robustness against parameter variations, fast dynamic response, and simplicity in implementation [1]. Although the sliding mode control method gives a satisfactory performance for the second-order systems, its performance for a fourth-order

coupled system is questionable. For example, in an inverted pendulum system controlled by the sliding-mode control, either the pole or cart can be successfully controlled, but not both. A remedy to this problem is to decouple the states and apply a suitable control law to stabilize the whole system. Recently, a decoupled sliding-mode control has been proposed to cope with this issue. It provides a simple way to decouple a class of fourth-order nonlinear systems in two second-order subsystems such that each subsystem has a separate control objective expressed in terms of a sliding surface [5, 6]. An important consequence of the using decoupled sliding-mode control is that the second subsystem is successfully incorporated into the first one via a two-level decoupling strategy.

It is very important to note that for design of the sliding-mode control and decoupled sliding-mode control, the sliding surface parameters should be determined, properly. This point is very crucial for the performance of the control system. The problem can be solved using evolutionary optimization techniques such as the genetic algorithm [7–10]. In this paper, a new intelligent decoupled sliding-mode control scheme based on an improved multiobjective genetic algorithm is proposed. Using this optimization algorithm, the important parameters of the decoupled sliding mode controller are optimized in a way to decrease the errors of the position and angle, simultaneously. The results obtained from this study illustrate that there are some important optimal design facts among objective functions which have been discovered via the Pareto optimum design approach. Such important design facts could not be found without using the multiobjective Pareto optimization process. In the end, simulations are presented to show the feasibility and efficiency of the proposed Pareto optimum decoupled sliding-mode control for the nonlinear systems.

2. Sliding-Mode Control

Sliding-mode controller is a powerful robust control strategy to treat the model uncertainties and external disturbances [11]. Furthermore, it has been widely applied to robust control of nonlinear systems [12–18]. In this section we recall the general concepts of sliding mode control for a second-order dynamic system. Suppose a nonlinear system is defined by the general state space equation as follows:

$$\dot{x} = f(x, u, t), \quad (2.1)$$

where $x \in R^n$ is the state vector, $u \in R^m$ the input vector, n is the order of the system, and m is the number of inputs. Then, the sliding surface $s(e, t)$ is given by the following:

$$s(e, t) = \{e : H^T e = 0\}, \quad (2.2)$$

where $H \in R^n$ represents the coefficients or slope of the sliding surface. Here,

$$e = x - x_d \quad (2.3)$$

is the negative tracking error vector.

Usually a time-varying sliding surface $s(t)$ is simply defined in the state-space R^n by the scalar equation as the following:

$$s(e, t) = \left(\frac{d}{dt} + \lambda \right)^{n-1} e = 0, \quad (2.4)$$

where λ is a strictly positive constant that can also be explained as the slope of the sliding surface. For instance, if $n = 2$ (for a second-order system) then,

$$s = \dot{e} + \lambda e, \quad (2.5)$$

and hence, s is simply a weighted sum of the position and velocity error from (2.4). The n th-order tracking problem is now being replaced by a first-order stabilization problem in which scalar s is to be kept at zero by a governing reaching condition. By choosing Lyapunov function $V(x) = (1/2)s^2$, the following equation can guarantee that the reaching condition is satisfied,

$$\dot{V}(x) = s\dot{s} < 0. \quad (2.6)$$

The existence and convergence conditions can be rewritten as follows:

$$s\dot{s} \leq -\eta|s|. \quad (2.7)$$

This equation permits a nonswitching region. Here, η is a strictly positive constant, and its value is usually chosen based on some knowledge of disturbances or system dynamics in terms of some known amplitudes.

In this control method, by changing the control law according to certain predefined rules which depend on the position of the error states of the system with respect to sliding surfaces, those states are switched between stable and unstable trajectories until they reach the sliding surface.

It can be shown that the sliding condition of (2.6) is always satisfied by the following:

$$u = u_{eq} - k \cdot \text{sgn}(s), \quad (2.8)$$

where u_{eq} is called equivalent-control input which is obtained by $\dot{s} = 0$. k is a design parameter and $k \geq \eta$.

Function sgn makes the high frequency chattering in control command. Using a proper definition of a thin boundary layer around the sliding surface, the chattering can be eliminated (Figure 1). This is accomplished by defining a boundary layer of thickness Φ , and replacing function sgn with function sat . This function is as the following and shown in Figure 2,

$$\text{sat}\left(\frac{s}{\Phi}\right) = \begin{cases} \text{sgn}\left(\frac{s}{\Phi}\right) & \text{if } \left|\frac{s}{\Phi}\right| \geq 1 \\ \left(\frac{s}{\Phi}\right) & \text{if } \left|\frac{s}{\Phi}\right| < 1 \end{cases}. \quad (2.9)$$

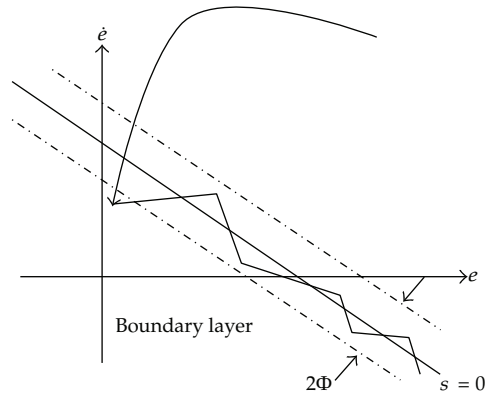


Figure 1: Sliding plant of a smooth controller.

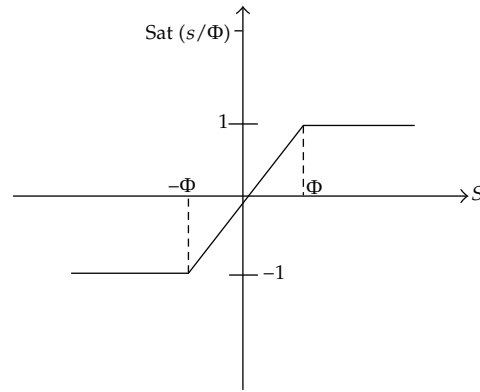


Figure 2: Function $\text{Sat}(s/\Phi)$ to eliminate the chattering phenomena in the sliding mode controller.

3. Inverted Pendulum System

In this section, the model of an inverted pendulum is recalled. In fact, the work deals with the stabilization control of a complicated, nonlinear, and unstable system. A pole, hinged to a cart moving on a track, is balanced upwards by motioning of the cart via a DC motor. The system observable state vector is $x = [x_1, x_2, x_3, x_4]^T$, including, respectively, the position of the cart, the angle of the pole with respect to the vertical axis, and their derivatives. The force to motion the cart may be expressed as $F = au$, where u is the input that is the limited motor supply voltage. The system dynamic model is as follows:

$$\begin{aligned}
 \dot{x}_1 &= x_3, \\
 \dot{x}_2 &= x_4, \\
 \dot{x}_3 &= f_1(x) + b_1(x)u, \\
 \dot{x}_4 &= f_2(x) + b_2(x)u,
 \end{aligned} \tag{3.1}$$

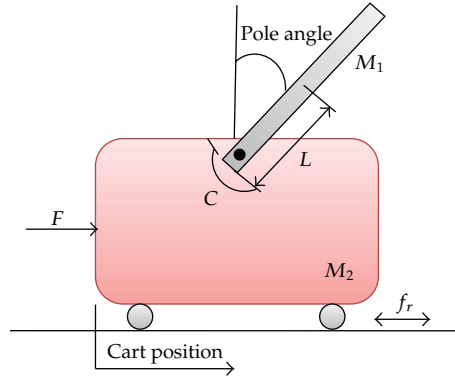


Figure 3: Inverted pendulum system.

where

$$\begin{aligned}
 f_1(x) &= \frac{a(-f_r - \mu x_4^2 \sin x_2) - l \cos(x_2)(\mu g \sin x_2 - Cx_4)}{J + \mu l \sin^2 x_2}, \\
 b_1(x) &= \frac{a\alpha}{J + \mu l \sin^2 x_2}, \\
 f_2(x) &= \frac{l \cos(x_2)(-f_r - \mu x_4^2 \sin x_2) + \mu g \sin x_2 - Cx_4}{J + \mu l \sin^2 x_2}, \\
 b_2(x) &= \frac{l \cos x_2 \alpha}{J + \mu l \sin^2 x_2},
 \end{aligned} \tag{3.2}$$

That is

$$l = \frac{LM_1}{2(M_2 + M_1)}, \quad a = l^2 + \frac{J}{M_2 + M_1}, \quad \mu = (M_2 + M_1)l. \tag{3.3}$$

Masses of the cart and pole are, respectively, M_2 and M_1 , g represents the gravity acceleration, L is the half length of the pole, and J is the overall inertia moment of the cart and pole with respect to the system centre of mass. C is the rotational friction coefficient of the pole, and f_r is the horizontal friction coefficient of the cart (Figure 3). This system is a nonlinear fourth-order system that includes two second-order subsystems in the canonical form with states $[x_1, x_3]^T$ and $[x_2, x_4]^T$.

4. Ball and Beam System

The ball and beam system is one of the most enduringly popular and important laboratory models for teaching control systems engineering. Because it is very simple to understand as a system, and control techniques that can stable it cover many important classical and modern design methods. The system has a very important property, it is open-loop unstable. The system is very simple, a steel ball rolling on the top of a long beam. The beam is mounted on

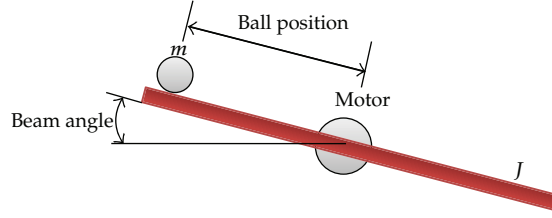


Figure 4: Ball and beam system.

the output shaft of an electrical motor, and so the beam can be tilted about its center axis by applying an electrical control signal to the motor amplifier. The control job is to automatically regulate the position of the ball on the beam by changing the angle of the beam. This is a difficult control task because the ball does not stay in one place on the beam, and moves with acceleration that is approximately proportional to the tilt of the beam. In control terminology, the system is open-loop unstable because the system output (the ball position) increases without any limitation for a fixed input (beam angle). Feedback control must be used to stabilize the system and to keep the ball in a desired position on the beam.

Consider a ball and beam system depicted in Figure 4 and its dynamic is described below:

$$\begin{aligned}
 \dot{x}_1 &= x_3, \\
 \dot{x}_2 &= x_4, \\
 \dot{x}_3 &= f_1(x) + b_1(x)u, \\
 \dot{x}_4 &= f_2(x) + b_2(x)u,
 \end{aligned} \tag{4.1}$$

where

$$\begin{aligned}
 f_1(x) &= -\frac{5}{7}(g \sin(x_2) - x_1 x_4^2), \\
 b_1(x) &= 0, \\
 f_2(x) &= \frac{-mx_1(2x_3x_4 - g \cos(x_2))}{mx_1^2 + J}, \\
 b_2(x) &= \frac{1}{mx_1^2 + J}.
 \end{aligned} \tag{4.2}$$

The mass of the ball is m , g represents the gravity acceleration, and J is the inertia moment of the beam (Figure 4). The system observable state vector is $x = [x_1, x_2, x_3, x_4]^T$, including, respectively, the position of the ball, the angle of the beam with respect to the horizontal axis, and their derivatives. This system is a nonlinear fourth-order system that includes two second-order subsystems in the canonical form with states $[x_1, x_3]^T$ and $[x_2, x_4]^T$.

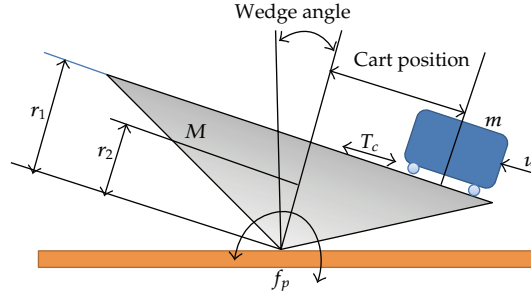


Figure 5: Seesaw system.

5. Seesaw System

According to the basic physical concepts, in the seesaw mechanism, if the vertical line along the centre of gravity of the inverted wedge is not passing through the fulcrum perpendicularly, then the inverted wedge will result in a torque and rotates until reaching the stable state. If we want to balance the inverted wedge, we have to put an external force to produce an appropriate opposite torque. For this reason, the inverted wedge is equipped with a cart to balance the unstable system. The cart can move to produce the appropriate torque against the internal force (Figure 5).

The observable state vector is $x = [x_1, x_2, x_3, x_4]^T$, including, respectively, the cart position, the wedge angle with respect to the vertical axis, and their derivatives. The system dynamic model is as the following:

$$\begin{aligned}\dot{x}_1 &= x_3, \\ \dot{x}_2 &= x_4, \\ \dot{x}_3 &= f_1(x) + b_1(x)u, \\ \dot{x}_4 &= f_2(x) + b_2(x)u,\end{aligned}\tag{5.1}$$

where

$$\begin{aligned}f_1(x) &= g \sin(x_2) - \frac{T_c}{m} x_3, \\ b_1(x) &= \frac{1}{m}, \\ f_2(x) &= \frac{Mgr_2 \sin(x_2) + mg\sqrt{x_1^2 + r_1^2} \sin(x_2 + \alpha) - f_p x_4}{J}, \\ b_2(x) &= \frac{r_1}{J},\end{aligned}\tag{5.2}$$

that is $\alpha = \tan^{-1}(x_1/r_1)$.

The cart and wedge masses are, respectively, m and M , g represents the gravity acceleration, r_1 is the height of the wedge, r_2 is the height of mass centre, J is the inertia

moment of the wedge, f_p is the rotational friction coefficient of the wedge, and T_c is the friction coefficient of the cart. This system is a nonlinear fourth-order system that includes two second-order subsystems in the canonical form with states $[x_1, x_3]^T$ and $[x_2, x_4]^T$.

6. Decoupled Sliding-Mode Control

Consider the nonlinear fourth-order coupled system expressed as the following.

$$\begin{aligned}\dot{x}_1 &= x_3, \\ \dot{x}_2 &= x_4, \\ \dot{x}_3 &= f_1(x) + b_1(x)u, \\ \dot{x}_4 &= f_2(x) + b_2(x)u.\end{aligned}\tag{6.1}$$

This system includes two second-order subsystems in the canonical form with states $[x_1, x_3]^T$ and $[x_2, x_4]^T$, and the sliding-mode control mentioned in the Section 2 can only control one of these subsystems. Hence, the basic idea of the decoupled sliding-mode control is proposed to design a control law such that the single input u simultaneously controls two coupled subsystems to accomplish the desired performance [5, 6, 19]. To achieve this goal, the following sliding surfaces are defined:

$$s_1(x) = \lambda_1 \cdot (x_2 - x_{2d} - z) + x_4 - x_{4d} = 0 \tag{6.2a}$$

$$s_2(x) = \lambda_2 \cdot (x_1 - x_{1d}) + x_3 - x_{3d} = 0. \tag{6.2b}$$

Here, z is a proportional value of s_2 and has a proper range with respect to x_2 . A comparison of (6.2a) with (2.5) shows the meaning of (6.2a): the control objective in the first subsystem of (6.1) changes from $x_2 = x_{2d}$ and $x_4 = x_{4d}$ to $x_2 = x_{2d} + z$ and $x_4 = x_{4d}$. On the other hand, (6.2b) has the same meaning of (2.5) and its control objectives are $x_1 = x_{1d}$ and $x_3 = x_{3d}$. Now, let the control law for (6.2a) be a sliding mode with a boundary layer, then:

$$u_1 = \hat{u}_1 - G_{f_1} \text{sat}(s_1(x)b_2(x)G_{s_1}), \quad G_{f_1}, G_{s_1} > 0, \tag{6.3}$$

with

$$\hat{u}_1 = -b_1^{-1}(x)(f_2(x) - \ddot{x}_{2d} + \lambda_1 x_4 - \lambda_1 \dot{x}_{2d}). \tag{6.4}$$

So

$$z = \text{sat}(s_2 \cdot G_{s_2}) \cdot G_{f_2}, \quad 0 < G_{f_2} < 1, \tag{6.5}$$

where G_{s_2} represents the inverse of the width of the boundary layer for s_2 , G_{f_2} transfers s_2 to the proper range of x_2 . Notice, in (6.5) z is a decaying oscillation signal since $G_{f_2} < 1$. Moreover, in (6.2a), if $s_1 = 0$, then $x_2 = x_{2d} + z$ and $x_4 = x_{4d}$.

Now, the control sequence is as follows: when $s_2 \neq 0$, then $z \neq 0$ in (6.2a) causes (6.3) to generate a control action that reduces s_2 ; as s_2 decreases, z decreases too. Hence, at the limit $s_2 \rightarrow 0$ with $x_1 \rightarrow x_{1d}$, then $z \rightarrow 0$ with $x_2 \rightarrow x_{2d}$; so, $s_1 \rightarrow 0$, and the control objective would be achieved [19].

7. Genetic Algorithm

Optimization in engineering design has always been of great importance and interest particularly in solving complex real-world design problems. Basically, the optimization process is defined as finding a set of values for a vector of design variables so that it leads to an optimum value of an objective or cost function. In such single-objective optimization problems, there may or may not exist some constraint functions on the design variables, and they are, respectively, referred to as constrained or unconstrained optimization problems. There are many calculus-based methods including gradient approaches to search for mostly local optimum solutions and these are well documented in [20, 21]. However, some basic difficulties in the gradient methods such as their strong dependence on the initial guess can cause them to find a local optimum rather than a global one. This has led to other heuristic optimization methods, particularly genetic algorithms (GAs) being used extensively during the last decade. Such nature-inspired evolutionary algorithms [22, 23] differ from other traditional calculus based techniques. The main difference is that GAs work with a population of candidate solutions, not a single point in search space. This helps significantly to avoid being trapped in local optima [24] as long as the diversity of the population is well preserved.

One of complex real-world problems is the controller design, because it is necessary to assign the control parameters. This parameter tuning is traditionally based on the trial and error procedure; however, this problem can be solved via evolutionary algorithms, for example, genetic algorithms. In the existing literature, several previous works have considered the evolutionary algorithms for control design. For an overview of evolutionary algorithms in the control engineering, [25] is appropriate. In particular, the pole placement procedure to design a discrete-time regulator in [26] and the observer-based feedback control design in [27] are formulated as multiobjective optimization problems and solved via genetic algorithms. Moreover, in [28], two decoupled sliding-mode control configurations are designed for a scale model of an oil platform supply ship while the genetic algorithm is used for optimization.

A simple genetic algorithm includes individual selection from population based on the fitness, crossover, and mutation with some probabilities to generate new individuals. With the genetic operation going on, the individual maximum fitness and the population average fitness are increased, steadily. When applied to a problem, GA uses a genetics-based mechanism to iteratively generate new solutions from currently available solutions. It then replaces some or all of the existing members of the current solution pool with the newly created members. The motivation behind the approach is that the quality of the solution pool should improve with the passage of time [22, 23].

8. Multiobjective Optimization

In multiobjective optimization problems which is also called multi-criteria optimization problems or vector optimization problems, there are several objective or cost functions

(a vector of objectives) to be optimized (minimized or maximized), simultaneously. These objectives often conflict with each other so that as one objective function improves, another deteriorates. Therefore, there is no single optimal solution that is best with respect to all the objective functions. Instead, there is a set of optimal solutions, well-known as Pareto optimal solutions [29–32], which distinguishes significantly the inherent natures between single-objective and multiobjective optimization problems.

In fact, multiobjective optimization has been defined as finding a vector of decision variables satisfying constraints to give acceptable values to all objective functions. Such multiobjective minimization based on Pareto approach can be conducted using some definitions [33].

8.1. Definition of Pareto Dominance

A vector $\vec{U} = [u_1, u_2, \dots, u_n]$, is dominance to vector $\vec{V} = [v_1, v_2, \dots, v_n]$ (denoted by $\vec{U} < \vec{V}$) if and only if for all $i \in \{1, 2, \dots, n\}$, $u_i \leq v_i \wedge \exists j \in \{1, 2, \dots, n\} : u_j < v_j$.

8.2. Definition of Pareto Optimality

A point $X^* \in \Omega$ (Ω is a feasible region in R^n) is said to be Pareto optimal (minimal) if and only if there is not $X \in \Omega$ which is dominance to X^* . Alternatively, it can be readily restated as following. For all $X \in \Omega$, $X \neq X^*$, $\exists i \in \{1, 2, \dots, m\} : f_i(X^*) < f_i(X)$.

8.3. Definition of Pareto Set

For a given multiobjective optimization problem, a Pareto set P^* is a set in the decision variable space consisting of all the Pareto optimal vectors. $P^* = \{X \in \Omega \mid \nexists X' \in \Omega : F(X') < F(X)\}$.

8.4. Definition of Pareto Front

For a given multiobjective optimization problem, the Pareto front PT^* is a set of vectors of objective functions which are obtained using the vectors of decision variables in the Pareto set P^* , that is $PT^* = \{F(X) = (f_1(X), f_2(X), \dots, f_m(X)) : X \in P^*\}$. In other words, the Pareto front PT^* is a set of the vectors of objective functions mapped from P^* .

In fact, evolutionary algorithms have been widely used for multiobjective optimization because of their natural properties suited for these types of problems. This is mostly because of their parallel or population-based search approach. Therefore, most of the difficulties and deficiencies within the classical methods in solving multiobjective optimization problems are eliminated. For example, there is no need for either several runs to find all individuals of the Pareto front or quantification of the importance of each objective using numerical weights. In this way, the original nondominated sorting procedure given by Goldberg [22] was the catalyst for several different versions of multiobjective optimization algorithms [29, 30]. However, it is very important that the genetic diversity within the population be preserved sufficiently. This main issue in multiobjective optimization problems has been addressed by many related research works [34]. Consequently, the premature convergence of multiobjective optimization evolutionary algorithms is prevented, and the

solutions are directed and distributed along the true Pareto front if such genetic diversity is well provided. The Pareto-based approach of NSGAII [33] has been used recently in a wide area of engineering multiobjective optimization problems because of its simple yet efficient non-dominance ranking procedure in yielding different level of Pareto frontiers. However, the crowding approach in such state-of-the-art multiobjective optimization problems [35] is not efficient as a diversity preserving operator [36]. In this paper, a new diversity preserving algorithm called ε -elimination diversity algorithm [36], as a multiobjective tool, searches the definition space of decision variables and returns the optimum answers in Pareto form. In this ε -elimination diversity approach that is used to replace the crowding distance assignment approach in NSGAII [33], all the clones and/or ε -similar individuals based on Euclidean norm of two vectors are recognized and simply eliminated from the current population. Therefore, based on a predefined value of ε as the elimination threshold ($\varepsilon = 0.01$ has been used in this paper) all the individuals in a front within this limit of a particular individual are eliminated. It should be noted that such ε -similarity must exist both in the space of objectives and in the space of the associated design variables. This will ensure that very different individuals in the space of design variables having ε -similarity in the space of objectives will not be eliminated from the population. Evidently, the clones or ε -similar individuals are replaced from the population with the same number of new randomly generated individuals. Meanwhile, this will additionally help to explore the search space of the given multiobjective optimization problems more efficiently [36].

9. Multiobjective Optimization of Decoupled Sliding Mode Control

As mentioned before this, it is necessary for the practical engineering applications to solve the optimization problems involving multiple design criteria which are also called objective functions. Furthermore, the design criteria may conflict with each other so that improving one of them will deteriorate since another. The inherent conflicting behavior of such objective functions lead to a set of optimal solutions named Pareto solutions. These types of problems can be solved using evolutionary multiobjective optimization techniques. Here, for multiobjective optimization of the decoupled sliding mode controller, vector $[G_{f_1}, G_{s_1}, \lambda_1, G_{f_2}, G_{s_2}, \lambda_2]$ is the vector of selective parameters of the decoupled sliding mode controller. G_{f_1} and G_{s_1} are positive constant. λ_1 and λ_2 are coefficients of sliding surfaces, and G_{s_2} represents the inverse of the width of the boundary layer of s_2 . G_{f_2} transfers s_2 to the proper range of x_2 . The error of the position and the error of the angle are functions of this vector's components. This means that by selecting various values for the selective parameters, we can make changes in the position and angel errors. In this paper, we are concerned in choosing values for the selective parameters to minimize above two functions. Clearly, this is an optimization problem with two object functions (errors of position and angle) and six decision variables $[G_{f_1}, G_{s_1}, \lambda_1, G_{f_2}, G_{s_2}, \lambda_2]$. The regions of the selective parameters are as follows:

$G_{s_2}, G_{f_1}, G_{s_1}$: positive constant, $G_{s_2}, G_{f_1}, G_{s_1} > 0$,

λ_1, λ_2 : coefficients of the sliding surface, $\lambda_1, \lambda_2 > 0$,

G_{f_2} : transfers s_2 to a proper range of x_2 , $0 < G_{f_2} < 1$.

The following parameters of the genetic algorithm are considered.

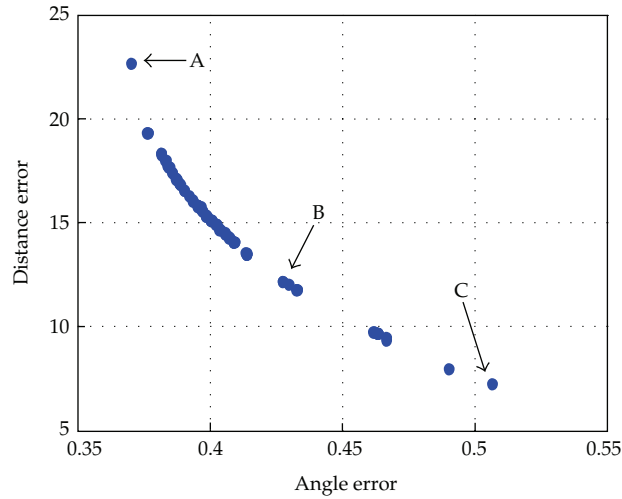


Figure 6: Pareto front of the angle error and distance error for the inverted pendulum.

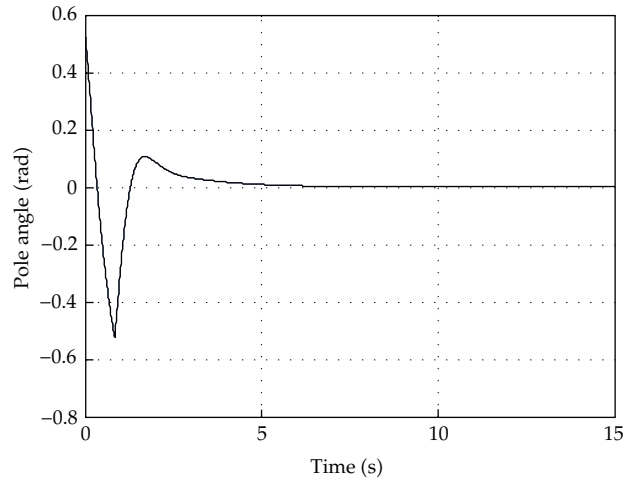


Figure 7: Simulation results for the pole angle.

Population size = 100, chromosome length = 48, generations = 300, crossover probability = 0.8, and mutation probability = 0.02. Also, the stopping criterion for this algorithm is the maximum number of generations.

10. Simulation and Results for the Inverted Pendulum System

The simulation for the inverted pendulum system considered here is carried out by MATLAB software. The initial values are as the following:

$$x_1(0) = 0, \quad x_2(0) = \frac{\pi}{6} \text{ rad}, \quad x_3(0) = 0, \quad x_4(0) = 0. \quad (10.1)$$

The system parameters and constants used in the simulation are given in Table 1.

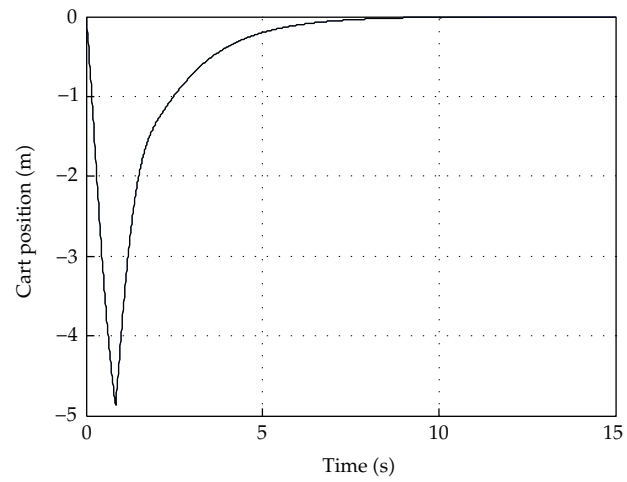


Figure 8: Simulation results for the cart position.

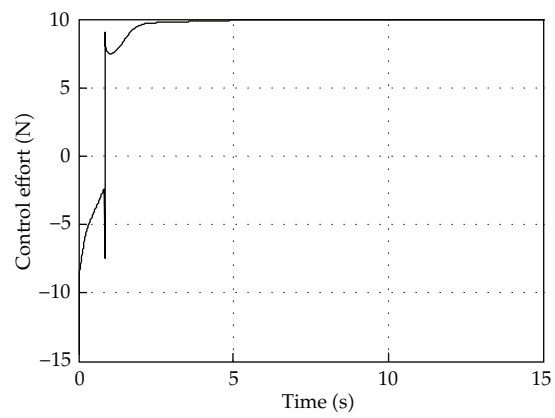


Figure 9: Simulation results for the control action.

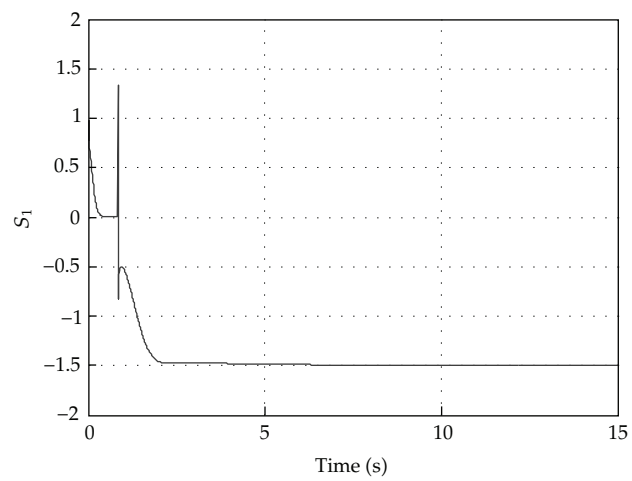


Figure 10: Sliding surface $s_1(x)$.

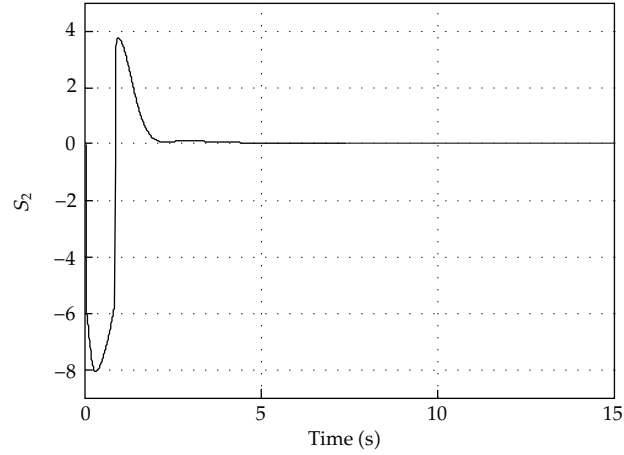


Figure 11: Sliding surface $s_2(x)$.

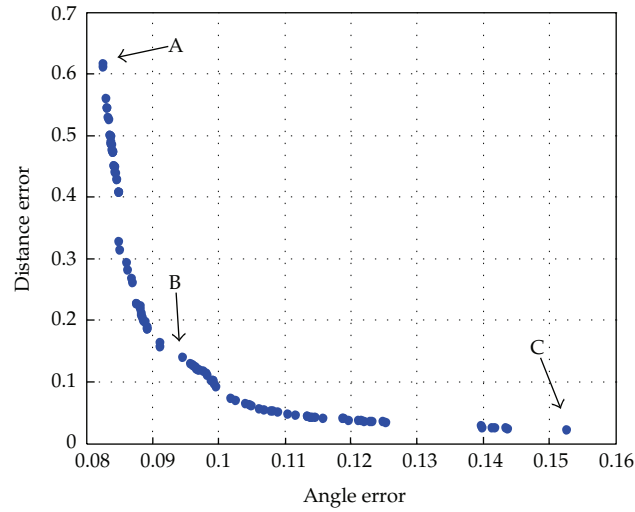


Figure 12: Pareto front of the angle error and distance error for the ball and beam system.

Table 1: Inverted pendulum parameters.

The mass of the pole	M_1	0.5
The mass of the cart	M_2	2
The half length of the pole	L	0.5
The inertia moment of the cart and pole	J	0.4
The friction constant of the pole	C	0.1
The friction constant of the cart	f_r	0.25
The gravity acceleration	g	9.81
The force coefficient	α	3

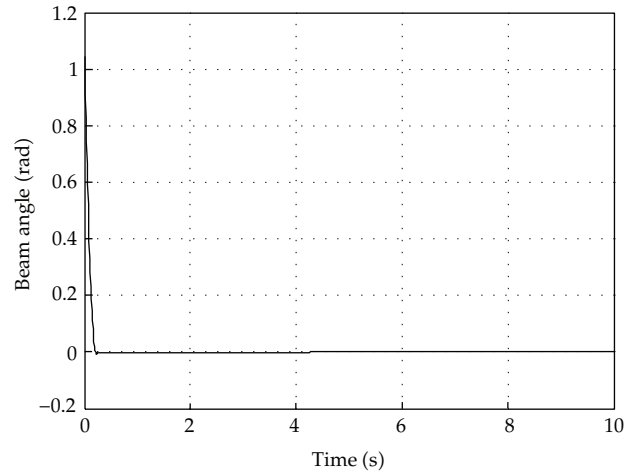


Figure 13: Simulation results for the beam angle.

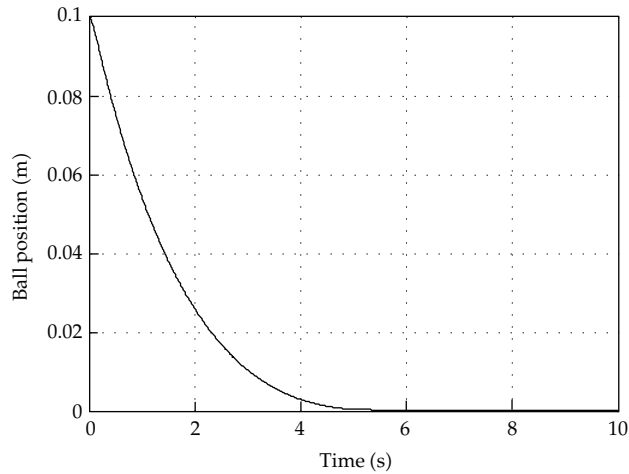


Figure 14: Simulation results for the ball position.

When we apply the multiobjective genetic algorithm, we achieve a Pareto front of the angle error and distance error as demonstrated in Figure 6.

Figure 6 is the chart resulted from multiobjective optimization which all the presented points are nondominated to each other. Each point in this chart is a representative of a vector of selective parameters which if we choose it for the decoupled sliding-mode controller, the analysis tends to objective functions corresponding to that point of chart. The design variables and objective functions of the optimum design points A, B, and C are presented in Table 2.

Achieving several solutions, all of which are considered optimum is a unique property of multiobjective optimization. Designer in facing to Pareto charts, among several different optimum points can choose a suitable multisided design point, easily. According to the Pareto chart, we applied point C for simulation, as shown in Figures 7, 8, 9, 10, and 11.

The simulation results (Figures 7, 8, 9, 10, and 11) show that the pole and the cart can be stabilized to the equilibrium point.

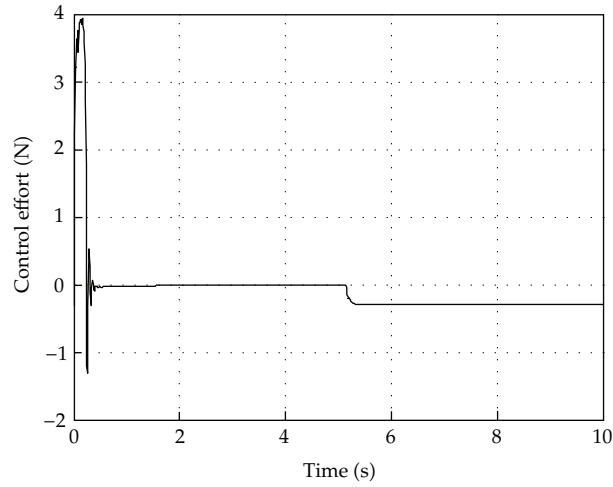


Figure 15: Simulation results for the control action.

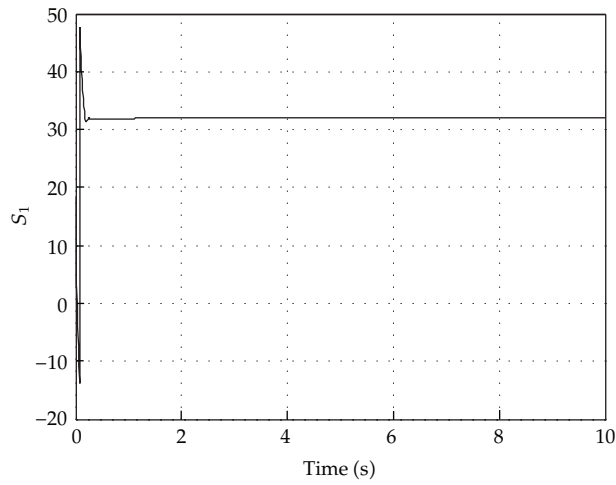


Figure 16: Sliding surface $s_1(x)$.

The numerical results show that the control action is bounded between -15 and 10 (N), and sliding surface $s_2(x)$ reaches to zero during the simulation.

11. Simulation and Results for the Ball and Beam System

The initial values of the ball and beam system are considered in the following form:

$$x_1(0) = 0.1 \text{ m}, \quad x_2(0) = \frac{\pi}{3} \text{ rad}, \quad x_3(0) = 0, \quad x_4(0) = 0. \quad (11.1)$$

The system parameters and constants used in the simulation are given in Table 3.

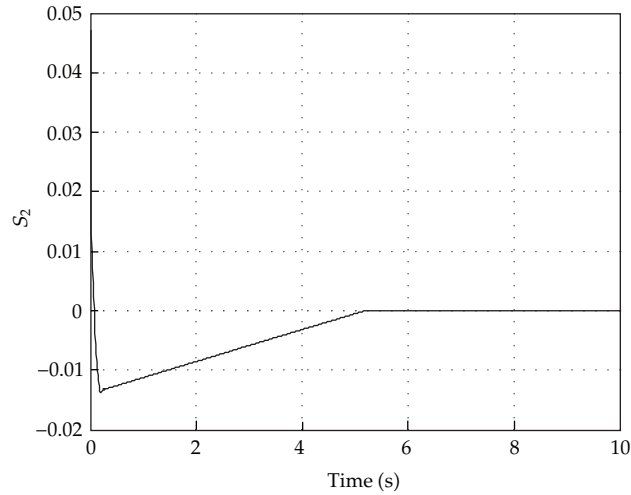


Figure 17: Sliding surface $s_2(x)$.

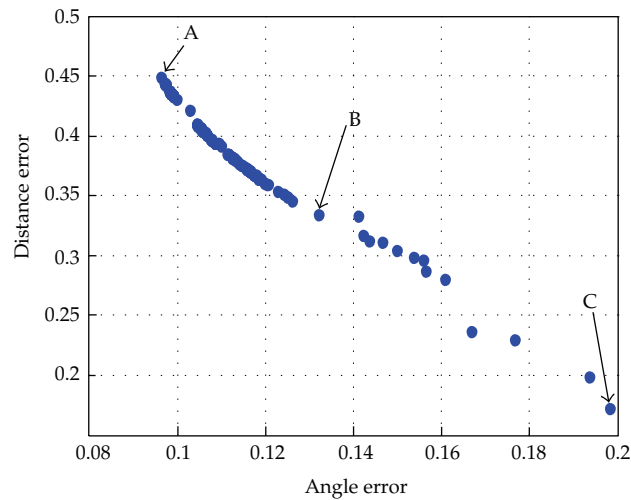


Figure 18: Pareto front of the angle error and distance error for the seesaw system.

When the multiobjective genetic algorithm is applied, a Pareto front of the angle error and distance error would be achieved (Figure 12).

Figure 12 shows the Pareto front obtained from the modified NSGAI algorithm in an arbitrary run for the ball and beam system. In this figure, points A and C stand for the best distance error and angle error, respectively. Furthermore, point B could be a trade-off optimum choice when considering minimum values of both angle error and distance error. Table 4 illustrates the design variables and objective functions corresponding to the optimum design points A, B, and C.

The time responses of the ball and beam system related to point B are shown in Figures 13, 14, 15, 16, and 17. These figures demonstrate that the ball and beam system can be stabilized to the equilibrium point.

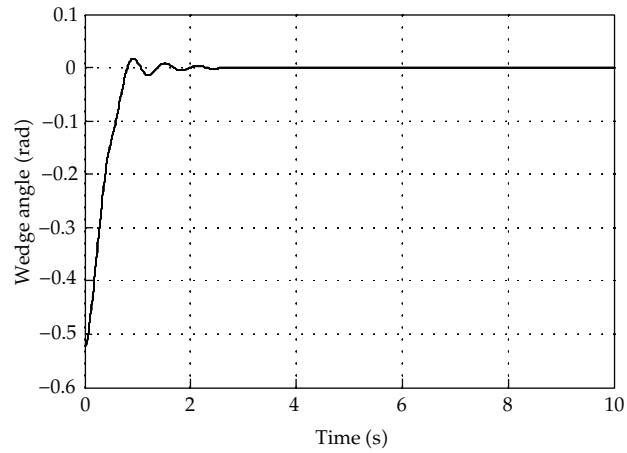


Figure 19: Simulation results for the wedge angle.

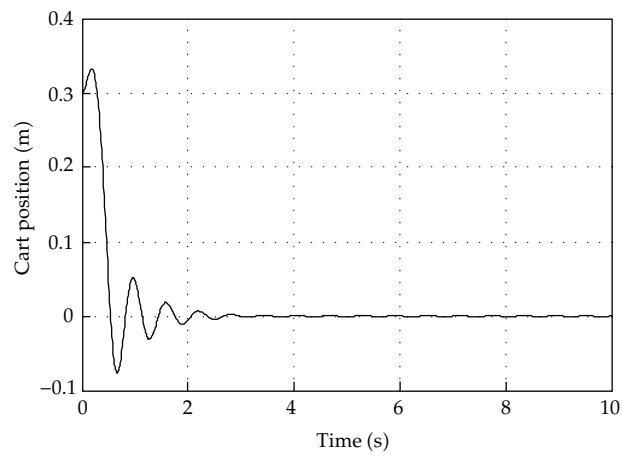


Figure 20: Simulation results for the cart position.

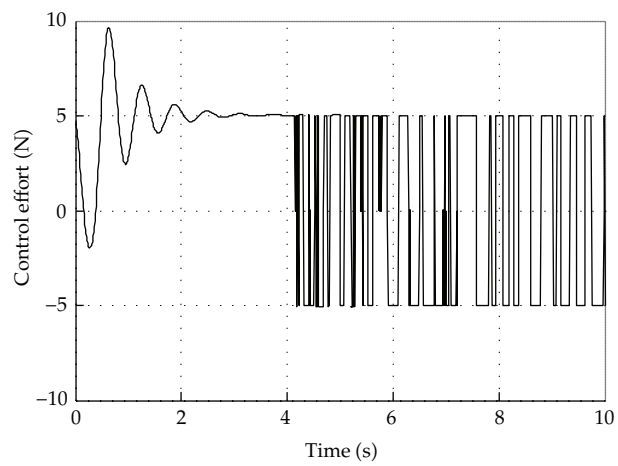


Figure 21: Simulation results for the control action.

Table 2: Comparison among points A, B, and C for Figure 6.

Point	G_{f_1}	G_{s_1}	λ_1	G_{f_2}	G_{s_2}	λ_2	Angle error	Distance error
A	9.9294	9.8941	1.4941	0.9608	0.1851	0.1459	0.3699	22.6927
B	9.9294	9.8941	1.4941	0.9608	0.1851	0.2987	0.4271	12.2111
C	9.8941	9.8941	1.4941	1.0000	0.1772	0.4946	0.5065	7.2827

Table 3: Ball and beam system parameters.

The mass of the ball	m	0.05
The inertia moment of the beam	J	0.0833
The gravity acceleration	g	9.81

Table 4: Comparison among points A, B, and C for Figure 12.

Point	G_{f_1}	G_{s_1}	λ_1	G_{f_2}	G_{s_2}	λ_2	Angle error	Distance error
A	29.8443	1.5451	48.4345	0.6823	46.8690	0.1074	0.0824	0.6174
B	0.2957	0.4158	48.2388	0.6647	41.5855	0.4725	0.0944	0.1402
C	9.8843	1.0040	41.5855	0.6612	2.0569	1.5008	0.1525	0.0225

Table 5: Seesaw system parameters.

The mass of the cart	m	0.46
The mass of the wedge	M	1.52
The height of the wedge	r_1	0.148
The height of center of mass	r_2	0.123
The inertia moment of the wedge	J	0.044
The friction coefficient of the wedge	f_p	0.3
The friction coefficient of the cart	T_c	0.7
The gravity acceleration	g	9.8

Table 6: Comparison among points A, B, and C for Figure 18.

Point	G_{f_1}	G_{s_1}	λ_1	G_{f_2}	G_{s_2}	λ_2	Angle error	Distance error
A	9.9612	7.6706	9.7671	0.0088	0.0060	0.0245	0.0963	0.4489
B	9.9612	0.2165	5.1082	0.0088	0.0557	0.0010	0.1319	0.3345
C	5.0306	0.29416	8.75766	0.60046	0.09306	0.9491	0.1984	0.1729

Furthermore, the simulation shows that the control action is bounded between -1.2 and 4 (N), and sliding surface $s_2(x)$ reaches to zero during simulation.

12. Simulation and Results for the Seesaw System

In this section, the simulation results for seesaw system are investigated. The initial values of this system are described by the following equations:

$$x_1(0) = 0.3 \text{ m}, \quad x_2(0) = -\frac{\pi}{6} \text{ rad}, \quad x_3(0) = 0, \quad x_4(0) = 0. \quad (12.1)$$

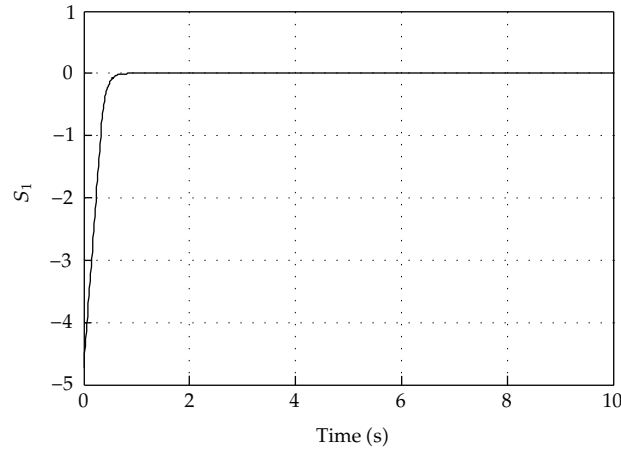


Figure 22: Sliding surface $s_1(x)$.

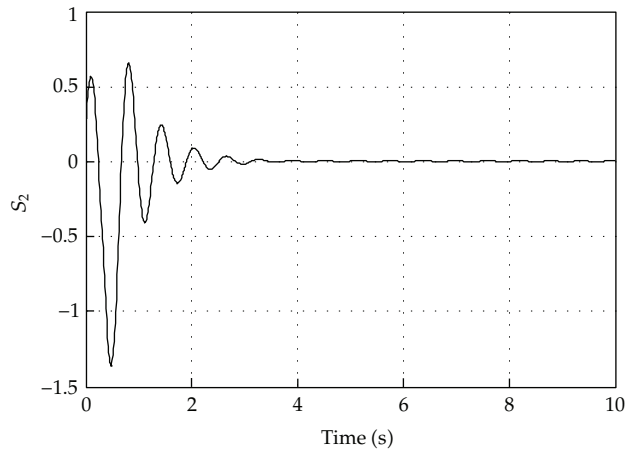


Figure 23: Sliding surface $s_2(x)$.

The system parameters used in the simulation are given in Table 5.

Figure 18 demonstrates a Pareto front of two objective functions (angle error and distance error) which is achieved of the multiobjective genetic algorithm (e.g. modified NSGAI).

It is clear that all points in Figure 18 are nondominated to each other, and each point in this chart is a representative of a vector of selective parameters for the decoupled sliding mode controller. Moreover, choosing a better value for any objective function in the Pareto front would cause a worse value for another objective function. Here, point B has been chosen from Figure 18 to design an optimum decoupled sliding mode controller (Figures 19, 20, 21, 22, and 23). Design variables and objective functions related to the optimum design points A, B, and C are detailed in Table 6.

The simulations (Figures 19, 20, 21, 22, and 23) shows that the seesaw system is stabilized to the equilibrium point after 3 seconds, and the control effort is bounded between -5 and 10 (N).

13. Conclusion

This paper proposes the decoupled sliding-mode technique for stabilising the coupled nonlinear systems while the multiobjective genetic algorithm is employed in order to optimize two objective functions. This method is a universal design method and suitable to various kinds of control objects. Usage this method includes two steps. The first step is to design the decoupled sliding-mode controller for the nonlinear system. The second step is to apply the multiobjective optimization tool to search the definition space of decision variables and to return the optimum answers in the Pareto form. The simulation results on three different and typical control systems show good control and robust performance of the proposed strategy.

References

- [1] J. J. E. Slotine and W. Li, *Applied Nonlinear Control*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1991.
- [2] Z. Gao and S. X. Ding, "Actuator fault robust estimation and fault-tolerant control for a class of nonlinear descriptor systems," *Automatica*, vol. 43, no. 5, pp. 912–920, 2007.
- [3] Z. Gao, X. Shi, and S. X. Ding, "Fuzzy state/disturbance observer design for T-S fuzzy systems with application to sensor fault estimation," *IEEE Transactions on Systems, Man, and Cybernetics B*, vol. 38, no. 3, pp. 875–880, 2008.
- [4] M. J. Mahmoodabadi, A. Bagheri, S. Arabani Mostaghim, and M. Bisheban, "Simulation of stability using Java application for Pareto design of controllers based on a new multi-objective particle swarm optimization," *Mathematical and Computer Modelling*, vol. 54, no. 5-6, pp. 1584–1607, 2011.
- [5] J. C. Lo and Y. H. Kuo, "Decoupled fuzzy sliding-mode control," *IEEE Transactions on Fuzzy Systems*, vol. 6, no. 3, pp. 426–435, 1998.
- [6] A. Bagheri and J. J. Moghaddam, "Decoupled adaptive neuro-fuzzy (DANF) sliding mode control system for a Lorenz chaotic problem," *Expert Systems with Applications*, vol. 36, no. 3, pp. 6062–6068, 2009.
- [7] N. H. Moin, A. S. I. Zinober, and P. J. Harley, "Sliding mode control design using genetic algorithms," in *Proceedings of the 1st IEE/IEEE International Conference on Genetic Algorithms in Engineering Systems: Innovations and Applications (GALESIA '95)*, vol. 414, pp. 238–244, September 1995.
- [8] C. C. Wong and S. Y. Chang, "Parameter selection in the sliding mode control design using genetic algorithms," *Tamkang Journal of Science and Engineering*, vol. 1, no. 2, pp. 115–122, 1998.
- [9] P. C. Chen, C. W. Chen, and W. L. Chiang, "GA-based fuzzy sliding mode controller for nonlinear systems," *Mathematical Problems in Engineering*, vol. 2008, Article ID 325859, 16 pages, 2008.
- [10] J. Javadi-Moghaddam and A. Bagheri, "An adaptive neuro-fuzzy sliding mode based genetic algorithm control system for under water remotely operated vehicle," *Expert Systems with Applications*, vol. 37, no. 1, pp. 647–660, 2010.
- [11] H. K. Khalil, *Nonlinear Systems*, MacMillan, New York, NY, USA, 1992.
- [12] N. Yagiz and Y. Hacioglu, "Robust control of a spatial robot using fuzzy sliding modes," *Mathematical and Computer Modelling*, vol. 49, no. 1-2, pp. 114–127, 2009.
- [13] W. S. Lin and C. S. Chen, "Robust adaptive sliding mode control using fuzzy modelling for a class of uncertain MIMO nonlinear systems," *IEE Proceedings: Control Theory and Applications*, vol. 149, no. 3, pp. 193–202, 2002.
- [14] J. Jing and Q. H. Wuan, "Intelligent sliding mode control algorithm for position tracking servo system," *International Journal of Information Technology*, vol. 12, no. 7, pp. 57–62, 2006.
- [15] V. I. Utkin and H. C. Chang, "Sliding mode control on electro-mechanical systems," *Mathematical Problems in Engineering*, vol. 8, no. 4-5, pp. 451–473, 2002.
- [16] N. F. Al-Muthairi and M. Zribi, "Sliding mode control of a magnetic levitation system," *Mathematical Problems in Engineering*, vol. 2004, no. 2, pp. 93–107, 2004.
- [17] Z. L. Wan, Y. Y. Hou, T. L. Liao, and J. J. Yan, "Partial finite-time synchronization of switched stochastic Chua's circuits via sliding-mode control," *Mathematical Problems in Engineering*, vol. 2011, Article ID 162490, 13 pages, 2011.
- [18] C. Pukdeboon, "Optimal sliding mode controllers for attitude stabilization of flexible spacecraft," *Mathematical Problems in Engineering*, vol. 2011, Article ID 863092, 20 pages, 2011.

- [19] M. Dotoli, P. Lino, and B. Turchiano, "A decoupled fuzzy sliding mode approach to swing-up and stabilize an inverted pendulum, The CSD03," in *Proceedings of the 2nd IFAC Conference on Control Systems Design*, pp. 113–120, Bratislava, Slovak Republic, 2003.
- [20] J. S. Arora, *Introduction to Optimum Design*, McGraw-Hill, New York, NY, USA, 1989.
- [21] S. S. Rao, *Engineering Optimization: Theory and Practice*, Wiley, New York, NY, USA, 1996.
- [22] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, Mass, USA, 1989.
- [23] T. Back, D. B. Fogel, and Z. Michalewicz, *Handbook of Evolutionary Computation*, Institute of Physics Publishing, New York, NY, USA, Oxford University Press, Oxford, UK, 1997.
- [24] G. Renner and A. Ekárt, "Genetic algorithms in computer aided design," *Computer Aided Design*, vol. 35, no. 8, pp. 709–726, 2003.
- [25] P. J. Fleming and R. C. Purshouse, "Evolutionary algorithms in control systems engineering: a survey," *Control Engineering Practice*, vol. 10, no. 11, pp. 1223–1241, 2002.
- [26] C. M. Fonseca and P. J. Fleming, "Multiobjective optimal controller design with genetic algorithms," in *Proceedings of the International Conference on Control*, vol. 1, pp. 745–749, March 1994.
- [27] G. Sánchez, M. Villasana, and M. Strefezza, "Multi-objective pole placement with evolutionary algorithms," *Lecture Notes in Computer Science*, vol. 4403, pp. 417–427, 2007.
- [28] E. Alfaro-Cid, E. W. McGookin, D. J. Murray-Smith, and T. I. Fossen, "Genetic algorithms optimisation of decoupled Sliding Mode controllers: simulated and real results," *Control Engineering Practice*, vol. 13, no. 6, pp. 739–748, 2005.
- [29] N. Srinivas and K. Deb, "Multiobjective optimization using nondominated sorting in genetic algorithms," *Evolutionary Computation*, vol. 2, no. 3, pp. 221–248, 1994.
- [30] C. M. Fonseca and P. J. Fleming, "Genetic algorithms for multi-objective optimization: formulation, discussion and generalization," in *Proceedings of the 5th International Conference On genetic Algorithms*, S. Forrest, Ed., pp. 416–423, Morgan Kaufmann, San Mateo, Calif, USA, 1993.
- [31] C. A. Coello and A. D. Christiansen, "Multiobjective optimization of trusses using genetic algorithms," *Computers and Structures*, vol. 75, no. 6, pp. 647–660, 2000.
- [32] C. A. Coello Coello, D. A. Van Veldhuizen, and G. B. Lamont, *Evolutionary Algorithms for Solving Multi-Objective Problems*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.
- [33] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.
- [34] A. Toffolo and E. Benini, "Genetic diversity as an objective in multi-objective evolutionary algorithms," *Evolutionary Computation*, vol. 11, no. 2, pp. 151–167, 2003.
- [35] C. A. Coello Coello and R. L. Becerra, "Evolutionary multiobjective optimization using a cultural algorithm," in *Proceedings of the IEEE Swarm Intelligence Symposium*, pp. 6–13, IEEE Service Center, Piscataway, NJ, USA, 2003.
- [36] K. Atashkari, N. Nariman-Zadeh, A. Pilechi, A. Jamali, and X. Yao, "Thermodynamic Pareto optimization of turbojet engines using multi-objective genetic algorithms," *International Journal of Thermal Sciences*, vol. 44, no. 11, pp. 1061–1071, 2005.

Research Article

Synchronization of Coupled Networks with Mixed Delays by Intermittent Control

Kun Yuan,¹ Jinde Cao,² and Shumin Fei¹

¹ School of Automation, Southeast University, Nanjing 210096, China

² Department of Mathematics, Southeast University, Nanjing 210096, China

Correspondence should be addressed to Kun Yuan, kyuan@seu.edu.cn

Received 3 May 2012; Revised 26 June 2012; Accepted 26 June 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 Kun Yuan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The synchronization of coupled networks with mixed delays is investigated by employing Lyapunov functional method and intermittent control. A sufficient condition is derived to ensure the global synchronization of coupled networks, which is controlled by the designed intermittent controller. Finally, a numerical simulation is constructed to justify the theoretical analysis.

1. Introduction

Various large-scale and complicated systems can be modelled by complex networks, such as the Internet, genetic networks, ecosystems, electrical power grids, and the social networks. A complex network is a large set of interconnected nodes, which can be described by the graph with the nodes representing individuals in the graph and the edges representing the connections among them. The most remarkable recent advances in study of complex networks are the developments of the small-world network model [1] and scale-free network model [2], which have been shown to be very closer to most real-world networks as compared with the random-graph model [3, 4]. Thereafter, small-world and scale-free networks have been extensively investigated.

The dynamical behaviors of complex networks have become a focal topic of great interest, particularly the synchronization phenomena, which is observed in natural, social, physical, and biological systems and has been widely applied in a variety of fields, such as secure communication, image processing, and harmonic oscillation generation. It is noted that the dynamical behavior of a complex network is determined not only by the dynamical

rules governing the isolated nodes, referred to as self-dynamics, but also by information flow along the edges, which depends on the topology of the network. Synchronization in an array of linearly coupled dynamical systems was investigated in [5]. Later, many results on local, global, and partial synchronization in various coupled systems have also been obtained in [6–15]. As a special case of coupled systems, coupled neural networks with time delay have also been found to exhibit complex behaviors. The estimation and diagnosis for time delay systems are discussed in [16, 17], and synchronization for coupled neural networks with time delay has been investigated by many researchers, for example, [8–15].

In the case that the whole network cannot synchronize by itself, some controllers should be designed and applied to force the network to synchronize. Recently, another interesting intermittent control was introduced and studied, that is, the control time is periodic, and in any period the time is composed of work time and rest time. It is a straightforward engineering approach to process control of any typelan approach that has been used for a variety of purposes in such engineering fields as manufacturing, transportation, and communication. Intermittent control has been introduced to control nonlinear dynamical systems [18] and has been studied in [19–26]. In [18], the authors investigated numerically chaos synchronization under the condition that the interacting systems, that is, master and slave systems are coupled intermittently. In [19, 20], the stabilization problems of chaotic systems with or without delays by periodically intermittent control were discussed. Huang et al. discussed the synchronization of coupled chaotic systems with delay by using intermittent state feedback in [21]. In [25], the authors synchronize coupled networks using pinning control and intermittent control. In [26], cluster synchronization was studied for coupled networks without time delay using adaptive intermittent control.

Another type of time delays, namely, distributed delays, has begun to receive research attention. The main reason is that a neural network usually has a spatial nature due to the presence of an amount of parallel pathways of a variety of axon sizes and lengths, and it is desirable to model them by introducing continuously distributed delays over a certain duration of time, such that the distant past has less influence compared to the recent behavior of the state [27]. Therefore, both discrete and distributed time delays should be taken into account [28–33]. Although synchronization has been investigated under intermittent control, [25, 26], there is still no theoretical result of synchronization for coupled networks with mixed delay.

Motivated by the above discussion, the intermittent controller will be designed to achieve the synchronization for coupled networks with mixed delay. The rest of the paper is organized as follows. In Section 2, some preliminary definitions and lemmas are briefly outlined. Some synchronization criteria are given and intermittent controller are designed in Section 3. An illustrative simulation is given to verify the theoretical analysis in Section 4. Conclusions are finally drawn.

Notations

\mathbb{R}^n is the n -dimensional Euclidean space; $\mathbb{R}^{m \times n}$ denotes the set of $m \times n$ real matrix. I is the identity matrix with appropriate dimension, and the superscript “ T ” represents the transpose. Matrix dimensions, if not explicitly stated, are assumed to be compatible for algebraic operations.

2. Model Description and Preliminaries

Consider a dynamical network consisting of N identical and diffusively coupled nodes, with each node being an n -dimensional delayed neural network. The state equations of the network are

$$\begin{aligned} \dot{x}_i(t) = & -Dx_i(t) + Af(x_i(t)) + Bg(x_i(t-\tau)) + C \int_{t-\tau}^t h(x_i(v))dv \\ & + I(t) + \sum_{j=1, j \neq i}^N G_{ij}\Gamma(x_j(t) - x_i(t)), \end{aligned} \quad (2.1)$$

where $x_i(t) = (x_{i1}(t), x_{i2}(t), \dots, x_{in}(t))^T \in \mathbb{R}^n$ is the state vector of the i th node; $D = \text{diag}(d_1, d_2, \dots, d_n) > 0$ denotes the rate with which the cell i resets its potential to the resting state when isolated from other cells and inputs; $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times n}$, and $C \in \mathbb{R}^{n \times n}$ represent the connection weight matrix, the discretely delayed connection weight matrix, and the distributively delayed connection weights, respectively; $f(x_i(\cdot)) = [f_1(x_{i1}(\cdot)), f_2(x_{i2}(\cdot)), \dots, f_n(x_{in}(\cdot))]^T \in \mathbb{R}^n$, $g(x_i(\cdot)) = [g_1(x_{i1}(\cdot)), g_2(x_{i2}(\cdot)), \dots, g_n(x_{in}(\cdot))]^T \in \mathbb{R}^n$ and $h(x_i(\cdot)) = [h_1(x_{i1}(\cdot)), h_2(x_{i2}(\cdot)), \dots, h_n(x_{in}(\cdot))]^T \in \mathbb{R}^n$ are activation functions; $I(t)$ is the input vector of each node; $\Gamma \in \mathbb{R}^{n \times n}$ is the inner coupling matrix; $G = (G_{ij})_{N \times N}$ is the coupling configuration matrix representing the topological structure of the network, where G_{ij} is defined as follows: if there exists a connection between node i and node j , $G_{ij} > 0$, otherwise $G_{ij} = 0$ ($j \neq i$), and the diagonal elements of matrix G are defined by

$$G_{ii} = - \sum_{j=1, j \neq i}^N G_{ij}, \quad (2.2)$$

which ensures the diffusion that $\sum_{j=1}^N G_{ij} = 0$. Equivalently, network (2.1) can be rewritten in a form as follows:

$$\begin{aligned} \dot{x}_i(t) = & -Dx_i(t) + Af(x_i(t)) + Bg(x_i(t-\tau)) + C \int_{t-\tau}^t h(x_i(v))dv \\ & + I(t) + \sum_{j=1}^N G_{ij}\Gamma x_j(t), \quad i = 1, 2, \dots, N. \end{aligned} \quad (2.3)$$

Suppose that the coupled network (2.3) is connected in the sense that there are no isolated clusters, then the coupling matrix G is irreducible.

Note that a solution to an isolated node satisfies

$$\frac{ds(t)}{dt} = -Ds(t) + Af(s(t)) + Bg(s(t-\tau)) + C \int_{t-\tau}^t h(s(v))dv + I(t). \quad (2.4)$$

To realize the synchronization of network (2.3), the intermittent strategy is selected, and the controlled network can be described by

$$\begin{aligned} \dot{x}_i(t) = & -Dx_i(t) + Af(x_i(t)) + Bg(x_i(t-\tau)) + C \int_{t-\tau}^t h(x_i(v))dv \\ & + I(t) + \sum_{j=1}^N G_{ij}\Gamma x_j(t) + u_i, \quad i = 1, 2, \dots, N, \end{aligned} \quad (2.5)$$

where

$$u_i = -k_i(t)(x_i(t) - s(t)), \quad (2.6)$$

$k(t)$ is the intermittent linear state feedback control gain defined as follows:

$$k_i(t) = \begin{cases} k_i, & n\omega \leq t \leq n\omega + \delta, \\ 0, & n\omega + \delta < t \leq (n+1)\omega, \end{cases} \quad (2.7)$$

where $k_i \in R$ is a constant control gain, $\omega > 0$ is the control period, and $\delta > 0$ is called the control width. In this paper, our goal is to design suitable δ , ω , and k_i such that network (2.5) synchronize with respect to the isolated node $s(t)$. Denote $e_i(t) = x_i(t) - s(t)$, then the following error dynamical system is obtained:

$$\begin{aligned} \dot{e}_i(t) = & -De_i(t) + A[f(x_i(t)) - f(s(t))] + B[g(x_i(t-\tau)) - g(s(t-\tau))] \\ & + C \int_{t-\tau}^t [h(x_i(v)) - h(s(v))]dv + \sum_{j=1}^N G_{ij}\Gamma e_j(t) - k_i e_i(t), \quad n\omega \leq t \leq n\omega + \delta, \\ \dot{e}_i(t) = & -De_i(t) + A[f(x_i(t)) - f(s(t))] + B[g(x_i(t-\tau)) - g(s(t-\tau))] \\ & + C \int_{t-\tau}^t [h(x_i(v)) - h(s(v))]dv + \sum_{j=1}^N G_{ij}\Gamma e_j(t), \quad n\omega + \delta < t \leq (n+1)\omega. \end{aligned} \quad (2.8)$$

(H) We assume that f , g , and h are Lipschitz continuous functions; there exist positive constants L_f , L_g and L_h such that, for all $x, y \in \mathbb{R}^m$,

$$\begin{aligned} \|f(x) - f(y)\| & \leq L_f \|x - y\|, \\ \|g(x) - g(y)\| & \leq L_g \|x - y\|, \\ \|h(x) - h(y)\| & \leq L_h \|x - y\|. \end{aligned} \quad (2.9)$$

Definition 2.1. For any positive integers p, q, r, s , we define the Kronecker product of two matrices $A \in \mathbb{R}^{p \times q}, B \in \mathbb{R}^{r \times s}$ as follows:

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1q}B \\ \vdots & \ddots & \vdots \\ a_{p1}B & \cdots & a_{pq}B \end{bmatrix} \in \mathbb{R}^{pr \times qs}. \quad (2.10)$$

Lemma 2.2. By the definition of Kronecker product, the following properties hold:

- (1) $(A \otimes B)^T = A^T \otimes B^T$;
- (2) $(\alpha A) \otimes B = A \otimes (\alpha B)$, where α is a real number;
- (3) $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$.

Lemma 2.3. For any vectors $x, y \in \mathbb{R}^m$, and positive-definite matrix $Q \in \mathbb{R}^{m \times m}$, the following matrix inequality holds:

$$2x^T y \leq x^T Q x + y^T Q^{-1} y. \quad (2.11)$$

Lemma 2.4 (Jensen's inequality [34]). For any constant matrix $V \in \mathbb{R}^{m \times m}$, $V > 0$, scalar $0 < r(t) < r$, vector function $v : [0, r] \rightarrow \mathbb{R}^m$ such that the integrations concerned are well defined, then

$$r(t) \int_0^{r(t)} v^T(s) V v(s) ds \geq \left(\int_0^{r(t)} v(s) ds \right)^T V \left(\int_0^{r(t)} v(s) ds \right). \quad (2.12)$$

Lemma 2.5 (Halanay inequality [35]). Let $V : [\mu - \tau, \infty) \rightarrow [0, \infty)$ be a continuous function such that

$$\frac{dV(t)}{dt} \leq -aV(t) + b \max V_t \quad (2.13)$$

is satisfied for $t \geq \mu$. If $a > b > 0$, then

$$V(t) \leq [\max V_\mu] \exp\{-r(t - \mu)\}, \quad t \geq \mu, \quad (2.14)$$

where $\max V_t = \sup_{t-\tau \leq \theta \leq t} V(\theta)$, and $r > 0$ is the smallest real root of the following equation:

$$-r = -a + b \exp\{r\tau\}. \quad (2.15)$$

Lemma 2.6 (see [19]). Let $V : [\mu - \tau, \infty) \rightarrow [0, \infty)$ be a continuous function, such that

$$\frac{dV(t)}{dt} \leq aV(t) + b \max V_t \quad (2.16)$$

is satisfied for $t \geq \mu$. If $a > 0$, $b > 0$, then

$$V(t) \leq \max V_t \leq [\max V_\mu] \exp\{(a+b)(t-\mu)\}, \quad t \geq \mu, \quad (2.17)$$

where $\max V_t = \sup_{t-\tau \leq \theta \leq t} V(\theta)$.

3. Criteria for Synchronization

Theorem 3.1. Suppose that assumption (H) holds. The controlled coupled network (2.5) globally synchronizes to (2.4) if there are positive definite matrix P , positive constants $\alpha, \beta, \gamma, a_1, a_2, b_1, b_2$ such that the following conditions hold:

$$(a) \quad I_N \otimes (Q + a_1 P) + G \otimes \Gamma - K \otimes I_n \leq 0,$$

$$(b) \quad I_N \otimes (Q - a_2 P) + G \otimes \Gamma \leq 0,$$

$$(c) \quad \beta^{-1} L_g^2 I_n - b_1 P \leq 0,$$

$$(d) \quad \gamma^{-1} L_h^2 I_n - b_2 P \leq 0,$$

$$(e) \quad a_1 > b = b_1 + \tau^2 b_2$$

$$(f) \quad \rho = r(\delta - \tau) - (a_2 + b)(\omega - \delta) > 0,$$

where $Q = -PD + (\alpha/4)PAA^T P + \alpha^{-1}L_f^2 I_n + (\beta/4)PBB^T P + (\gamma/4)PCC^T P + a_1 P$, $K = \text{diag}(k_1, k_2, \dots, k_N)$ and r is the positive solution of $-r = -a_1 + be^{r\tau}$.

Proof. Consider the following Lyapunov function:

$$V(t) = \frac{1}{2} \sum_{i=1}^N e_i(t)^T P e_i(t) = \frac{1}{2} e^T (I_N \otimes P) e(t), \quad (3.1)$$

where $e(t) = [e_1^T(t), e_2^T(t), \dots, e_N^T(t)]^T$. Calculate the derivative $V(t)$ with respect to time t along the trajectory of error system (2.8), and estimate it.

For $lw \leq t \leq lw + \delta$, using Lemma 2.3 and assumption, we have the following:

$$\begin{aligned}
\dot{V}(t) &= \sum_{i=1}^N e_i^T(t) P \dot{e}_i(t) \\
&= \sum_{i=1}^N e_i^T(t) P \left[-De_i(t) + A(f(x_i(t)) - f(s(t))) + B(g(x_i(t-\tau)) - g(s(t-\tau))) \right. \\
&\quad \left. + C \left(\int_{t-\tau}^t h(x_i(v)) dv - \int_{t-\tau}^t h(s(v)) dv \right) + \sum_{j=1}^N G_{ij} \Gamma e_j(t) - k_i e_i(t) \right] \\
&\leq \sum_{i=1}^N \left[-e_i^T(t) P D e_i(t) + \frac{\alpha}{4} e_i^T(t) P A A^T P e_i(t) + \alpha^{-1} \|f(x_i(t)) - f(s(t))\|^2 \right. \\
&\quad \left. + \frac{\beta}{4} e_i^T(t) P B B^T P e_i(t) + \beta^{-1} \|g(x_i(t-\tau)) - g(s(t-\tau))\|^2 \right. \\
&\quad \left. + \frac{\gamma}{4} e_i^T(t) P C C^T P e_i(t) + \gamma^{-1} \left\| \int_{t-\tau}^t h(x_i(v)) dv - \int_{t-\tau}^t h(s(v)) dv \right\|^2 \right] \quad (3.2) \\
&\quad + \sum_{i=1}^N \sum_{j=1}^N G_{ij} e_i^T(t) \Gamma e_j(t) - \sum_{i=1}^N k_i e_i^T(t) e_i(t) \\
&\leq e^T(t) I_N \otimes \left(-PD + \frac{\alpha}{4} P A A^T P + \alpha^{-1} L_f^2 I_n \right. \\
&\quad \left. + \frac{\beta}{4} P B B^T P + \frac{\gamma}{4} P C C^T P + a_1 P \right) e(t) - a_1 e^T(t) (I_N \otimes P) e(t) \\
&\quad + e^T(t-\tau) \left[I_N \otimes (\beta^{-1} L_g^2 I_n - b_1 P) \right] e(t-\tau) + b_1 e^T(t-\tau) (I_N \otimes P) e(t-\tau) \\
&\quad + \left[\int_{t-\tau}^t e(v) dv \right]^T \left[I_N \otimes (\gamma^{-1} L_h^2 I_n - b_2 P) \right] \left[\int_{t-\tau}^t e(v) dv \right] \\
&\quad + b_2 \left[\int_{t-\tau}^t e(v) dv \right]^T (I_N \otimes P) \left[\int_{t-\tau}^t e(v) dv \right] + e^T(t) (G \otimes \Gamma - K \otimes I_n) e(t).
\end{aligned}$$

From Jensen's inequality in Lemma 2.4, we have

$$b_2 \left[\int_{t-\tau}^t e(v) dv \right]^T (I_N \otimes P) \left[\int_{t-\tau}^t e(v) dv \right] \leq \tau b_2 \int_{t-\tau}^t e^T(v) (I_N \otimes P) e(v) dv. \quad (3.3)$$

By condition (a), (c), (d), and (3.3), one has

$$\dot{V}(t) \leq -a_1 V(t) + b \max V_t \quad (3.4)$$

For $l\omega + \delta \leq t \leq (l+1)\omega$, from conditions (b), (c), and (d), one has

$$\dot{V}(t) < a_2 V(t) + bV_t. \quad (3.5)$$

Next, we will prove the error $\|e(t)\| \rightarrow 0$.

From Lemma 2.5 and (3.2), one has

$$V(t) \leq \|V(0)\|_\tau e^{-rt}, \quad \text{for } 0 \leq t \leq \delta, \quad (3.6)$$

where r is the unique positive solution of $-r = -a_1 + be^{r\tau}$.

From Lemma 2.6, one obtain the following:

$$\begin{aligned} V(t) &\leq \|V(\delta)\|_\tau e^{(a_2+b)(t-\delta)} \\ &= \max_{\delta-\tau \leq t \leq \delta} |V(t)| e^{(a_2+b)(t-\delta)} \\ &\leq \|V(0)\|_\tau e^{-r(\delta-\tau)} e^{(a_2+b)(t-\delta)}, \end{aligned} \quad (3.7)$$

for $\delta \leq t \leq \omega$.

Suppose that $\omega - \tau > \delta$, then

$$\begin{aligned} \|V(\omega)\|_\tau &= \max_{\omega-\tau \leq t \leq \omega} |V(t)| \\ &\leq \max_{\omega-\tau \leq t \leq \omega} \left\{ \|V(0)\|_\tau e^{-r(\delta-\tau)} e^{(a_2+b)(t-\delta)} \right\} \\ &= \|V(0)\|_\tau e^{-r(\delta-\tau)} e^{(a_2+b)(\omega-\delta)} \\ &= \|V(0)\|_\tau e^{-\rho}. \end{aligned} \quad (3.8)$$

Using mathematical induction, we can prove, for any positive integer l ,

$$\|V(l\omega)\|_\tau \leq \|V(0)\|_\tau e^{-l\rho}. \quad (3.9)$$

Assume (3.9) holds when $k \leq l$. Now, we prove (3.9) is true when $k = l+1$.

First, we have

$$\|V(l\omega)\|_\tau \leq \|V(0)\|_\tau e^{-l\rho}. \quad (3.10)$$

When $t \in [l\omega, l\omega + \delta]$,

$$V(t) \leq \|V(l\omega)\|_\tau e^{-r(t-l\omega)} \leq \|V(0)\|_\tau e^{-l\rho} e^{-r(t-l\omega)}. \quad (3.11)$$

Thus, for $t \in [l\omega + \delta, (l+1)\omega]$, we have

$$\begin{aligned}
 V(t) &\leq \|V(l\omega + \delta)\|_{\tau} e^{(a_2+b)(t-l\omega-\delta)} \\
 &= \left[\max_{l\omega+\delta-\tau \leq t \leq l\omega+\delta} |V(t)| \right] e^{(a_2+b)(t-l\omega-\delta)} \\
 &\leq \left[\max_{l\omega+\delta-\tau \leq t \leq l\omega+\delta} \|V(0)\|_{\tau} e^{-l\rho} e^{-r(t-l\omega)} \right] \\
 &\quad \times e^{(a_2+b)(t-l\omega-\delta)} \\
 &\leq \|V(0)\|_{\tau} e^{-l\rho} e^{-r(\delta-\tau)} e^{(a_2+b)(t-l\omega-\delta)}, \tag{3.12} \\
 \|V((l+1)\omega)\|_{\tau} &= \max_{(l+1)\omega-\tau \leq t \leq (l+1)\omega} |V(t)| \\
 &\leq \max_{(l+1)\omega-\tau \leq t \leq (l+1)\omega} \left[\|V(0)\|_{\tau} e^{-l\rho} e^{-r(\delta-\tau)} e^{(a_2+b)(t-l\omega-\delta)} \right] \\
 &= \|V(0)\|_{\tau} e^{-l\rho} e^{-r(\delta-\tau)} e^{(a_2+b)(\omega-\delta)} \\
 &= \|V(0)\|_{\tau} e^{-(l+1)\rho}.
 \end{aligned}$$

Thus, (3.9) holds for all positive integers k .

For any $t > 0$, there is $n_0 \geq 0$, such that $n_0\omega \leq t \leq (n_0 + 1)\omega$;

$$\begin{aligned}
 V(t) &\leq \|V(n_0\omega)\|_{\tau} e^{(a_2+b)(t-n_0\omega)} \\
 &\leq \|V(0)\|_{\tau} e^{-n_0\rho} e^{(a_2+b)\omega} \\
 &\leq \|V(0)\|_{\tau} e^{(a_2+b)\omega} e^{\rho} \exp\left(-\frac{\rho}{\omega}t\right). \tag{3.13}
 \end{aligned}$$

Let $M = \|V(0)\|_{\tau} e^{(a_2+b)\omega} e^{\rho}$, one has the following inequality:

$$\lambda_m(P) \|e(t)\|^2 \leq V(t) \leq M \exp\left(-\frac{\rho}{\omega}t\right), \quad \text{for } t \geq 0. \tag{3.14}$$

Obviously,

$$\|e(t)\| \leq \sqrt{\frac{M}{\lambda_m(P)}} \exp\left(-\frac{\rho}{2\omega}t\right), \tag{3.15}$$

which means the coupled networks (2.5) achieve synchronization. The proof is completed. \square

Corollary 3.2. For given control period ω and control duration δ , coupled networks (2.5) achieve synchronization, if the control gain $K = kI_N$ satisfies

$$k > \Phi(r^*), \tag{3.16}$$

where

$$\begin{aligned}\Phi(r) &= M + r + (2L_g^2 + 2\tau^2 L_h^2)e^{r\tau}, \\ r^* &= \frac{\omega - \delta}{\delta - \tau} (M + 2L_g^2 + 2\tau^2 L_h^2).\end{aligned}\tag{3.17}$$

Proof. In Theorem 3.1, let $P = I$, $\beta = \gamma = 1$, $b_1 = 2L_g^2$, $b_2 = 2L_h^2$, obviously, (c) and (d) in Theorem 3.1 hold.

Furthermore, let $M = \lambda_M(Q) + \lambda_M(G) \times \lambda_M(\Gamma)$, where $\lambda_M(\cdot)$ is the maximum of eigenvalue, $a_1 = k - M > 0$ and $a_2 = M$, (a) and (b) in Theorem 3.1 hold. From the above parameters and (f) in Theorem 3.1 hold if $r > r^*$ and r is the positive solution of $-r = -a_1 + (2L_g^2 + 2\tau^2 L_h^2)e^{r\tau}$, that is, $k = r + M + L_f + (2L_g^2 + 2\tau^2 L_h^2)e^{r\tau} = \Phi(r)$. Obviously, $\Phi(r)$ is increasing function.

Therefore, (a)–(f) hold if

$$k > \max[\Phi(r^*), M] = \Phi(r^*).\tag{3.18}$$

□

Remark 3.3. Corollary 3.2 shows us how to determine the control gain in a simple way provided that the control period ω and control duration δ are given.

4. Numerical Example

Consider the following coupled networks:

$$\begin{aligned}\dot{x}_i(t) &= -Dx_i(t) + Af(x_i(t)) + Bg(x_i(t - \tau)) \\ &\quad + C \int_{t-\tau}^t h(x_i(v))dv + \sum_{j=1}^N G_{ij}\Gamma x_j(t), \\ x_i(t) &= \phi(t), \quad -\tau \leq t \leq 0,\end{aligned}\tag{4.1}$$

where $x_i(t) = [x_{i1}(t), x_{i2}(t)]^T$, $i = 1, 2, 3$ are the state variable of the i th neural network. Choose $\tau = 1$, $f(x_i(t)) = g(x_i(t)) = h(x_i(t)) = (3/5)[\tanh(x_{i1}), \tanh(x_{i2})]^T$, and

$$\begin{aligned}D &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A = \begin{bmatrix} 2.0 & -0.1 \\ -5.0 & 4.5 \end{bmatrix}, \\ B &= \begin{bmatrix} -1.5 & -0.1 \\ -0.2 & -2 \end{bmatrix}, \quad C = \begin{bmatrix} -0.2 & -0.1 \\ -1.6 & -3.2 \end{bmatrix}, \quad G = \begin{bmatrix} -4 & 2 & 2 \\ 1 & -2 & 1 \\ 1 & 2 & -3 \end{bmatrix},\end{aligned}\tag{4.2}$$

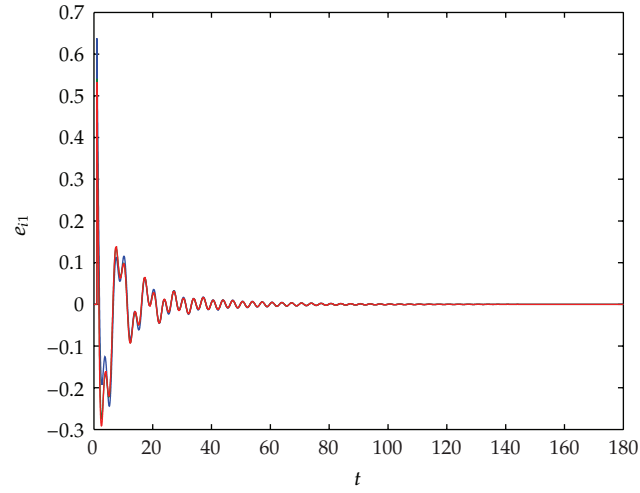


Figure 1: Error state $e_{i1}(t)$.

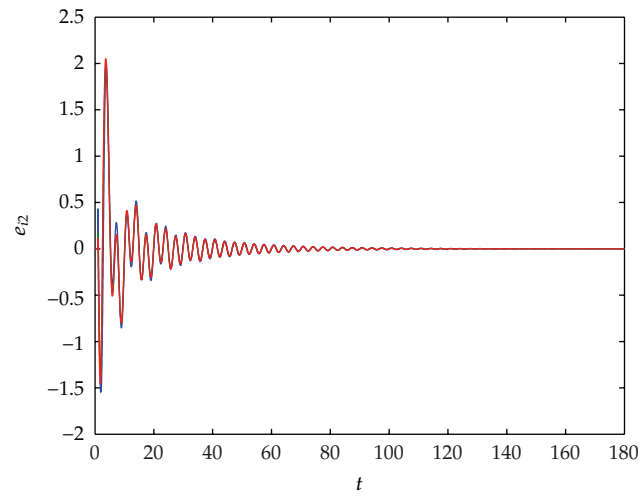


Figure 2: Error state $e_{i2}(t)$.

and input vectors $I = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, and $k_i(t)$ is the intermittent linear state feedback control gain defined as the following:

$$k_i(t) = \begin{cases} k_i, & k\omega \leq t \leq k\omega + \delta, \\ 0, & k\omega + \delta < t \leq (k+1)\omega, \end{cases} \quad (4.3)$$

where the control gain $k_1 = k_2 = k_3 = 0.1$, the control period $\omega = 3$, and the control width $\delta = 1.3$. The above suitable δ , ω and K such that (4.1) synchronize. The synchronize errors are given in Figures 1 and 2.

5. Conclusion

In this paper, synchronization of coupled networks with mixed time delay has been investigated via intermittent control. Some criteria for ensuring coupled networks synchronization have been derived, and some analytical techniques have been proposed to obtain appropriate control period ω , control width δ , and control gain for achieving network synchronization. Finally, the simulation confirmed the effectiveness of the proposed intermittent controller.

Acknowledgments

This work was jointly supported by the National Natural Science Foundation of China under Grant 61004043 and the Specialized Research Fund for the Doctoral Program of Higher Education under Grant 2009092120066.

References

- [1] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [2] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
- [3] P. Erdős and A. Rényi, "On random graphs. I," *Publicationes Mathematicae Debrecen*, vol. 6, pp. 290–297, 1959.
- [4] P. Erdős and A. Rényi, "On the evolution of random graphs," *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, vol. 5, pp. 17–61, 1960.
- [5] C. W. Wu and L. O. Chua, "Synchronization in an array of linearly coupled dynamical systems," *IEEE Transactions on Circuits and Systems*, vol. 42, no. 8, pp. 430–447, 1995.
- [6] C. W. Wu, "Synchronization in coupled arrays of chaotic oscillators with nonreciprocal coupling," *IEEE Transactions on Circuits and Systems*, vol. 50, no. 2, pp. 294–297, 2003.
- [7] E. M. Izhikevich and F. C. Hoppensteadt, "Slowly coupled oscillators: phase dynamics and synchronization," *SIAM Journal on Applied Mathematics*, vol. 63, no. 6, pp. 1935–1953, 2003.
- [8] W. Lu and T. Chen, "Synchronization of coupled connected neural networks with delays," *IEEE Transactions on Circuits and Systems*, vol. 51, no. 12, pp. 2491–2503, 2004.
- [9] G. Chen, J. Zhou, and Z. Liu, "Global synchronization of coupled delayed neural networks and applications to chaotic CNN models," *International Journal of Bifurcation and Chaos in Applied Sciences and Engineering*, vol. 14, no. 7, pp. 2229–2240, 2004.
- [10] W. Wang and J. Cao, "Synchronization in an array of linearly coupled networks with time-varying delay," *Physica A*, vol. 366, pp. 197–211, 2006.
- [11] J. Cao, G. Chen, and P. Li, "Global synchronization in an array of delayed neural networks with hybrid coupling," *IEEE Transactions on Systems, Man, and Cybernetics B*, vol. 38, no. 2, pp. 488–498, 2008.
- [12] W. Yu, J. Cao, and J. Lü, "Global synchronization of linearly hybrid coupled networks with time-varying delay," *SIAM Journal on Applied Dynamical Systems*, vol. 7, no. 1, pp. 108–133, 2008.
- [13] W. Yu, J. Cao, G. Chen, J. Lü, J. Han, and W. Wei, "Local synchronization of a complex network model," *IEEE Transactions on Systems, Man, and Cybernetics B*, vol. 39, no. 1, pp. 230–241, 2009.
- [14] K. Yuan, "Robust synchronization in arrays of coupled networks with delay and mixed coupling," *Neurocomputing*, vol. 72, no. 4–6, pp. 1026–1031, 2009.
- [15] J. Cao, P. Li, and W. Wang, "Global synchronization in arrays of delayed neural networks with constant and delayed coupling," *Physics Letters A*, vol. 353, no. 4, pp. 318–325, 2006.
- [16] Z. Gao and S. X. Ding, "State and disturbance estimator for time-delay systems with application to fault estimation and signal compensation," *IEEE Transactions on Signal Processing*, vol. 55, no. 12, pp. 5541–5551, 2007.
- [17] Z. Gao, T. Breikin, and H. Wang, "Reliable observer-based control against sensor failures for systems with time delays in both state and input," *IEEE Transactions on Systems, Man, and Cybernetics A*, vol. 38, no. 5, pp. 1018–1029, 2008.
- [18] M. Zochowski, "Intermittent dynamical control," *Physica D*, vol. 145, no. 3–4, pp. 181–190, 2000.

- [19] C. Li, X. Liao, and T. Huang, "Exponential stabilization of chaotic systems with delay by periodically intermittent control," *Chaos*, vol. 17, no. 1, article 013103, 2007.
- [20] C. Li, G. Feng, and X. Liao, "Stabilization of nonlinear systems via periodically intermittent control," *IEEE Transactions on Circuits and Systems II*, vol. 54, no. 11, pp. 1019–1023, 2007.
- [21] T. Huang, C. Li, and X. Liu, "Synchronization of chaotic systems with delay using intermittent linear state feedback," *Chaos*, vol. 18, no. 3, article 033122, 2008.
- [22] T. Huang, C. Li, W. Yu, and G. Chen, "Synchronization of delayed chaotic systems with parameter mismatches by using intermittent linear state feedback," *Nonlinearity*, vol. 22, no. 3, pp. 569–584, 2009.
- [23] X. Yang and J. Cao, "Stochastic synchronization of coupled neural networks with intermittent control," *Physics Letters A*, vol. 373, no. 36, pp. 3259–3272, 2009.
- [24] S. Cai, Z. Liu, F. Xu, and J. Shen, "Periodically intermittent controlling complex dynamical networks with time-varying delays to a desired orbit," *Physics Letters A*, vol. 373, no. 42, pp. 3846–3854, 2009.
- [25] W. Xia and J. Cao, "Pinning synchronization of delayed dynamical networks via periodically intermittent control," *Chaos*, vol. 19, no. 1, article 013120, 2009.
- [26] X. Liu and T. Chen, "Cluster synchronization in directed networks via intermittent pinning control," *IEEE Transactions on Neural Networks*, vol. 22, no. 7, pp. 1009–1020, 2011.
- [27] J. C. Principe, J. M. Kuo, and S. Celebi, "Analysis of the gamma memory in dynamic neural networks," *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 331–337, 1994.
- [28] Z. Wang, Y. Liu, and X. Liu, "On global asymptotic stability of neural networks with discrete and distributed delays," *Physics Letters A*, vol. 345, no. 4–6, pp. 299–308, 2005.
- [29] S. Ruan and R. S. Filfil, "Dynamics of a two-neuron system with discrete and distributed delays," *Physica D*, vol. 191, no. 3–4, pp. 323–342, 2004.
- [30] J. Cao, K. Yuan, D. W. C. Ho, and J. Lam, "Global point dissipativity of neural networks with mixed time-varying delays," *Chaos*, vol. 16, no. 1, article 013105, 2006.
- [31] Z. Wang, Y. Liu, M. Li, and X. Liu, "Stability analysis for stochastic Cohen-Grossberg neural networks with mixed time delays," *IEEE Transactions on Neural Networks*, vol. 17, no. 3, pp. 814–820, 2006.
- [32] Y. Liu, Z. Wang, and X. Liu, "Global exponential stability of generalized recurrent neural networks with discrete and distributed delays," *Neural Networks*, vol. 19, no. 5, pp. 667–675, 2006.
- [33] K. Yuan, J. Cao, and H. X. Li, "Robust stability of switched Cohen-Grossberg neural networks with mixed time-varying delays," *IEEE Transactions on Systems, Man, and Cybernetics B*, vol. 36, no. 6, pp. 1356–1363, 2006.
- [34] K. Gu, V. Kharitonov, and J. Chen, *Stability of Time-Delay Systems*, Birkhauser, Boston, Mass, USA, 2003.
- [35] A. Halanay, *Differential Equations: Stability, Oscillations, Time Lags*, Academic Press, New York, NY, USA, 1966.

Research Article

Dynamic Analysis of a Hybrid Squeeze Film Damper Mounted Rub-Impact Rotor-Stator System

Cai-Wan Chang-Jian

*Department of Mechanical and Automation Engineering, I-Shou University, No. 1, Section 1,
Hsueh-Cheng Road, Ta-Hsu District, Kaohsiung 840, Taiwan*

Correspondence should be addressed to Cai-Wan Chang-Jian, cwchangjian@mail.isu.edu.tw

Received 2 March 2012; Revised 8 May 2012; Accepted 8 May 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 Cai-Wan Chang-Jian. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

An investigation is carried out on the systematic analysis of the dynamic behavior of the hybrid squeeze-film damper (HSFD) mounted a rotor-bearing system with strongly nonlinear oil-film force and nonlinear rub-impact force in the present study. The dynamic orbits of the system are observed using bifurcation diagrams plotted using the dimensionless rotating speed ratio as control parameters. The onset of chaotic motion is identified from the phase diagrams, power spectra, Poincaré maps, bifurcation diagrams, maximum Lyapunov exponents, and fractal dimension of the rotor-bearing system. The dynamic behaviors are unlike the usual ways into chaos ($1T \Rightarrow 2T \Rightarrow 4T \Rightarrow 8T \Rightarrow 16T \Rightarrow 32T \cdots \Rightarrow$ chaos or periodic \Rightarrow quasi-periodic \Rightarrow chaotic), it suddenly gets in chaos from the periodic motion without any transition. The results presented in this study provide some useful insights into the design and development of a rotor-bearing system for rotating machinery that operates in highly rotating speed and highly nonlinear regimes.

1. Introduction

Squeeze-film damper (SFD) bearing is actually a special type of journal bearing with its journal mechanically prevented from rotating but free to vibrate within the clearance space. The hybrid squeeze-film damper (HSFD) and the porous squeeze-film damper (PSFD) are the well-known applications of SFD and also useful for industry. Some literatures discussed dynamic behaviors in SFD bearings and also found many interesting and useful results. Holmes et al. [1] published a paper dealing with aperiodic behavior in journal bearings and what may very well have been the first paper about aperiodic behavior in journal bearing systems. Nikolajsent and Holmes [2] reported their observation of

nonsynchronous vibrations in a test rig of a flexible, symmetric rotor on two identical plain journal bearings supported by centralized squeeze-film dampers. Sykes and Holmes [3] showed experimental observations of subharmonic motion in squeeze film bearings and linked this to possible precursors of chaotic motion. At the same time, Kim and Noah [4] analyzed the bifurcation of a modified Jeffcott rotor with bearing clearance. Ehrich [5] used a simple numerical model of a Jeffcott rotor mounted on a nonlinear spring. It was found that the vibratory response in the transition zone midway between adjacent zones of subharmonic response has all the characteristics of chaotic behavior. Zhao et al. [6] discussed the subharmonic and quasiperiodic motions of an eccentric squeeze film damper-mounted rigid rotor system. Brown [7] studied a simple model of a rigid and hydrodynamically supported journal bearing, using a short bearing theory. Theoretical and experimental investigations were reported by Adiletta et al. [8–10] in which a rigid rotor in short bearings would have subharmonic, quasiperiodic, and chaotic motion for suitable values of the system parameters. Sundararajan and Noah [11] proposed a simple shooting scheme along with an arc-length continuation algorithm with applications to periodically forced rotor systems. The occurrence of periodic, quasiperiodic and chaotic motions was predicted for various ranges of rotor speeds. Chang-Jian and Chen [12–16] presented a series of papers discussing about flexible rotor supported by journal bearings under nonlinear suspension and also combined with rub-impact effect, turbulent effect and micropolar lubricant into consideration. They found very bountiful nonperiodic responses occurring in rotor-bearing systems, and the studies would help engineers or scientists escape undesired motions in either designing or operating rotor-bearing systems.

Although virtually all physical phenomena in the real world can be regarded as nonlinear, most of these phenomena can be simplified to a linear form given a sufficiently precise linearization technique. However, this simplification is inappropriate for high-power, high rotating speed system and its application during the design and analysis stage may result in a flawed or potentially dangerous operation. As a result, nonlinear analysis methods are generally preferred within engineering and academic circles. The current study performs a nonlinear analysis of the dynamic behavior of a rotor-bearing system equipped with hybrid squeeze-film damper under nonlinear rub-impact force effect. The nondimensional equation of the rotor-bearing system is then solved using the fourth-order Runge-Kutta method. The nonperiodic behavior of this system is characterized using phase diagrams, power spectra, Poincaré maps, bifurcation diagrams, Lyapunov exponents, and the fractal dimension of the system.

2. Mathematical Modeling

Figure 1 shows a rotor supported on HSFDs in parallel with retaining springs. The bearing consists of four hydrostatic chambers and four hydrodynamic regions. The oil film supporting force is dependent on the integrated action of hydrodynamic pressure and hydrostatic pressure of HSFD. Figure 2(a) represents the cross-section of HSFD and rub-impact rotor-stator model. The structure of this kind bearing should be popularized to consist of $2N$ ($N = 2, 3, 4 \dots$) hydrostatic chambers and $2N$ hydrodynamic regions. In this study, oil pressure distribution model in the HSFD is proposed to integrate the pressure distribution of dynamic pressure region and static pressure region.

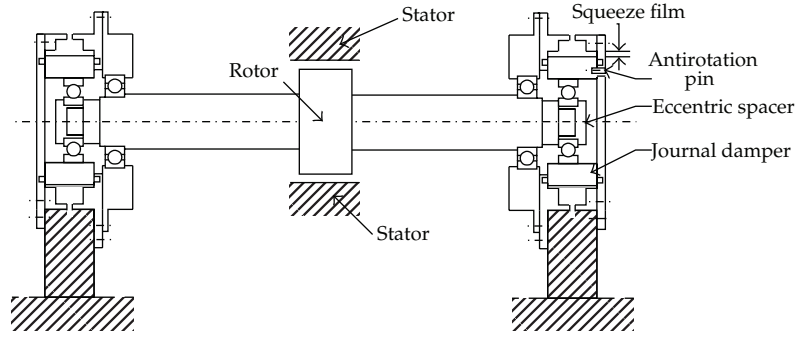


Figure 1: Schematic illustration of hybrid squeeze-film damper mounted the rotor-bearing system.

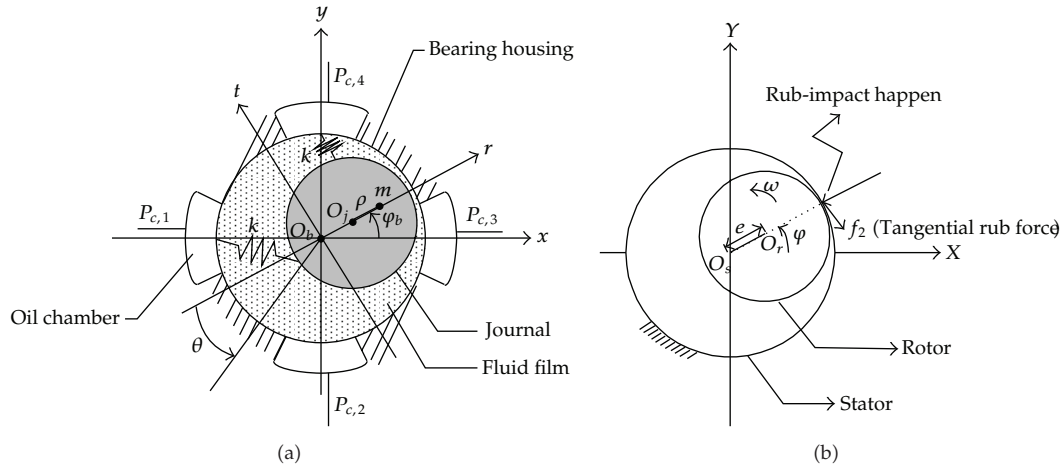


Figure 2: Cross-section of HSFD model and rub-impact rotor-stator model.

2.1. The Instant Oil Film Supporting Force for HSFD

To analyze the pressure distribution, the Reynolds equation for constant lubricant properties and noncompressibility should be assumed, then the Reynolds equation is introduced as follows [12]:

$$\frac{1}{R^2} \frac{\partial}{\partial \theta} \left(h^3 \frac{\partial p}{\partial \theta} \right) + \frac{1}{R^2} \frac{\partial}{\partial z} \left(h^3 \frac{\partial p}{\partial z} \right) = -12\mu\Omega \frac{\partial h}{\partial \theta} + 12\mu \frac{\partial h}{\partial t}. \quad (2.1)$$

The supporting region of HSFD should be divided into three regions: static pressure region, rotating direction dynamic pressure region, and axial direction dynamic pressure region, as shown in Figure 2. In the part of HSFD with $-a \leq z \leq a$, the long bearing theory is assumed and Reynolds equation is solved with the boundary condition of static pressure region $p_{c,i}$ acquiring the pressure distribution $p_0(\theta)$. In the part of HSFD with $a \leq |z| \leq L/2$,

the short bearing theory is assumed and solves the Reynolds equation with the boundary condition of $p(z, \theta)|_{z=\pm a} = p_0(\theta)$ and $p(z, \theta)|_{z=\pm L/2} = 0$, yielding the pressure distribution in axis direction dynamic pressure region $p(z, \theta)$. Finally, a formula of pressure distribution in whole supporting region is obtained.

According to the above conditions, the instant oil film pressure distribution is as follows. The instant pressure in rotating direction within the range of $-a \leq z \leq a$ is

$$p_0(\theta) = \begin{cases} p_{c,i}; & \frac{\pi}{2}(i-1) - \frac{\beta}{2} - \varphi_b \leq \theta \leq \frac{\pi}{2}(i-1) + \frac{\beta}{2} - \varphi_b \\ p_i(\theta); & \frac{\pi}{2}(i-1) + \frac{\beta}{2} - \varphi_b \leq \theta \leq \frac{\pi}{2}i - \frac{\beta}{2} - \varphi_b, \end{cases} \quad i = 1, 2, 3, 4, \quad (2.2)$$

where

$$p_i(\theta) = p_{c,i} + \frac{6\mu R^2}{\delta^2} \frac{\dot{\varepsilon}}{\varepsilon} \left[\frac{1}{(1 + \varepsilon \cos \theta)^2} - \frac{1}{(1 + \varepsilon \cos \theta_{i1})^2} \right] \quad (2.3)$$

$$+ c_1 \int_{\theta_{i1}}^{\theta} \frac{1}{\delta^3 (1 + \varepsilon \cos \theta)^3} d\theta - \int_{\theta_{i1}}^{\theta} \frac{12\dot{\varphi}_b \mu \delta \varepsilon R^2 \cos \theta}{\delta^3 (1 + \varepsilon \cos \theta)^3} d\theta \quad i = 1, 2, 3, 4,$$

$$c_1 = \frac{p_{c,i+1} - p_{c,i} - (6\mu R^2 / \delta^2) (\dot{\varepsilon} / \varepsilon) \left[1 / (1 + \varepsilon \cos \theta_{i2})^2 - 1 / (1 + \varepsilon \cos \theta_{i1})^2 \right]}{\int_{\theta_{i1}}^{\theta_{i2}} (1 / \delta^3 (1 + \varepsilon \cos \theta)^3) d\theta} + \frac{\int_{\theta_{i1}}^{\theta_{i2}} (12\dot{\varphi}_b \mu \delta \varepsilon R^2 \cos \theta / \delta (1 + \varepsilon \cos \theta)^3) d\theta}{\int_{\theta_{i1}}^{\theta_{i2}} (1 / \delta^3 (1 + \varepsilon \cos \theta)^3) d\theta}, \quad (2.4)$$

$$\theta_{i1} = (i-1) \frac{\pi}{2} + \frac{\beta}{2} - \varphi_b, \quad \theta_{i2} = i \frac{\pi}{2} - \frac{\beta}{2} - \varphi_b, \quad i = 1, 2, 3, 4.$$

The instant pressure in the axis direction within the range of $a \leq |z| \leq L/2$ is

$$p(\theta, z) = \left(\frac{L}{2} - |z| \right) \left\{ [A_1(\theta) \dot{\varphi}_b \varepsilon + A_2(\theta) \dot{\varepsilon}] (a - |z|) + p_0(\theta) \frac{1}{L/2 - a} \right\}, \quad (2.5)$$

where

$$A_1(\theta) = \frac{6\mu \delta \sin \theta}{\delta^3 (1 + \varepsilon \cos \theta)^3}; \quad A_2(\theta) = \frac{6\mu \delta \cos \theta}{\delta^3 (1 + \varepsilon \cos \theta)^3}. \quad (2.6)$$

The instant oil film forces of the different elements are determined by integrating (2.2) and (2.5) over the area of the journal sleeve. In the static pressure region, the forces are

$$\begin{aligned} F_{rs} &= \sum_{i=1}^4 p_{ci} 2aR \left[\sin\left(\frac{\pi}{2}(i-1) + \frac{\beta}{2} - \varphi_b\right) - \sin\left(\frac{\pi}{2}(i-1) - \frac{\beta}{2} - \varphi_b\right) \right], \\ F_{\tau s} &= \sum_{i=1}^4 p_{ci} 2aR \left[\cos\left(\frac{\pi}{2}(i-1) - \frac{\beta}{2} - \varphi_b\right) - \cos\left(\frac{\pi}{2}(i-1) + \frac{\beta}{2} - \varphi_b\right) \right]. \end{aligned} \quad (2.7)$$

In the rotating direction dynamics pressure region, the forces are

$$\begin{aligned} F_{rc} &= \sum_{i=1}^4 \int_{\theta_{i1}}^{\theta_{i2}} p_i(\theta) R 2a \cos \theta d\theta, \\ F_{\tau c} &= \sum_{i=1}^4 \int_{\theta_{i1}}^{\theta_{i2}} p_i(\theta) R 2a \sin \theta d\theta. \end{aligned} \quad (2.8)$$

In the axial direction dynamic pressure region, the forces are

$$\begin{aligned} F_{ra} &= \int_{-L/2}^{-a} dz \int_0^{2\pi} p(\theta, z) R \cos \theta d\theta + \int_{L/2}^a dz \int_0^{2\pi} p(\theta, z) R \cos \theta d\theta, \\ F_{\tau a} &= \int_{-L/2}^{-a} dz \int_0^{2\pi} p(\theta, z) R \sin \theta d\theta + \int_{L/2}^a dz \int_0^{2\pi} p(\theta, z) R \sin \theta d\theta. \end{aligned} \quad (2.9)$$

The resulting damper forces in the radial and tangential directions are determined by summing the above supporting forces. It is as follows:

$$\begin{aligned} F_r &= F_{rs} + F_{rc} + F_{ra}, \\ F_{\tau} &= F_{\tau s} + F_{\tau c} + F_{\tau a}. \end{aligned} \quad (2.10)$$

2.2. Rub-Impact Force

Figure 2(b) shows the radial impact force f_1 and the tangential rub force f_2 . f_1 and f_2 could be expressed as [17]

$$\begin{aligned} f_1 &= (e - \delta)k_c \\ f_2 &= (f + bv)f_1, \quad \text{if } e \geq \delta. \end{aligned} \quad (2.11)$$

Then we could get the rub-impact forces in the horizontal and vertical directions as follows:

$$\begin{aligned} R_x &= -\frac{(e - \delta)k_c}{e} [X - (f + bv)Y] \\ R_y &= -\frac{(e - \delta)k_c}{e} [(f + bv)X + Y]. \end{aligned} \quad (2.12)$$

2.3. Dynamics Equation

The equations of rotor motion in the Cartesian coordinates can be written as

$$\begin{aligned} m\ddot{x} + d\dot{x} + kx &= m\rho\omega^2 \cos \omega t + f_x + kx_0 + R_x, \\ m\ddot{y} + d\dot{y} + ky &= m\rho\omega^2 \sin \omega t + f_y + ky_0 + R_y. \end{aligned} \quad (2.13)$$

The origin of the o -xyz-coordinate system is taken to be the bearing center O_b . Dividing these two equations by $m\omega^2$ and defining a nondimensional time $\phi = \omega t$ and a speed parameter $s = \omega/\omega_n$, one obtains the following nondimensionalized equations of motion:

$$X'' + \frac{D}{s}X' + \frac{1}{s^2}X = U \cos \phi + \frac{B}{s} \left(\frac{XF_r - YF_\tau}{\varepsilon} \right) + \frac{X_0}{s^2} + \frac{R_x}{m\omega^2}, \quad (2.14)$$

$$Y'' + \frac{D}{s}Y' + \frac{1}{s^2}Y = U \sin \phi + \frac{B}{s} \left(\frac{YF_r + XF_\tau}{\varepsilon} \right) + \frac{Y_0}{s^2} + \frac{R_y}{m\omega^2}. \quad (2.15)$$

Equations (2.14)~(2.15) describe a nonlinear dynamic system. In the current study, the approximate solutions of these coupled nonlinear differential equations are obtained using the fourth-order Runge-Kutta numerical scheme.

3. Analytical Tools for Observing Nonlinear Dynamics of Rotor-Bearing System

In the present study, the nonlinear dynamics of the rotor-bearing system equipped with HSFD shown in Figure 1 are analyzed using Poincaré maps, bifurcation diagrams, the Lyapunov exponent and the fractal dimension. The basic principles of each analytical method are reviewed in the following subsections.

3.1. Dynamic Trajectories and Poincaré Maps

The dynamic trajectories of the rotor-bearing system provide a basic indication as to whether the system behavior is periodic or nonperiodic. However, they are unable to identify the onset of chaotic motion. Accordingly, some other form of analytical method is required. In the current study, the dynamics of the rotor-bearing system are analyzed using Poincaré maps derived from the Poincaré section of the rotor system. A Poincaré section is a hypersurface in the state-space transverse to the flow of the system of interest. In nonautonomous systems, points on the Poincaré section represent the return points of a time series corresponding to a constant interval T , where T is the driving period of the excitation force. The projection of the Poincaré section on the $y(nT)$ plane is referred to as the Poincaré map of the dynamic system. When the system performs quasiperiodic motion, the return points in the Poincaré map form a closed curve. For chaotic motion, the return points form a fractal structure comprising many irregularly distributed points. Finally, for nT -periodic motion, the return points have the form of n discrete points.

3.2. Power Spectrum

In this study, the spectrum components of the motion performed by the rotor-bearing system are analyzed by using the Fast Fourier Transformation method to derive the power spectrum of the displacement of the dimensionless dynamic transmission error. In the analysis, the frequency axis of the power spectrum plot is normalized using the rotating speed, ω .

3.3. Bifurcation Diagram

A bifurcation diagram summarizes the essential dynamics of a rotor-train system and is therefore a useful means of observing its nonlinear dynamic response. In the present analysis, the bifurcation diagrams are generated using two different control parameters, namely the dimensionless unbalance coefficient, β , and the dimensionless rotating speed ratio, s , respectively. In each case, the bifurcation control parameter is varied with a constant step, and the state variables at the end of one integration step are taken as the initial values for the next step. The corresponding variations of the $y(nT)$ coordinates of the return points in the Poincaré map are then plotted to form the bifurcation diagram.

3.4. Lyapunov Exponent

The Lyapunov exponent of a dynamic system characterizes the rate of separation of infinitesimally close trajectories and provides a useful test for the presence of chaos. In a chaotic system, the points of nearby trajectories starting initially within a sphere of radius ε_0 form after time t an approximately ellipsoidal distribution with semiaxes of length $\varepsilon_j(t)$. The Lyapunov exponents of a dynamic system are defined by $\lambda_j = \lim_{t \rightarrow \infty} (1/t) \log(\varepsilon_j(t)/\varepsilon_0)$, where λ_j denotes the rate of divergence of the nearby trajectories. The exponents of a system are usually ordered into a Lyapunov spectrum, that is, $\lambda_1 > \lambda_2 > \dots > \lambda_m$. A positive value of the maximum Lyapunov exponent (λ_1) is generally taken as an indication of chaotic motion [16].

3.5. Fractal Dimension

The presence of chaotic vibration in a system is generally detected using either the Lyapunov exponent or the fractal dimension property. The Lyapunov exponent test can be used for both dissipative systems and nondissipative (i.e. conservative) systems, but is not easily applied to the analysis of experimental data. Conversely, the fractal dimension test can only be used for dissipative systems but is easily applied to experimental data. In contrast to Fourier transform-based techniques and bifurcation diagrams, which provide only a general indication of the change from periodic motion to chaotic behavior, dimensional measures allow chaotic signals to be differentiated from random signals. Although many dimensional measures have been proposed, the most commonly applied measure is the correlation dimension d_G defined by Grassberger and Procaccia due to its computational speed and the consistency of its results. However, before the correlation dimension of a dynamic system flow can be evaluated, it is first necessary to generate a time series of one of the system variables using a time-delayed pseudo-phase-plane method. Assume an original time series of $x_i = \{x(i\tau); i = 1, 2, 3, \dots, N\}$, where τ is the time delay (or sampling time). If the system is acted upon by an excitation force with a frequency ω , the sampling time, τ , is generally chosen such that it is much smaller than the driving period. The delay coordinates are then used to construct an n -dimensional vector $X = (x(j\tau), x[(j+1)\tau], x[(j+2)\tau], \dots, x[(j+n-1)\tau])$, where $j = 1, 2, 3, \dots, (N - n + 1)$. The resulting vector comprises a total of $(N - n + 1)$ vectors, which are then plotted in an n -dimensional embedding space. Importantly, the system flow in the reconstructed n -dimensional phase space retains the dynamic characteristics of the system in the original phase space. In other words, if the system flow has the form of a closed orbit in the original phase plane, it also forms a closed path in the n -dimensional embedding space. Similarly, if the system exhibits a chaotic behavior in the original phase plane, its path in the embedding space will also be chaotic. The characteristics of the attractor in the n -dimensional embedding space are generally tested using the function $\sum_{i,j=1}^N H(r - |x_i - x_j|)$ to determine the number of pairs (i, j) lying within a distance $|x_i - x_j| < r$ in $\{x_i\}_{i=1}^N$, where H denotes the Heaviside step function, N represents the number of data points, and r is the radius of an n -dimensional hypersphere. For many attractors, this function exhibits a power law dependence on r as $r \rightarrow 0$, that is $c(r) \propto r^{d_G}$. Therefore, the correlation dimension, d_G , can be determined from the slope of a plot of $[\log c(r)]$ versus $[\log r]$. Chen and Yau [18] showed that the correlation dimension represents the lower bound to the capacity or fractal dimension d_c and approaches its value asymptotically when the attracting set is distributed more uniformly in the embedding phase space. A set of points in the embedding space is said to be fractal if its dimension has a finite noninteger value. Otherwise, the attractor is referred to as a "strange attractor." To establish the nature of the attractor, the embedding dimension is progressively increased, causing the slope of the characteristic curve to approach a steady-state value. This value is then used to determine whether the system has a fractal structure or a strange attractor structure. If the dimension of the system flow is found to be fractal (i.e. to have a noninteger value), the system is judged to be chaotic.

In the current study, the attractors in the embedding space were constructed using a total of 60000 data points taken from the time series corresponding to the displacement of the system. Via a process of trial and error, the optimum delay time when constructing the time series was found to correspond to one third of a revolution of the system. The reconstructed attractors were placed in embedding spaces with dimensions of $n = 2, 4, 6, 8, 10, 12, 14, 16, 18$, and 20, respectively, yielding 10 different $[\log c(r)]$ versus $[\log r]$ plots for each attractor. The number of data points chosen for embedding purposes (i.e., 60000)

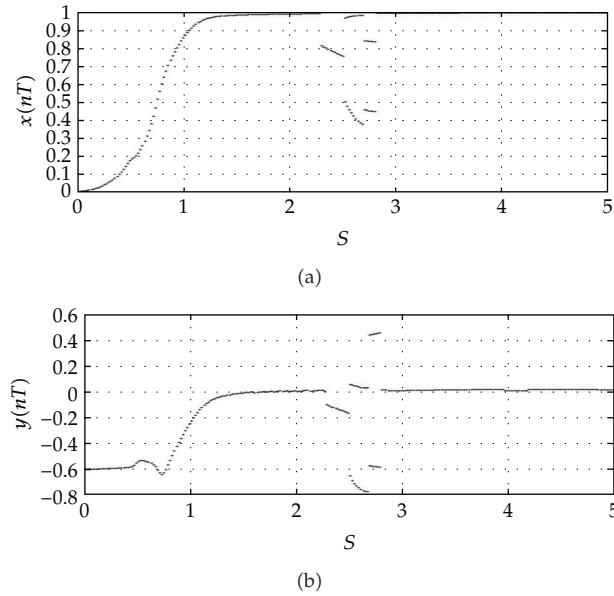


Figure 3: Bifurcation diagram of $X(nT)$ (a) and $Y(nT)$ (b) versus rotor speed s (without rub-impact effect).

reflects the need for a compromise between the computation time and the accuracy of the results. In accordance with Grassberger and Procaccia [19], the number of points used to estimate the intrinsic dimension of the attracting set in the current analysis is less than 42^M , where M is the greatest integer value less than the fractal dimension of the attracting set.

4. Numerical Results and Discussions

The nonlinear dynamic equations presented in (2.14) to (2.15) for the HSFD rotor-bearing system with strongly nonlinear oil-film force and nonlinear rub-impact force were solved using the fourth-order Runge-Kutta method. The time step in the iterative solution procedure was assigned a value of $\pi/300$, and the termination criterion was specified as an error tolerance of less than 0.0001. The time series data corresponding to the first 800 revolutions of the rotor was deliberately excluded from the dynamic analysis to ensure that the analyzed data related to steady-state conditions. The sampled data were used to generate the dynamic trajectories, Poincaré maps, and bifurcation diagrams of the spur rotor system in order to obtain a basic understanding of its dynamic behavior. The maximum Lyapunov exponent and the fractal dimension measure were then used to identify the onset of chaotic motion. The rotating speed ratio s is one of the most significant and commonly used as a control parameter in analyzing dynamic characteristics of bearing systems. Accordingly, the dynamic behavior of the current rotor-bearing system was examined using the dimensionless rotating speed ratio s as a bifurcation control parameter.

The bifurcation diagram in Figure 3 shows the long-term values of the rotational angle, plotted with rotor displacement against the dimensionless speed s without rub-impact effect.

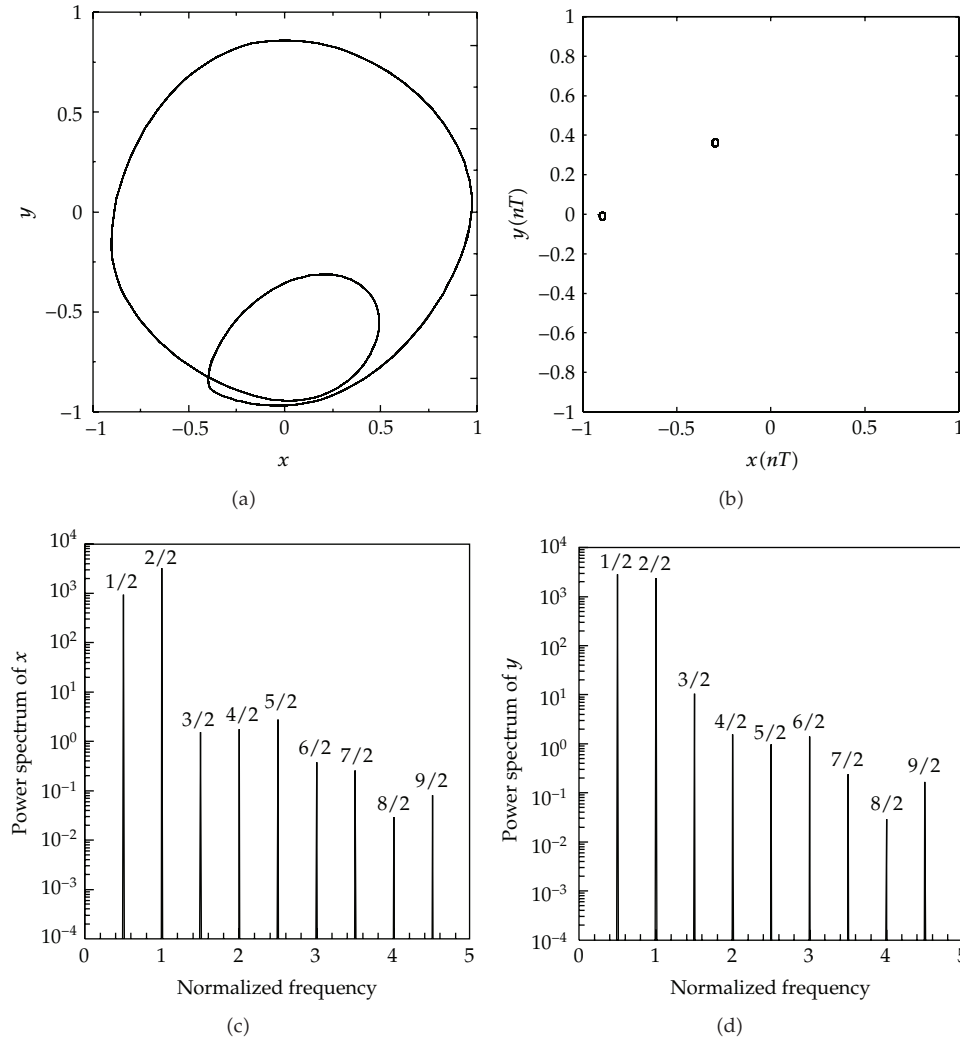


Figure 4: Subharmonic motion at $s = 2.6$ (case 1); (a) Rotor trajectory; (b) Poincaré map; (c) and (d) Displacement power in X and Y directions (without rub-impact effect).

Qualitatively different behavior was observed at values of s within the range of $0 < s < 5$. It can be seen that the dynamic motion of rotor trajectory in low speed is T -periodic motion both in X and Y directions, and it drops to a lower spatial displacement mode at the speed $s = 2.3$. As the speed is increased, the T -period motion loses its stability at $s = 2.52$, and a $2T$ -periodic motion starts to build up. The jump phenomenon is also occurred under $2T$ -periodic motion at $s = 2.7$. As the speed is further increased, the $2T$ -periodic motion loses its stability at $s = 2.82$, and a T -periodic motion suddenly appears. The rotor trajectory, the Poincaré map, and the displacement power spectrum in the X and Y directions at $s = 2.6$ are given in Figure 4, from which the 0.5 -subharmonic motion is shown by the double loops of the rotor trajectory, two discrete points in the Poincaré map and peaks at 0.5 in the power spectrum. The pressure distributions in the four oil chambers are shown in Figure 5. It can be seen that

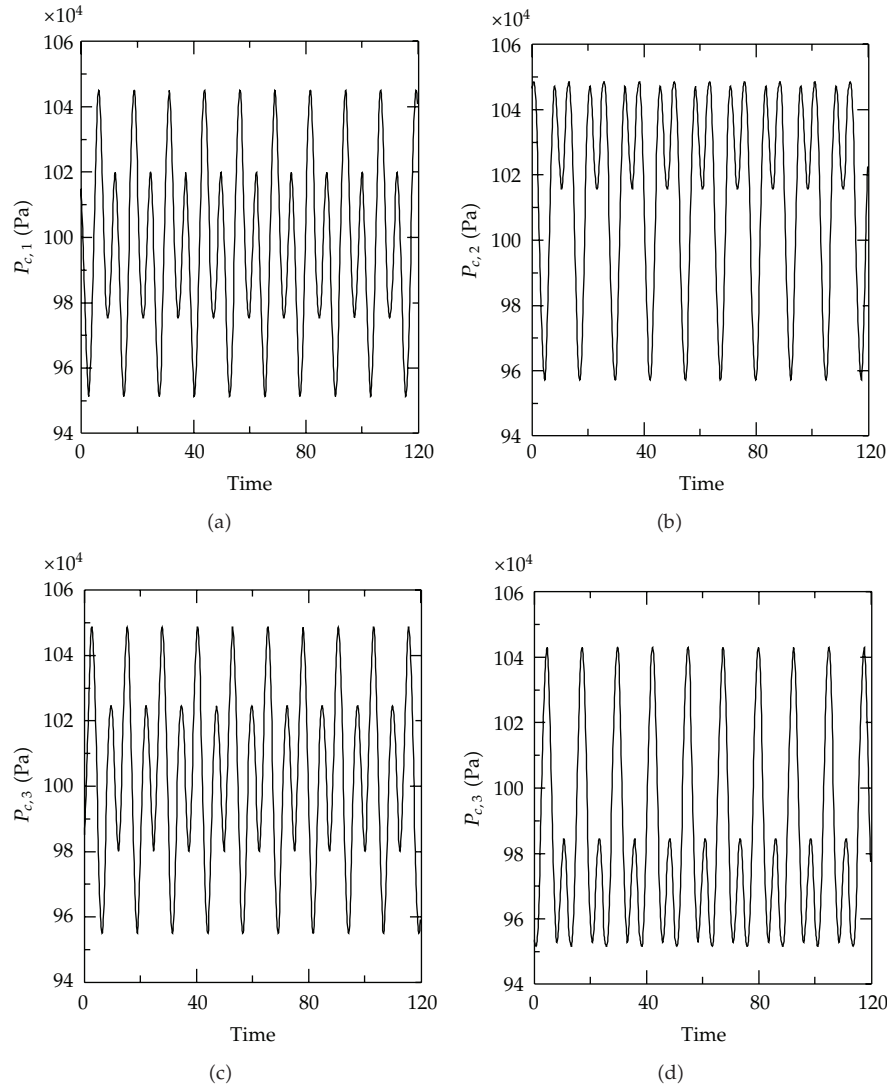


Figure 5: Pressure distribution in the static pressure chamber at $s = 2.6$ (without rub-impact effect).

the variations of pressure distributions are periodic, and the period is the same with the rotor trajectory.

Figures 6(a) and 6(b) show the bifurcation diagrams for the rotor displacement against the dimensionless rotating speed ratio with rub-impact effect. Compared with bifurcation results without rub-impact effect, bifurcation results with rub-impact effect show that dynamic trajectories perform strongly nonperiodic at low rotating speeds, but it would escape nonperiodic motions to periodic motions. The bifurcation diagrams show that the geometric centers of rotor in the horizontal and vertical directions perform nonperiodic motion or the so-called chaotic motion at low values of the rotating speed ratio, that is, $s < 0.61$. Figures 7, 8, and 9 represent phase diagrams, power spectra, Poincaré maps, Lyapunov exponents, and the fractal dimensions of pinion center with $s = 0.32, 0.36$,

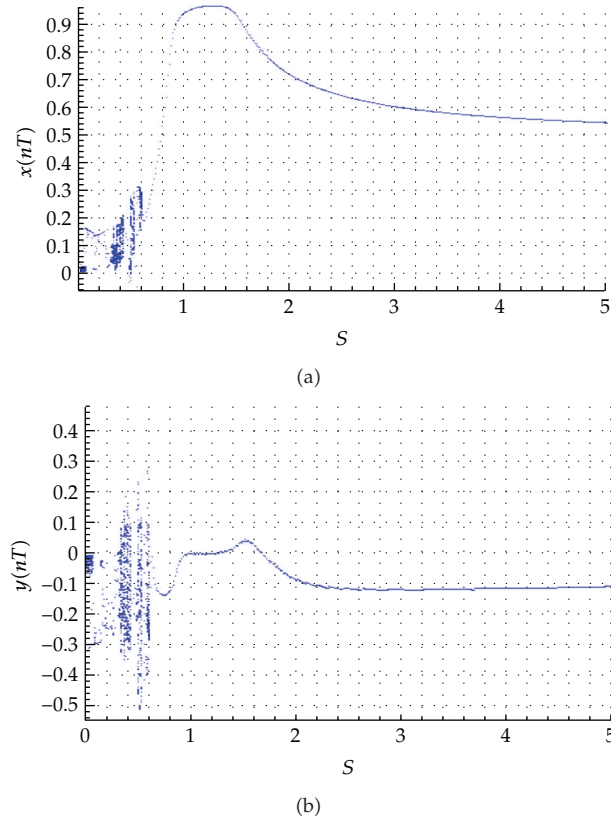


Figure 6: Bifurcation diagram of HSFD rotor-bearing system using dimensionless rotating speed coefficient, s , as bifurcation parameter (with rub-impact effect).

and $s = 0.42$, respectively. The simulation results show that phase diagrams show disordered dynamic behaviors with $s = 0.32, 0.36$ and $s = 0.42$; power spectra reveal numerous excitation frequencies; the return points in the Poincaré maps form some geometrically fractal structures, but the maximum Lyapunov exponent is positive with $s = 0.36$ and maximum Lyapunov exponent is negative with $s = 0.32$ and $s = 0.42$. Thus, the results show that the dynamic trajectory performs chaotic motion with $s = 0.36$, but they present no chaotic motions with $s = 0.32$ and $s = 0.42$. Figures 10 and 11 are phase diagrams and Poincaré maps for the route of subharmonic motion into chaos, out of chaos to periodic response at different rotating speed ratios of s (with rub-impact effect). Unlike the usual ways into chaos ($1T \Rightarrow 2T \Rightarrow 4T \Rightarrow 8T \Rightarrow 16T \Rightarrow 32T \dots \Rightarrow$ chaos or periodic \Rightarrow quasi-periodic \Rightarrow chaotic), it suddenly gets in chaos from the periodic motion without any transition or suddenly escape from irregular motions into periodic motions in accordance with phase diagrams and Poincaré maps.

5. Conclusions

A hybrid squeeze-film damper mounted rotor-bearing system with nonlinear oil-film force and nonlinear rub-impact force has been presented and studied by a numerical analysis

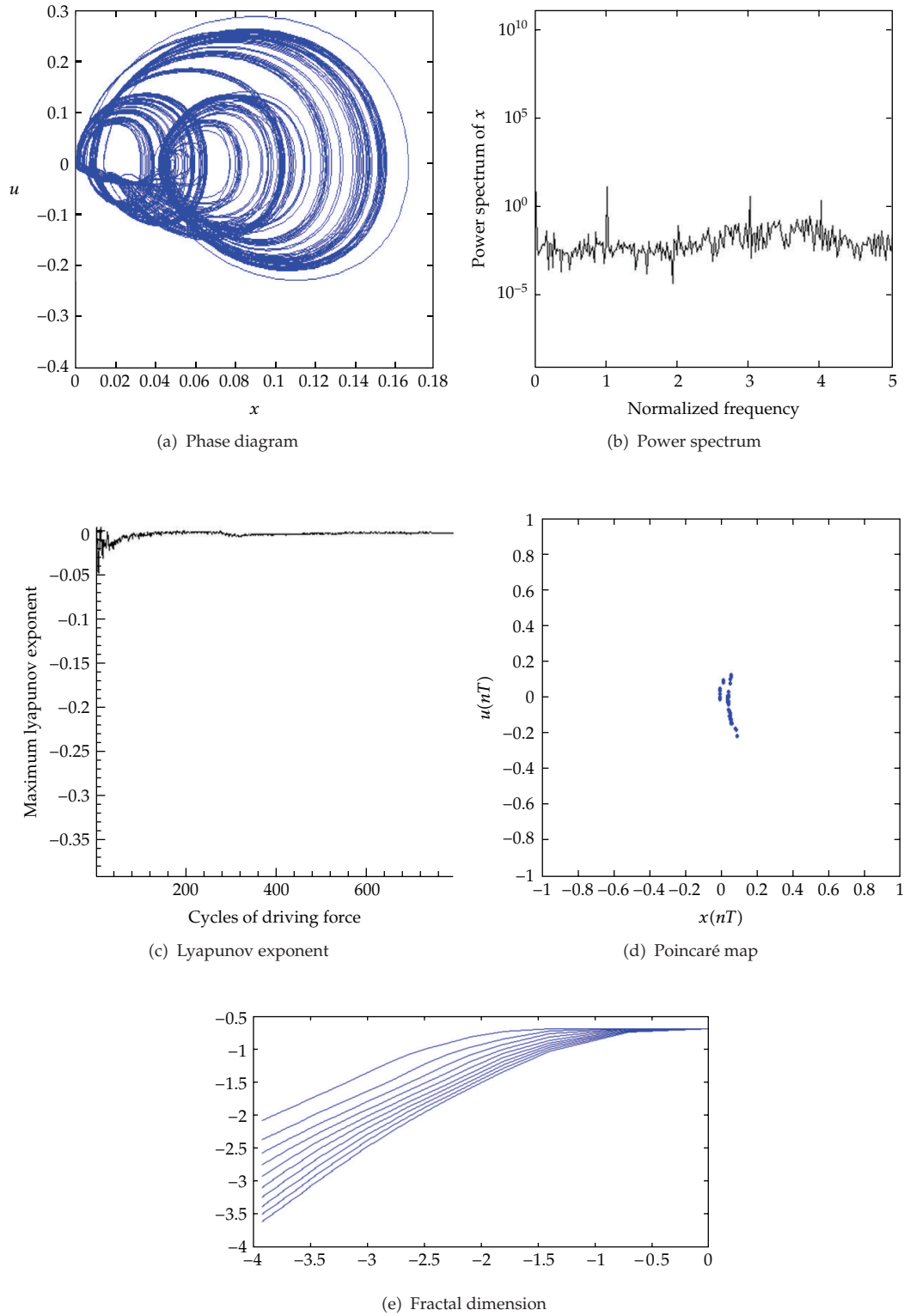


Figure 7: Simulation results obtained for rotor-bearing system with $s = 0.32$ (with rub-impact effect).

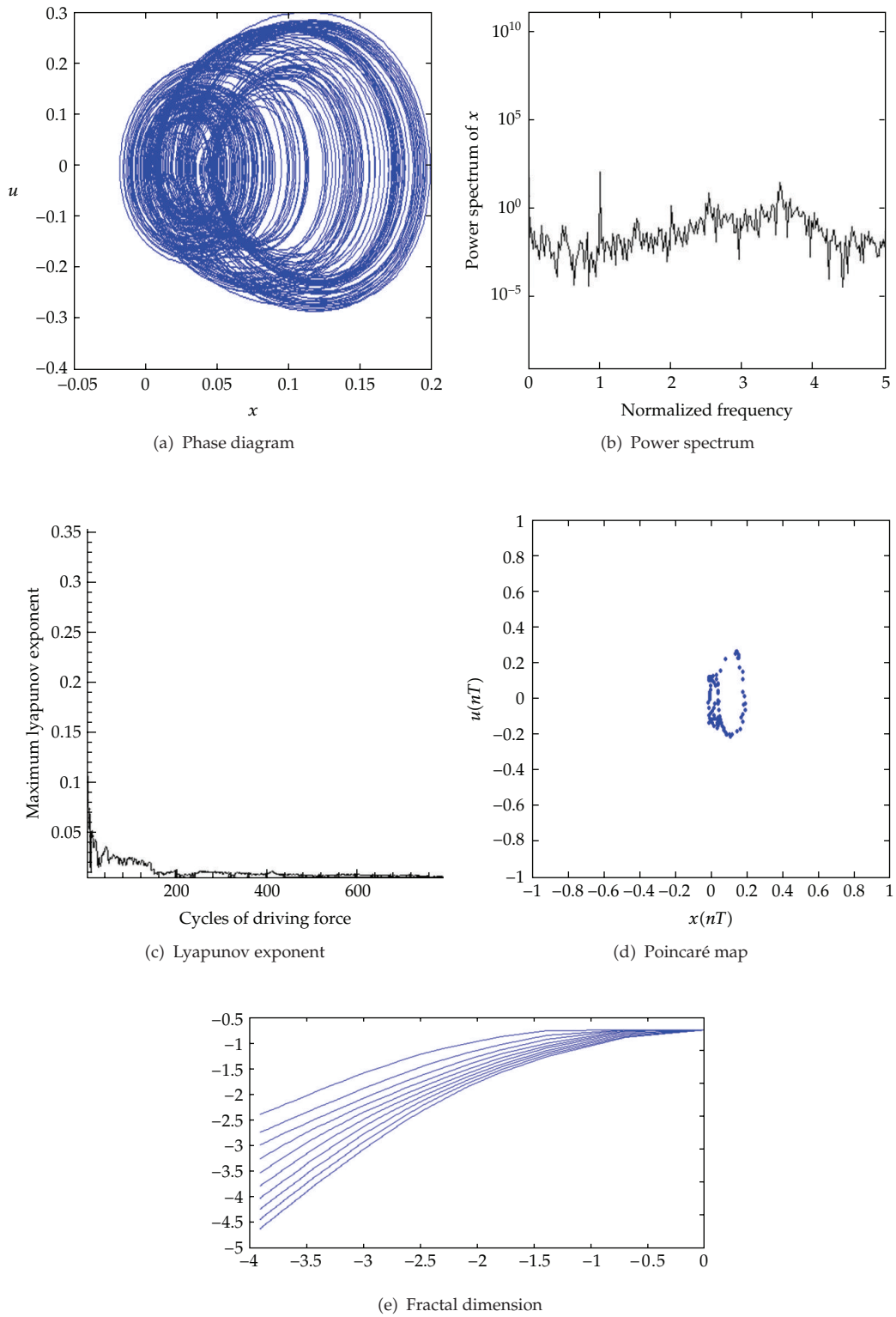


Figure 8: Simulation results obtained for rotor-bearing system with $s = 0.36$ (with rub-impact effect).

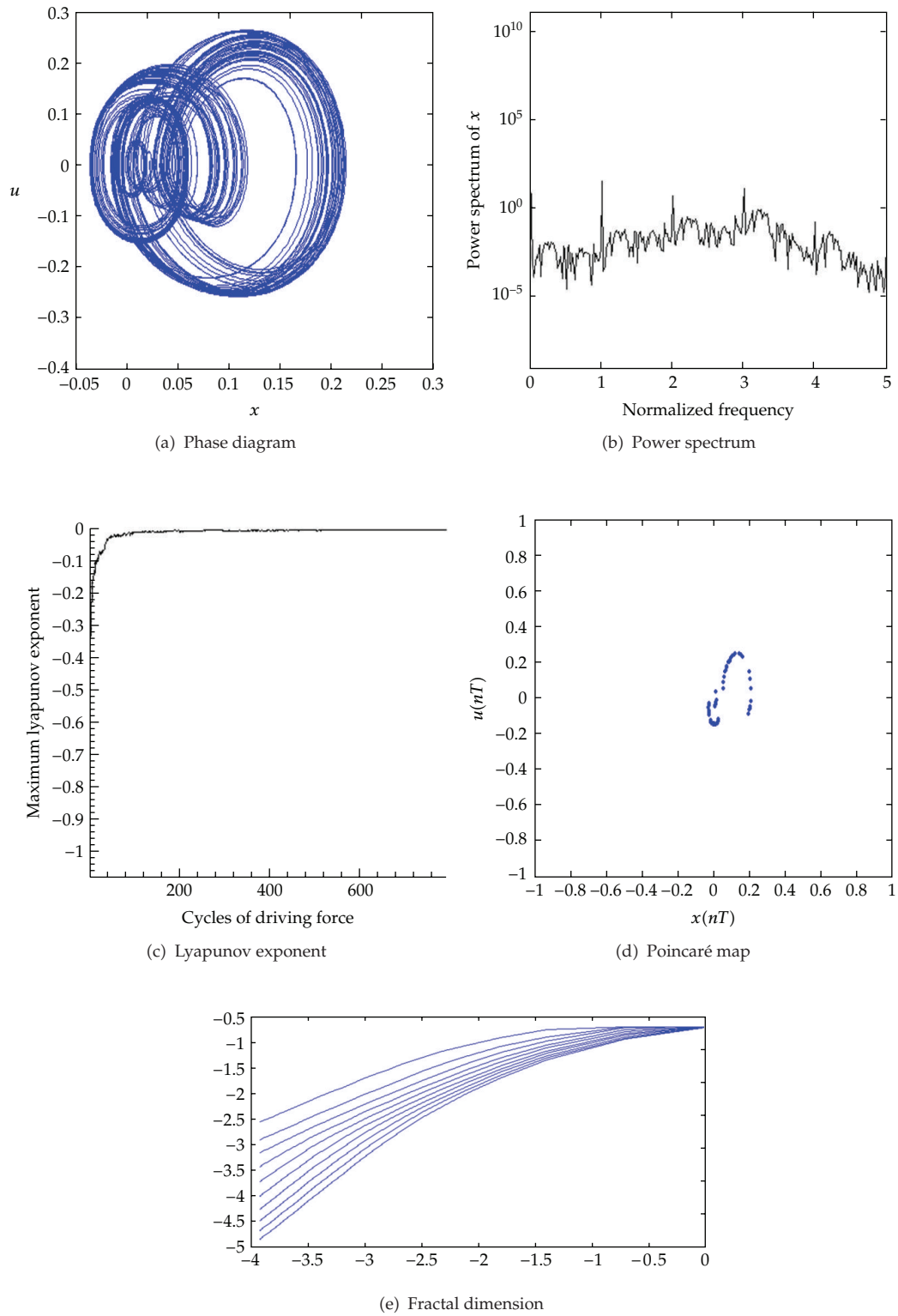


Figure 9: Simulation results obtained for rotor-bearing system with $s = 0.42$ (with rub-impact effect).

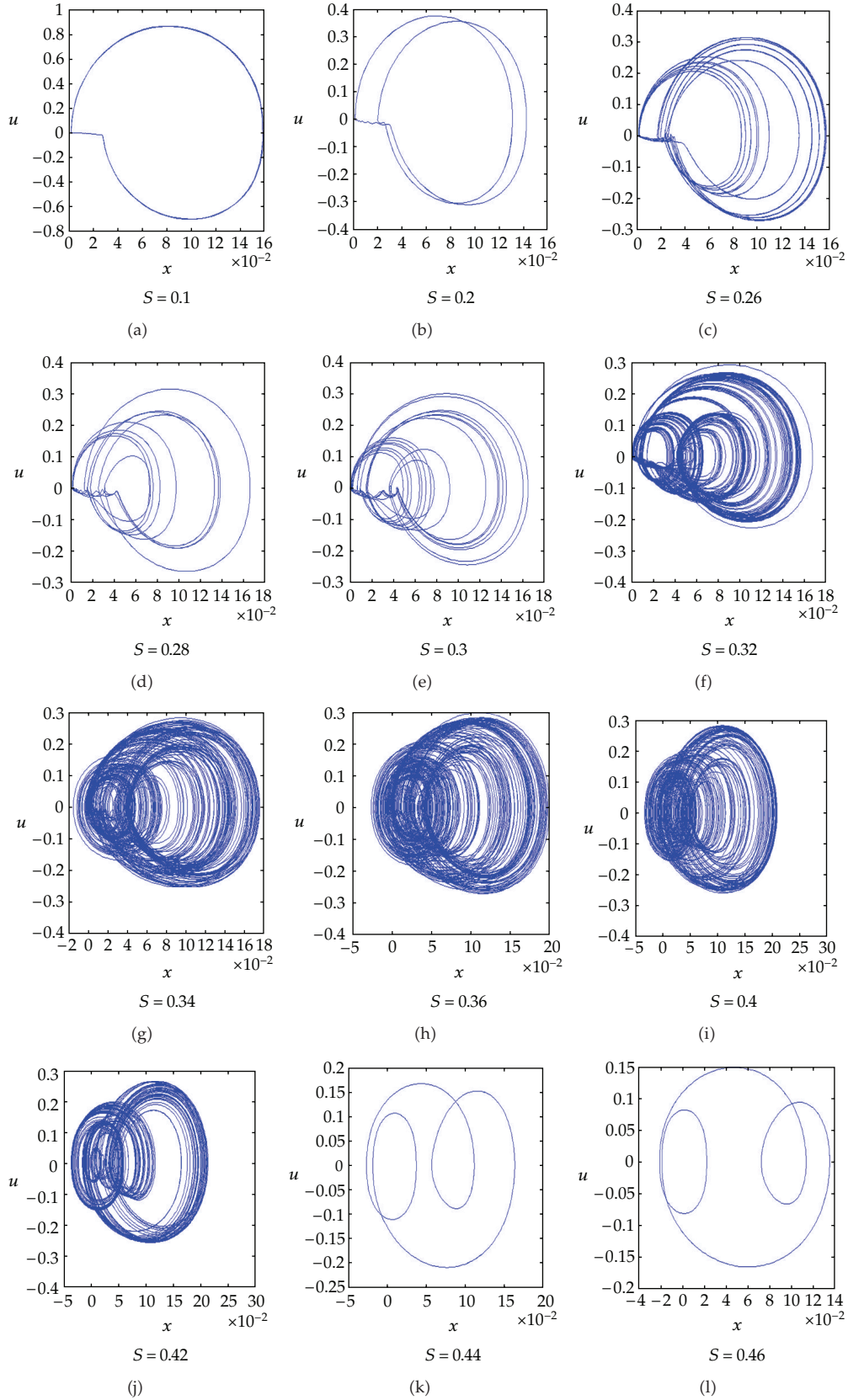


Figure 10: Continued.

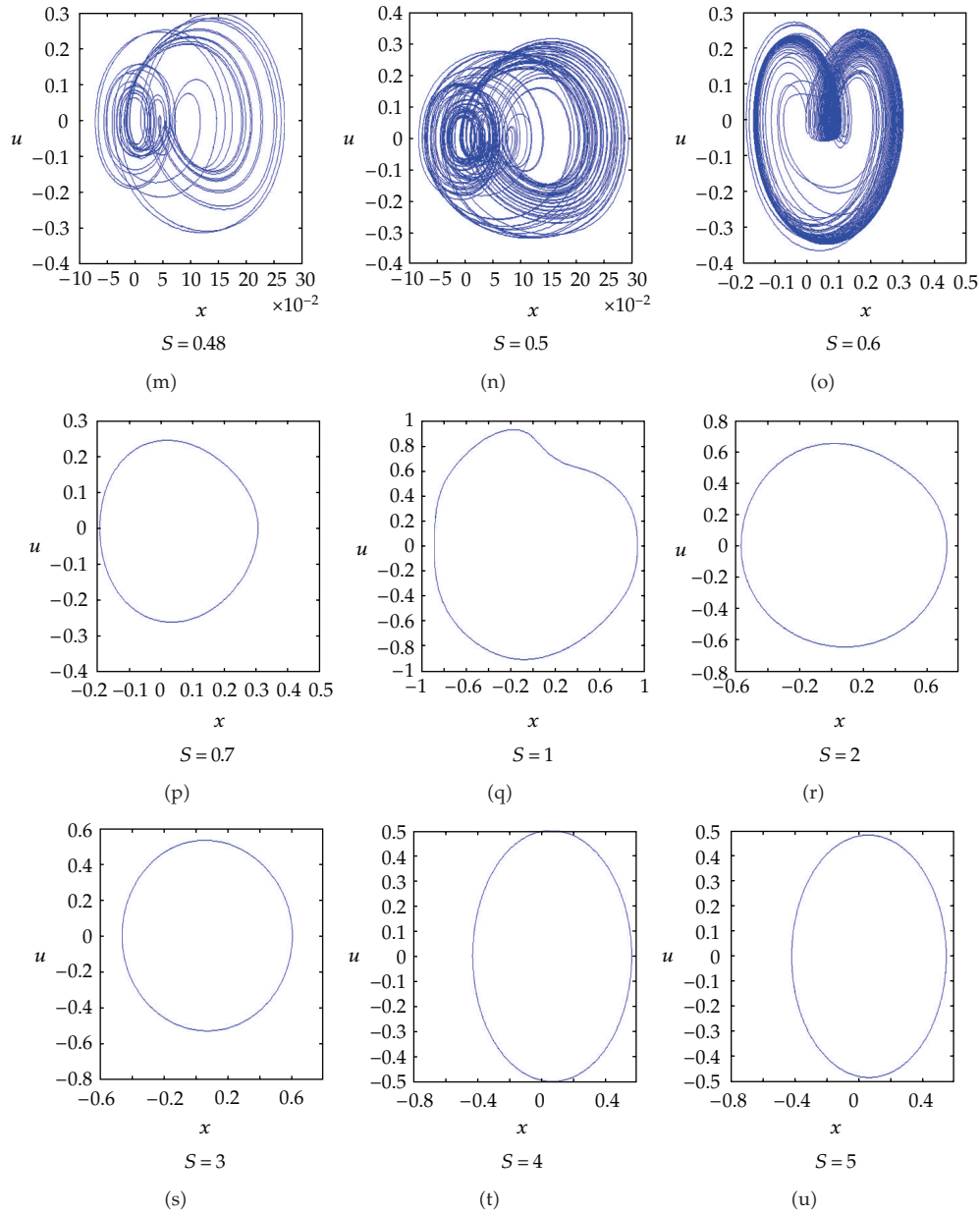


Figure 10: Phase diagrams for the route of subharmonic motion into chaos, out of chaos to periodic response at different rotating speed ratios of s (with rub-impact effect).

of the nonlinear dynamic response in this study. The dynamics of the system have been analyzed by reference to its dynamic trajectories, power spectra, Poincaré maps, bifurcation diagrams, maximum Lyapunov exponents, and fractal dimensions. The bifurcation results can be observed that HSFD may be used to improve dynamic irregularity. The system with rub-impact force effect may be a strongly nonlinear effect, and the bifurcation results show

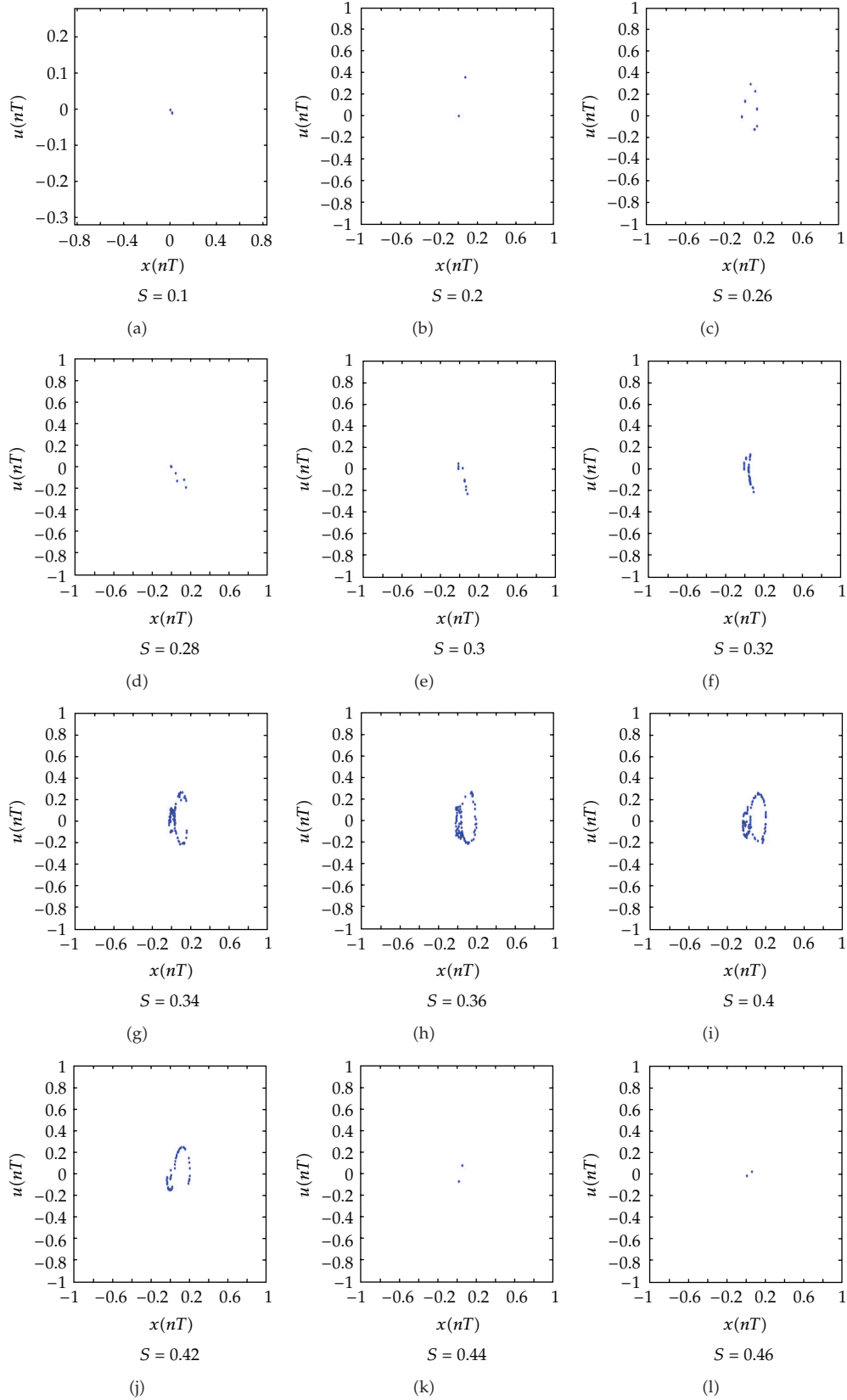


Figure 11: Continued.

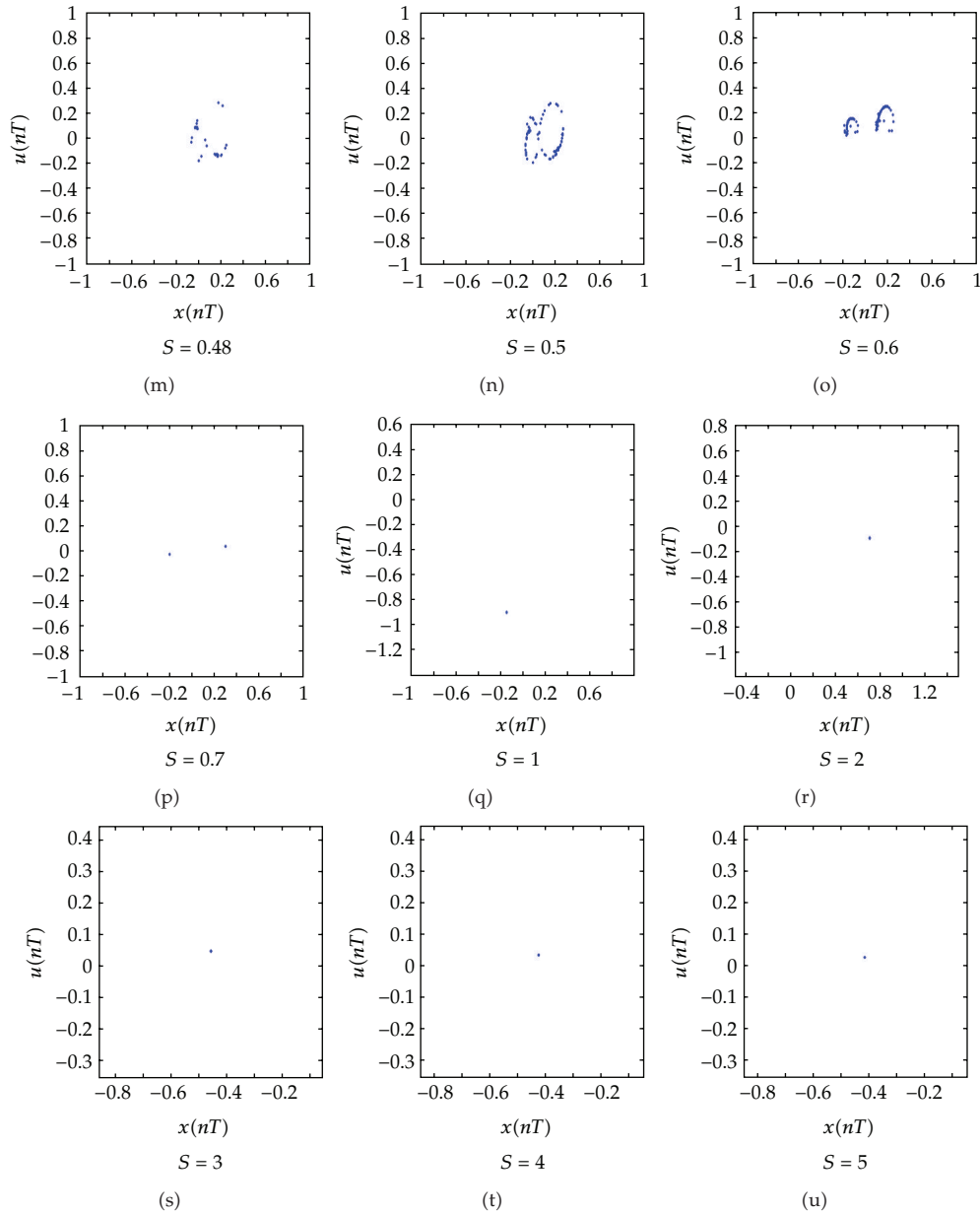


Figure 11: Poincaré maps for the route of subharmonic motion into chaos, out of chaos to periodic response at different rotating speed ratios of s (with rub-impact effect).

that HSFD mounted rotor-bearing system with rub-impact force effect present nonperiodic motions at low rotating speeds and perform periodic motions at high rotating speeds. The results will enable suitable values of the rotating speed ratio to be specified such that chaotic behavior can be avoided, thus reducing the amplitude of the vibration within the system and extending the system life.

Nomenclature

B :	Bearing parameter = $6\mu R^2 L^2 / m\delta^3 \omega_n$
d :	Viscous damping of the rotor disk
D :	$d / m\omega_n$
e :	Damper eccentricity = $\varepsilon\delta$
f_x, f_y :	Components of the fluid film force in horizontal and vertical coordinates
F_r, F_τ :	Components of the fluid film force in radial and tangential directions
h :	Oil film thickness, $h = \delta(1 + \varepsilon \cos \theta)$
k :	Stiffness of the retaining springs
k_d :	Proportional gain of PD controller
k_p :	Derivative gain of PD controller
L :	Bearing length
m :	Masses lumped at the rotor mid-point
O_m :	Center of rotor gravity
O_b, O_j :	Geometric center of the bearing and journal
$p(\theta)$:	Pressure distribution in the fluid film
p_s :	Pressure of supplying oil
$p_{c,i}$:	Pressure in the static pressure chamber
R :	Inner radius of the bearing housing
r :	Radius of the journal.
r, t :	Radial and tangent coordinates
s :	Speed parameter = ω / ω_n
U :	ρ / δ
x, y, z :	Horizontal, vertical and axial coordinates
x_0, y_0 :	Damper static displacements
X, Y, X_0, Y_0 :	$x/\delta, y/\delta, x_0/\delta, y_0/\delta$
ρ :	Mass eccentricity of the rotor
ϕ :	Rotational angle ($\phi = \omega t$)
ω :	Rotational speed of the shaft
φ_b :	Angle displacement of line $O_b O_j$ from the x -coordinate (see Figure 1)
Ω :	$\dot{\varphi}_b$
δ :	Radial clearance = $R - r$,
θ :	The angular position along the oil film from line $O_1 O_3$ (see Figure 1)
μ :	Oil dynamic viscosity
ε :	e / δ
β :	Distribution angle of static pressure region
$(\bullet), ('):$	Derivatives with respect to t and ϕ .

References

- [1] A. G. Holmes, C. M. M. Ettles, and I. W. Mayes, "Aperiodic behavior of a rigid shaft in short journal bearings," *International Journal for Numerical Methods in Engineering*, vol. 12, no. 4, pp. 695–702, 1978.
- [2] J. L. Nikolajsen and R. Holmes, "Investigation of squeeze-film isolators for the vibration control of a flexible rotor," *Journal of Mechanical Engineering Science*, vol. 21, no. 4, pp. 247–252, 1979.
- [3] J. E. H. Sykes and R. Holmes, "The effect of bearing misalignment on the non-linear vibration of aero-engine rotor-damper assemblies," *Proceedings of the Institution of Mechanical Engineers*, vol. 204, pp. 83–99, 1990.

- [4] Y. B. Kim and S. T. Noah, "Bifurcation analysis for a modified Jeffcott rotor with bearing clearances," *Nonlinear Dynamics*, vol. 1, no. 3, pp. 221–241, 1990.
- [5] F. F. Ehrich, "Some observations of chaotic vibration phenomena in high-speed rotordynamics," *ASME Journal of Vibration and Acoustics*, pp. 50–57, 1991.
- [6] J. Y. Zhao, I. W. Linnett, and L. J. McLean, "Subharmonic and quasi-periodic motions of an eccentric squeeze film damper-mounted rigid rotor," *Journal of Vibration and Acoustics*, vol. 116, no. 3, pp. 357–363, 1994.
- [7] R. D. Brown, P. Addison, and A. H. C. Chan, "Chaos in the unbalance response of journal bearings," *Nonlinear Dynamics*, vol. 5, no. 4, pp. 421–432, 1994.
- [8] G. Adiletta, A. R. Guido, and C. Rossi, "Chaotic motions of a rigid rotor in short journal bearings," *Nonlinear Dynamics*, vol. 10, no. 3, pp. 251–269, 1996.
- [9] G. Adiletta, A. R. Guido, and C. Rossi, "Nonlinear dynamics of a rigid unbalanced rotor in journal bearings. Part I: theoretical analysis," *Nonlinear Dynamics*, vol. 14, no. 1, pp. 57–87, 1997.
- [10] G. Adiletta, A. R. Guido, and C. Rossi, "Nonlinear dynamics of a rigid unbalanced rotor in journal bearings. Part II: experimental analysis," *Nonlinear Dynamics*, vol. 14, no. 2, pp. 157–189, 1997.
- [11] P. Sundararajan and S. T. Noah, "Dynamics of forced nonlinear systems using shooting/Arc-length continuation method—application to rotor systems," *Journal of Vibration and Acoustics*, vol. 119, no. 1, pp. 9–20, 1997.
- [12] C. W. Chang-Jian and C. K. Chen, "Chaos and bifurcation of a flexible rub-impact rotor supported by oil film bearings with nonlinear suspension," *Mechanism and Machine Theory*, vol. 42, no. 3, pp. 312–333, 2007.
- [13] C. W. Chang-Jian and C. K. Chen, "Bifurcation and chaos of a flexible rotor supported by turbulent journal bearings with non-linear suspension," *Proceedings of the Institution of Mechanical Engineers J*, vol. 22, no. 6, pp. 549–561, 2006.
- [14] C. W. Chang-Jian and C. K. Chen, "Nonlinear dynamic analysis of a flexible rotor supported by micropolar fluid film journal bearings," *International Journal of Engineering Science*, vol. 44, no. 15–16, pp. 1050–1070, 2006.
- [15] C. W. Chang-Jian and C. K. Chen, "Bifurcation and chaos analysis of a flexible rotor supported by turbulent long journal bearings," *Chaos, Solitons and Fractals*, vol. 34, no. 4, pp. 1160–1179, 2007.
- [16] C. W. Chang-Jian and C. K. Chen, "Nonlinear numerical analysis of a flexible rotor equipped with squeeze couple stress fluid film journal bearings," *Acta Mechanica Solida Sinica*, vol. 20, no. 4, pp. 309–316, 2007.
- [17] Z. P. Li, Y. G. Luo, H. L. Yao, and B. C. Wen, "Nonlinear dynamic study of the elastic rotor-bearing system with rub-impact fault," *Journal of Northeastern University*, vol. 23, no. 10, pp. 980–983, 2002.
- [18] C. L. Chen and H. T. Yau, "Subharmonic and chaotic motions of a hybrid squeeze-film damper-mounted rigid rotor with active control," *Journal of Vibration and Acoustics*, vol. 124, no. 2, pp. 198–208, 2002.
- [19] P. Grassberger and I. Procaccia, "Characterization of strange attractors," *Physical Review Letters*, vol. 50, no. 5, pp. 346–349, 1983.

Research Article

Nonsmooth Recursive Identification of Sandwich Systems with Backlash-Like Hysteresis

Ruili Dong,¹ Yonghong Tan,¹ Hui Chen,² and Yangqiu Xie³

¹ College of Mechanical and Electronic Engineering, Shanghai Normal University, Shanghai 200234, China

² Department of Automation, Southeast University, Nanjing 210096, China

³ School of Electronic Engineering, Xi'dian University, Xi'an 710071, China

Correspondence should be addressed to Yonghong Tan, tany@shnu.edu.cn

Received 29 March 2012; Revised 10 June 2012; Accepted 10 June 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 Ruili Dong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A recursive gradient identification algorithm based on the bundle method for sandwich systems with backlash-like hysteresis is presented in this paper. In this method, a dynamic parameter estimation scheme based on a subgradient is developed to handle the nonsmooth problem caused by the backlash embedded in the system. The search direction of the algorithm is estimated based on the so-called bundle method. Then, the convergence of the algorithm is discussed. After that, simulation results on a nonsmooth sandwich system are presented to validate the proposed estimation algorithm. Finally, the application of the proposed method to an X-Y moving positioning stage is illustrated.

1. Introduction

Usually, a sandwich system with backlash-like hysteresis is defined as the system that a backlash-like hysteresis is sandwiched between two linear dynamic subsystems. In engineering applications, many mechanical systems such as mechanical transmission systems, servo control systems, and hydraulic valve systems can be described by the so-called sandwich systems with backlash-like hysteresis. The reason to cause the backlash-like hysteresis phenomenon is mainly due to the gaps existing in transmission mechanism systems such as gearbox and ball screw.

Recently, identification of sandwich systems has become one of the interesting issues in the domain of modeling and control for complex systems. References [1–3] proposed the recursive identification methods for the sandwich system with smooth nonlinearities. The main ideas of those approaches are to extend the linear system identification methods to smooth nonlinear cases. Moreover, there have been some methods for the identification of

Hammerstein or Wiener systems with backlash-like hysteresis [4–8], most of which are the modified linear system identification methods.

However, until today, there have been very few publications concerning the identification of the sandwich systems with backlash-like hysteresis. Reference [9] proposed a method to identify the sandwich systems with backlash-like hysteresis, but the approach is still based on idea to extend the linear system identification method to nonlinear cases. On the other hand, the switching functions in that method have significant influence on the convergence speed of the algorithm.

In this paper, a recursive gradient algorithm based on the bundle method is proposed to identify parameters of the sandwich model. In this algorithm, the effect of the nonsmoothness caused by the backlash-like hysteresis in sandwich system is considered. In order to obtain the optimizing search direction at the nonsmooth points of the system, the Clarke subgradient technique is utilized based on the idea of the bundle method [10–12]. By comparing with the above-mentioned available methods, the proposed method employs the nonsmooth optimization technique to identify the nonsmooth sandwich systems with backlash-like hysteresis. Thus, it will provide us with a new approach for dealing with on-line modeling of nonsmooth dynamic systems. A numerical example will be presented to evaluate the performance of the proposed approach. Finally, experimental results on an X-Y moving positioning stage are illustrated.

2. Brief Description of Sandwich Systems with Backlash

The structure of a sandwich system with backlash-like hysteresis is shown in Figure 1, in which a backlash-like hysteresis is embedded between the input and output linear subsystems, that is, $L_1(\cdot)$ and $L_2(\cdot)$. It is assumed that input $u(k)$ and output $y(k)$ can be measured directly, but the internal variables $x(k)$ and $v(k)$ are not measurable.

Suppose that both linear subsystems are stable, and the time delays q_1 and q_2 in $L_1(\cdot)$ and $L_2(\cdot)$ are known, respectively. The corresponding discrete-time models of $L_1(\cdot)$ and $L_2(\cdot)$ are, respectively, written as

$$\begin{aligned} x(k) &= -\sum_{i_2=1}^{n_a} a_{i_2} x(k-i_2) + \sum_{j_2=0}^{n_b} b_{j_2} u(k-q_1-j_2), \\ y(k) &= -\sum_{i_1=1}^{n_c} c_{i_1} y(k-i_1) + \sum_{j_1=0}^{n_d} d_{j_1} v(k-q_2-j_1), \end{aligned} \quad (2.1)$$

where n_a and n_b are the orders of $L_1(\cdot)$, q_1 is the time delay, and a_{i_2} as well as b_{j_2} are the coefficients of $L_1(\cdot)$; n_c and n_d are the orders of $L_2(\cdot)$, q_2 is the time delay, and c_{i_1} and d_{j_1} are the coefficients of $L_2(\cdot)$. Let both b_0 and d_0 be equal to unity for unique representation.

Note that the backlash-like hysteresis shown in Figure 1 is specified by the slopes m_1 and m_2 as well as the absolute thresholds, D_1 and D_2 , where $0 < m_1 < \infty$, $0 < m_2 < \infty$,

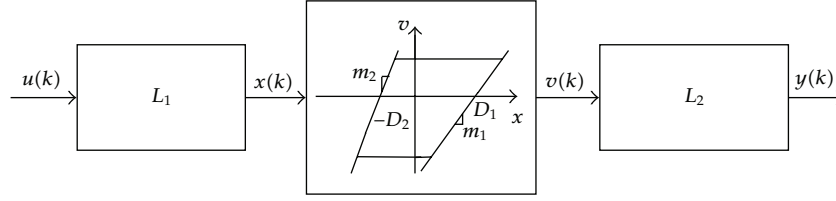


Figure 1: The structure of the sandwich system with backlash-like hysteresis.

$0 < D_1 < \infty$, and $0 < D_2 < \infty$. Hence, the discrete-time model of the backlash-like hysteresis is described as

$$v(k) = \begin{cases} m_1(x(k) - D_1), & x(k) > \frac{v(k-1)}{m_1} + D_1, \quad x(k) > x(k-1), \text{ increase zone,} \\ v(k-1), & \frac{v(k-1)}{m_2} - D_2 \leq x(k) \leq \frac{v(k-1)}{m_1} + D_1, \text{ memory zone,} \\ m_2(x(k) + D_2), & x(k) < \frac{v(k-1)}{m_2} - D_2, \quad x(k) < x(k-1), \text{ decrease zone.} \end{cases} \quad (2.2)$$

For the convenience to describe the system, the discrete-time model of the backlash-like hysteresis can be rewritten as

$$\begin{aligned} m(k) &= m_1 + (m_2 - m_1)g(k), \\ v1(k) &= m(k)(x(k) + g(k)x(k) - D_1g_1(k) + D_2g_2(k)), \\ v(k) &= v1(k) + [v(k-1) - v1(k)](g_1(k) - 1)(g_2(k) - 1), \end{aligned} \quad (2.3)$$

where the switching functions $g(k)$, $g_1(k)$, and $g_2(k)$ are, respectively, defined as

$$\begin{aligned} g(k) &= \begin{cases} 0, & \Delta x(k) > 0 \\ 1, & \Delta x(k) \leq 0, \end{cases} \\ g_1(k) &= \begin{cases} 1, & x(k) > \frac{v(k-1)}{m_1} + D_1, \quad x(k) > x(k-1), \\ 0, & \text{else,} \end{cases} \\ g_2(k) &= \begin{cases} 1, & x(k) < \frac{v(k-1)}{m_2} - D_2, \quad x(k) < x(k-1), \\ 0, & \text{else,} \end{cases} \end{aligned} \quad (2.4)$$

where $\Delta x(k) = x(k) - x(k-1)$.

Thus, (2.1)–(2.3) present the model to describe the sandwich system with backlash-like hysteresis. Hence, the unknown parameter vector of the model can be written as $\theta \in R^{n_a+n_b+n_c+n_d+4}$, where

$$\theta = [c_1, \dots, c_{n_c}, a_1, \dots, a_{n_a}, m_1, m_2, D_1, D_2, b_1, \dots, b_{n_b}, d_1, \dots, d_{n_d}]^T. \quad (2.5)$$

According to concept of the gradient algorithm, define the objective function as

$$Q(k, \hat{\theta}(k)) = \sum_{k=1}^n \frac{[y(k) - \hat{y}(k, \hat{\theta}(k))]^2}{2} = \frac{1}{2} \sum_{k=1}^n f(k, \hat{\theta}(k)), \quad (2.6)$$

where $\hat{\theta}$ is the estimate of θ , and $\hat{y}(k, \hat{\theta}(k))$ is the output of system model. The optimal estimate of $\hat{\theta}$ can be obtained by minimizing the above-mentioned criterion.

3. The Nonsmooth Estimation of the Sandwich Model with Backlash-Like Hysteresis

In this section, a gradient-based identification algorithm is proposed for identification of the sandwich system with backlash-like hysteresis. Due to the nonsmoothness of the backlash, the gradients of the system output with respect to the parameters of the backlash at nonsmooth points will not exist. The smooth gradient-based methods directly applied to nonsmooth systems may fail in convergence [13]. On the other hand, the genetic algorithms [14] or Powell's method [15], which are based on derivative-free techniques, may be unreliable and become inefficient when the system structure is complicated. Thus, we should find a special way for solving this problem. The simplest way to solve the problem is to apply the Clarke subgradients [11] to the approximation of the gradients at the nonsmooth points.

The basic idea of the bundle method is to approximate the subdifferential of $Q(k, \hat{\theta}(k))$ with respect to $\hat{\theta}(k)$ by gathering the subgradients from previous iterations into a bundle for the nonsmooth objective function $Q(k, \hat{\theta}(k))$. The gradient $\nabla Q(k, \hat{\theta}(k))$ can change discontinuously, and some change of the gradient may not be small in the neighborhood of the minimum of the function. So the values of $Q(k, \hat{\theta}(k))$ and $\partial Q(k, \hat{\theta}(k))$ at a single point $\hat{\theta}(k)$ do not offer sufficient information of the local behavior of $Q(k, \hat{\theta}(k))$. The detail of the bundle method can be found in [10–12] and reference therein.

Considering that the sandwich system with backlash-like hysteresis is locally Lipschitz continuous, we have the following definition.

Definition 3.1 (see [11]). Let $F: R^n \times R \rightarrow R$ be locally Lipschitz continuous. This allows one to define a Clarke subgradient of F at ξ as $dF(\xi)$:

$$dF(\xi) \in \partial F(\xi), \quad \text{subject to } \partial F(\xi) = \text{conv} \left\{ \nabla F(\xi^i) \mid \xi^i \rightarrow \xi, \nabla F(\xi^i) \text{ exists} \right\}, \quad (3.1)$$

where “conv” denotes the convex hull of a set.

The set of all the Clarke subgradients is the Clarke subdifferential of F at ξ which is denoted by $\partial F(\xi)$ [11].

Considering that backlash-like hysteresis is a nonsmooth mapping, the gradients of parameters in $L_1(\cdot)$ with respect to $v(k)$ do not exist at a nonsmooth point. Hence, we define the parameters of the backlash-like hysteresis and $L_1(\cdot)$ as $\sigma = \{m_1, m_2, D_1, D_2, a_1 \cdots a_{na}, b_1 \cdots b_{nb}\} \in R^{4+na+nb}$. Considering the cost function described by (2.6), the gradients of $f(\cdot)$ with respect to σ will not exist at the nonsmooth points. Hence,

at the nonsmooth points of $Q(\cdot)$, the Clarke subdifferential of $f(\cdot)$ with respect to σ , that is, $\partial f(\sigma)$, can be obtained by

$$\partial f(\sigma) = -\text{conv} \left\{ \left[y(k) - \hat{y}(k, \hat{\theta}(k)) \right] \left(\sum_{j_1=0}^{n_b} d_{j_1} \partial \hat{v}_{(k-j_1-q_2)}(\sigma) \right) \right\}, \quad (3.2)$$

where $\partial \hat{v}_{(k-j_1-q_2)}(\sigma) = \text{conv} \{ \nabla \hat{v}_{(k-j_1-q_2)}(\sigma) \}$, and $\nabla \hat{v}_{(k-j_1-q_2)}(\sigma)$ is the gradient of \hat{v} , the output of backlash-like hysteresis, with respect to σ at the smooth points. Thus, the corresponding gradients of $\hat{v}(k - j_1 - q_2)$ with respect to σ at the smooth points are

$$\nabla \hat{v}_{(k-j_1-q_2)}(\sigma) = \begin{cases} \begin{bmatrix} u(k - j_1 - q_2 - q_1) - \hat{D}_1(k-1), 0, \\ -\hat{m}_1(k-1), 0, -\hat{m}_1(k-1)\hat{x}(k-1-j_1-q_2), \dots, \\ -\hat{m}_1(k-1)\hat{x}(k-n_a-j_1-q_2), \\ \hat{m}_1(k-1)u(k-1-j_1-q_2-q_1), \dots, \\ \hat{m}_1(k-1)u(k-n_b-j_1-q_2-q_1) \end{bmatrix}^T, & \text{in increase zones,} \\ [0, 0, 0, 0, 0, \dots, 0, 0, \dots, 0]^T, & \text{in memory zones,} \\ \begin{bmatrix} 0, u(k - j_1 - q_2 - q_1) + \hat{D}_2(k-1), 0, \\ \hat{m}_2(k-1), -\hat{m}_2(k-1)\hat{x}(k-1-j_1-q_2), \dots, \\ -\hat{m}_2(k-1)\hat{x}(k-n_a-j_1-q_2), \\ \hat{m}_2(k-1)u(k-1-j_1-q_2-q_1), \dots, \hat{m}_2(k-1) \\ u(k-n_b-j_1-q_2-q_1) \end{bmatrix}^T, & \text{in decrease zones,} \end{cases} \quad (3.3)$$

where $\hat{x}(k) = -\sum_{i_2=1}^{n_a} \hat{a}_{i_2} \hat{x}(k-i_2) + \sum_{j_2=0}^{n_b} \hat{b}_{j_2} u(k-q_1-j_2)$, and the coefficients \hat{a}_{i_2} and \hat{b}_{j_2} are the corresponding estimated values at the previous step.

Hence, based on (3.2) and (3.3), the Clarke subdifferential of $f(\cdot)$ with respect to σ can be obtained at nonsmooth points of the system. Besides, as $L_2(\cdot)$ is a smooth function, the gradients of $f(\cdot)$ with respect to the parameters of the linear subsystems $L_2(\cdot)$ always exist. So, the Clarke subdifferential of $f(\cdot)$ with respect to all the unknown parameters of the sandwich system can be determined.

The proper Clarke subgradient direction $\mathbf{t}(k, \hat{\theta}(k))$ of $f(\cdot)$ with respect to the parameters to be estimated at nonsmooth points can be derived based on

$$\begin{aligned} \min_{\varphi, \mathbf{d}} \quad & \left(\varphi(k) + \frac{1}{2} \left\| \mathbf{t}(k, \hat{\theta}(k)) \right\|^2 \right) \\ \text{s.t.} \quad & -\beta_j(k) + \left\langle \mathbf{h}_j(k), \mathbf{t}(k, \hat{\theta}(k)) \right\rangle \leq \varphi(k), \quad \forall j \in J_k, \end{aligned} \quad (3.4)$$

where $\|\cdot\|$ denotes the Euclidean norm; J_k is a nonempty subset of $\{1, \dots, k\}$; set $\varphi(k)$ is the predicted amount of descent; $\mathbf{h}_j(k) \in \partial f(k, \hat{\theta}_j^*(k))$ for $j \in J_k$, and $\hat{\theta}_j^*(k)$ are some trail points (from the past iterations); $\beta_j(k) = \max\{|\alpha_j(k)|, \gamma(s_j(k))^2\}$ is the locality measure of subgradient; $\gamma \geq 0$ is the distance measure parameter ($\gamma = 0$ if $f(k, \hat{\theta}(k))$ is convex),

$\alpha_j(k) = f(k, \hat{\theta}(k)) - f(k, \hat{\theta}_j^*(k)) - \mathbf{h}_j(k)(\hat{\theta}(k) - \hat{\theta}_j^*(k))$ is the linearization error; $s_j(k) = \|\hat{\theta}_j(k) - \hat{\theta}_j^*(k)\| + \sum_{i=j}^{k-1} \|\hat{\theta}_{i+1}(k) - \hat{\theta}_i(k)\|$ is the distance measure to estimate $\|\hat{\theta}(k) - \hat{\theta}_j^*(k)\|$ without the requirement to store the trial point $\hat{\theta}^*(k)$.

According to formula (3.4), $\mathbf{t}(k)$ and $\varphi(k)$ are obtained, that is,

$$\mathbf{t}(k, \hat{\theta}(k)) = -\sum_{j \in J_k} \lambda_j^k \mathbf{h}_j(k) = \sum_{j \in J_k} [y(k) - \hat{y}(k, \hat{\theta}_j(k))] \lambda_j^k \mathbf{w}_j(k) = e(k, \hat{\theta}(k)) \bar{\mathbf{h}}_j(k), \quad (3.5)$$

$$\varphi(k) = -\|t(k, \hat{\theta}(k))\|^2 - \sum_{j \in J_k} \lambda_j^k \beta_j(k), \quad (3.6)$$

where $\mathbf{w}_j(k) = \partial \hat{y}(k, \hat{\theta}) / \partial \theta|_{\hat{\theta}=\hat{\theta}_j(k)}$, $e(k, \hat{\theta}(k)) = y(k) - \hat{y}(k, \hat{\theta}_j(k))$, $\bar{\mathbf{h}}_j(k) = \sum_{j \in J_k} \lambda_j^k \mathbf{w}_j(k)$, $\lambda_j^k \geq 0$, and $\sum_{j \in J_k} \lambda_j^k = 1$.

Remark 3.2. If $f(\cdot)$ is convex, the model $f(k, \hat{\theta}(k))$ is an underestimate for $f(\cdot)$, and the nonnegative linearization error $\alpha_j(k)$ measures the performance of an approximation of the model to the original cost function. If $f(\cdot)$ is nonconvex, these facts are not valid anymore because $\alpha_j(k)$ may have a small or even negative value, although the trial point $\hat{\theta}_j^*(k)$ locates far away from the current iteration point $\hat{\theta}(k)$, and thus, the corresponding subgradient $\mathbf{h}_j(k)$ is worthless. For these reasons, the locality measure of subgradient $\beta_j(k)$ is introduced.

Therefore, the proposed recursive gradient estimation algorithm based on bundle method for the sandwich model with backlash-like hysteresis is shown as follows.

Step 1. Select starting point $\theta_0 \in R^{n_a+n_b+n_c+n_d+4}$ and stopping parameter $\delta > 0$. Calculate $f(k, \theta_0)$ and vector $\mathbf{h}_j(k) \in \partial f(k, \theta_0)$, where $j \in J_k$, $J_k = \{k_0\}$, $|J_k| \leq k_1$, $|J_k|$ is the element number of J_k , and k_1 is a given positive number. Set $\beta_j(k) = 0$, $k = k_0$ and the line search parameters

$$q \in (0, 0.5), \quad q^* \in (q, 1), \quad \eta(0) \in (0, 1]. \quad (3.7)$$

Step 2. Calculate optimal solution $(\varphi(k), \mathbf{t}(k, \hat{\theta}(k)))$ based on formulas (3.2)–(3.6). If $\varphi(k) \geq -\delta$, then stop.

Step 3. Search for the largest step size $\eta(k) \in [0, 1]$ such that $\eta(k) \geq \eta(0)$ and if

$$f(k, \hat{\theta}(k) + \eta(k)\mathbf{t}(k)) \leq f(k, \hat{\theta}(k)) + q\eta(k)\varphi(k), \quad (3.8)$$

it holds

$$\varphi(k) = f(k, \hat{\theta}(k) + \mathbf{t}(k)) - f(k, \hat{\theta}(k)) < 0. \quad (3.9)$$

Then, we take a long step and set $\hat{\theta}(k+1) = \hat{\theta}(k) + \eta(k)\mathbf{t}(k)$ and $\hat{\theta}^*(k+1) = \hat{\theta}(k+1)$; go to Step 4.

Otherwise, if $0 < \eta(k) < \eta(0)$, and formula (3.8) holds, then we take a short step and set $\hat{\theta}(k+1) = \hat{\theta}(k) + \eta(k)\mathbf{t}(k)$, and $\hat{\theta}^*(k+1) = \hat{\theta}(k) + \eta^*(k)\mathbf{t}(k)$ where $\eta^*(k) > \eta(k)$. Go to Step 5.

If $\eta(k) = 0$, and formula (3.8) holds, we take a null step, and namely set $\hat{\theta}(k+1) = \hat{\theta}(k)$ and $\hat{\theta}^*(k+1) = \hat{\theta}(k) + \eta^*(k)t(k)$; go to Step 5.

Step 4. Let $J_k = J_k \cup \{k+1\}$, $k = k+1$; if $k \leq k_1$, then $J_k = \{1, \dots, k\}$, and if $k > k_1$, then $J_k = J_{k-1} \cup \{k\} \setminus \{k-k_1\}$, then go to Step 2.

Step 5. $J_k = J_k \cup \{k+1\}$, $k = k+1$; if $k \leq k_1$, then $J_k = \{1, \dots, k\}$, and if $k > k_1$, then $J_k = J_{k-1} \cup \{k\} \setminus \{k-k_1\}$, and the proper Clarke subgradient $\mathbf{h}_j(k)$ satisfies

$$-\beta_j(k, \hat{\theta}(k)) + \mathbf{h}_j^T(k, \hat{\theta}(k))\mathbf{t}(k, \hat{\theta}(k-1)) \geq q^*\varphi(k, \hat{\theta}(k-1)), \quad (3.10)$$

then go to Step 2.

Remark 3.3. In long step, there is an obvious decrease in the value of the objective function. Hence, it is unnecessary to detect discontinuities in the gradient of $f(\cdot)$. Thus, we just set $\mathbf{h}_j(k) \in \partial f(k, \hat{\theta}(k))$. On the other hand, in short steps and null steps, the gradient of $f(\cdot)$ is discontinuous. Then, based on (3.10), both $\hat{\theta}(k)$ and $\hat{\theta}^*(k)$ located on the opposite sides of this discontinuity are guaranteed, and the new subgradient $\mathbf{h}_j(k) \in \partial f(k, \hat{\theta}^*(k))$ will force an obvious modification of the next search direction. Hence, the algorithm approximates the effectively searching direction at nonsmooth points based on the bundle method, which cannot be realized by the smooth optimization techniques.

Remark 3.4. If the value of $\eta(0)$ is too small, the convergence speed will be very sluggish, while $\eta(0)$ is too large, and the algorithm may not be convergent. Hence, it is important for $\eta(0)$ to be chosen properly. Usually, $\eta(0)$ is chosen based on an empirical method.

Remark 3.5. If all the Clarke subgradients are included in J_k , the corresponding storage capacity is infinite. Hence, the number of the subgradients in J_k must be constrained. In the proposed algorithm, we give the upper bound of $|J_k| \leq k_1$, and the upper bound k_1 is specified by empirical method.

4. Convergence of the Estimation

For the convergence of the above-mentioned estimation algorithm, we have the following

Theorem 4.1. Suppose that $\eta(k)$ and $\beta_j(k)$ satisfy

$$0 \leq \eta(k) \leq \frac{2e^2(k)\bar{\mathbf{h}}(k)\bar{\mathbf{h}}^T(k) - \beta_j(k)}{e^2(k)\bar{\mathbf{h}}^T(k)\bar{\mathbf{h}}(k)\left(1 + \left[\bar{\mathbf{h}}(k)\bar{\mathbf{h}}^T(k)\right]\right)}, \quad (4.1)$$

$$\beta_j(k) \leq 2e^2(k)\left[\bar{\mathbf{h}}(k)\bar{\mathbf{h}}^T(k)\right], \quad (4.2)$$

respectively, then the parameters θ can be convergent to a local optimal value.

Proof. The proof of this theorem can be found in Appendix. \square

5. Simulation

The proposed approach is used to identify a numerical sandwich system with backlash-like hysteresis based on the measured system input and output. Suppose that the parameters of the backlash-like hysteresis in the system are $m_1 = 1$, $m_2 = 1.2$, $D_1 = 0.5$, and $D_2 = 0.6$. The linear subsystems $L_1(\cdot)$ and $L_2(\cdot)$ are

$$x(k) = -0.1x(k-1) - 0.2x(k-2) + 1.5u(k-1), \quad (5.1)$$

$$y(k) = -1.2y(k-1) - 0.32y(k-2) + 2v(k-1) - 0.1v(k-2), \quad (5.2)$$

respectively.

That implies $a_1 = 0.1$, $a_2 = 0.2$, $b_0 = 1.5$, $c_1 = 1.2$, $c_2 = 0.32$, $d_0 = 2$, and $d_1 = -0.1$. In the simulation, both b_0 and d_0 are assumed to be equal to unity for model uniqueness, which implies that the corresponding equivalent true values of the coefficients are $\tilde{a}_1 = 0.1$, $\tilde{a}_2 = 0.2$, $\tilde{c}_1 = 1.2$, $\tilde{c}_2 = 0.32$, $\tilde{d}_1 = -0.05$, $\tilde{m}_1 = 3$, $\tilde{m}_2 = 3.6$, $\tilde{D}_1 = 0.33$, and $\tilde{D}_2 = 0.4$, respectively, but this does not affect the properties of the whole system.

In the simulation, the signal to excite the system is a random sequence with variance $\sigma^2 = 0.49$. Choose $\delta = 1.0 \times 10^{-4}$. In the proposed algorithm, based on Remark 3.4, select $\eta(0) = 0.015$, $k_1 = 6$, $\theta_0 = [0, 0, 0, 0, 0.1, 0.1, 0.1, 0.1, 0]^T$, and $\beta_1(k_0) = 0$, respectively. For comparison, the traditional gradient method is also used to estimate the parameters of the system. In this method, the nonsmooth points of the system are omitted for the gradients of the system do not exist at nonsmooth points. The initialized values of the parameters are the same as those used in the proposed method. The optimizing step is chosen as 0.009.

Figure 2 illustrates the comparison of the estimated parameter convergence procedures between the proposed method and the traditional gradient method. In Figure 2, blue and solid lines denote the convergence procedures of the parameters estimated by the proposed method, while red and dotted lines show the convergence procedures of the parameters determined by the traditional gradient method. From Figure 2, we note that the parameters of the backlash-like hysteresis converge slower than those of the linear submodels especially the input linear submodel. Moreover, the proposed method has achieved faster convergence than that of the traditional gradient method. It is noticed that the oscillation and sharp jumps happened in the estimation procedure of the traditional gradient approach.

In the case that the system is affected by random noise, the proposed strategy can still obtain better convergence of parameter estimation. In the simulation with noise, the signal to noise rate (SNR) is equal to 46.5. All the initial values of the parameters are the same as those in the noise-free case.

Figure 3 shows the comparison of the convergence procedures of the estimated parameters in the case with noise between the proposed method and the traditional gradient approach. Similar to the noise-free case, the blue and solid lines denote the convergence procedures of the parameters estimated by the proposed method, while the red and dotted lines show the convergence procedures of the parameters estimated by the traditional gradient method. Obviously, the proposed method has obtained faster convergent results than the traditional gradient method.

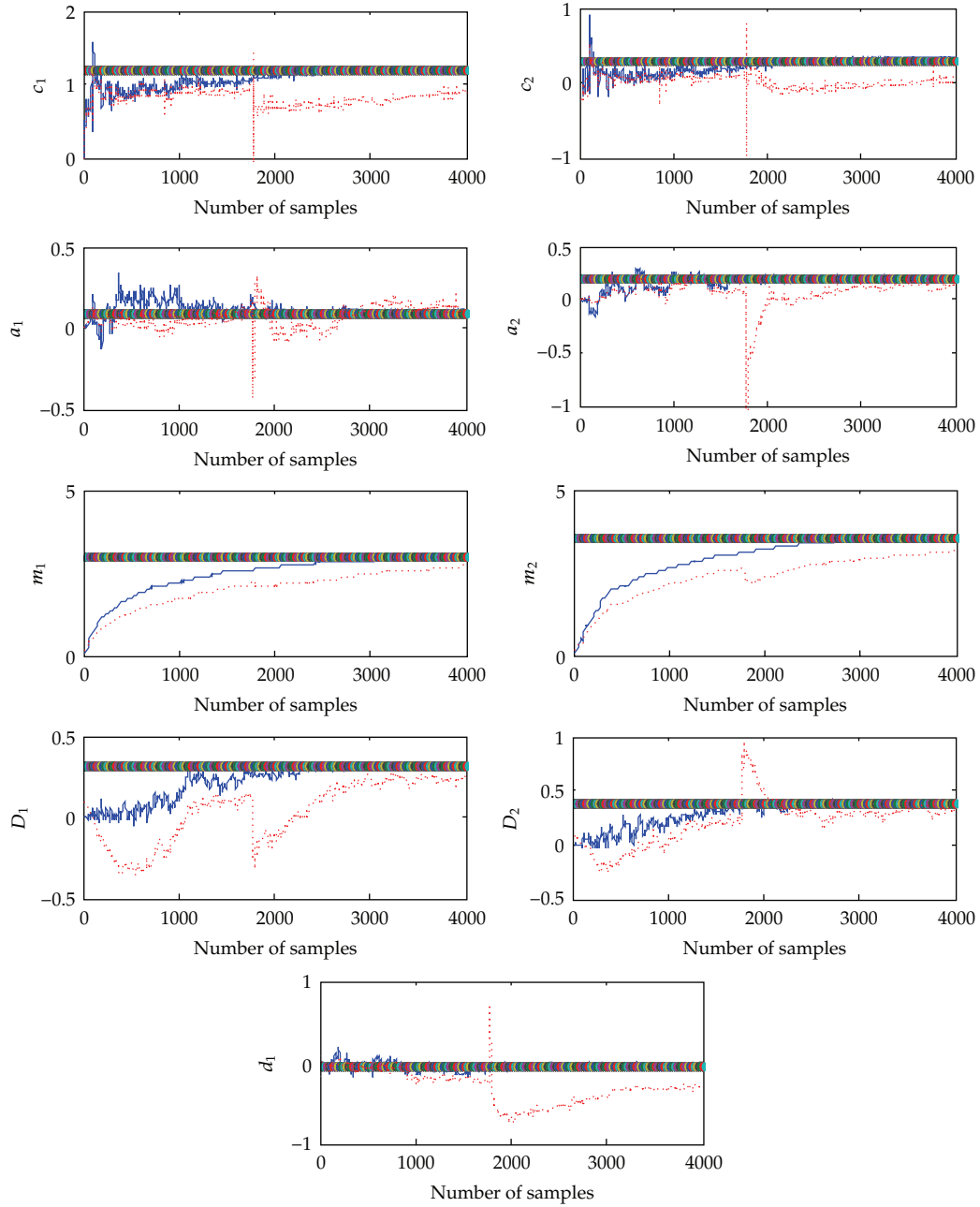


Figure 2: The convergence of the estimated parameters (noise-free case): proposed method: blue and solid line; traditional gradient method: red and dotted lines.

6. Application to an X-Y Moving Positioning Stage

The proposed identification approach is also applied to the modeling of an X-Y moving positioning stage with the architecture shown in Figure 4. In this equipment, the movement of the work platform of each axis is driven by a DC servomotor through a ball-screw-nut

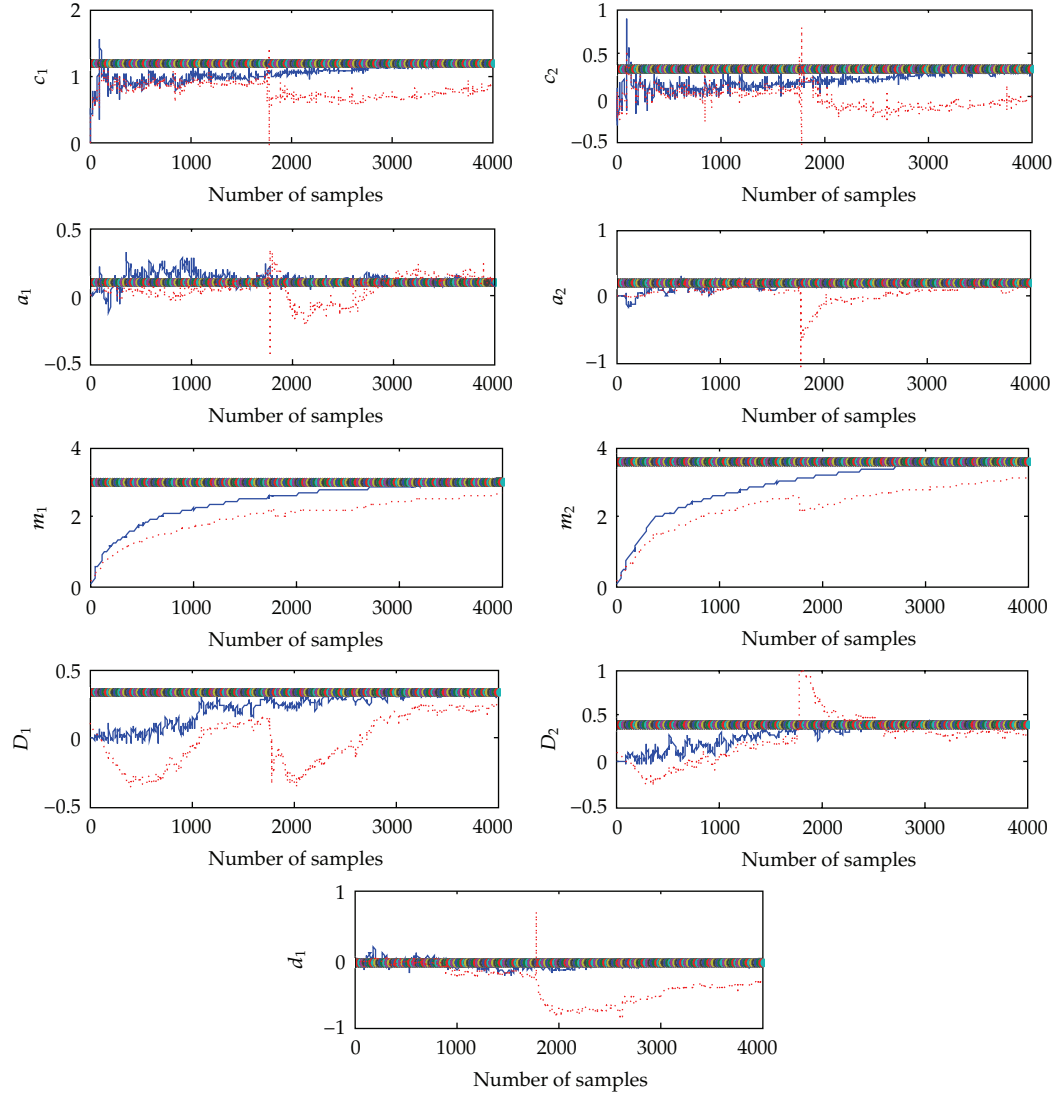


Figure 3: The convergence of the estimated parameters (noisy case): proposed method: blue and solid lines; traditional gradient method: red and dotted line.

mechanism which transforms the rotational shaft movement into linear displacement. The servomotor is controlled by a digital signal processor (TMS320LF-2407A). The displacement of each axis is measured by a linear encoder (RGF2000H125B). The signals of both phase A and phase B encoders are decoded by a quadrature decoding circuit which is based on the decoding chip (Agilent HCTL-2020).

In this system, the servomotor can be considered as a second-order linear dynamic subsystem. The movement of the work platform is also described by a linear second-order dynamic model. Due to the inherent characteristic, both dead zone and backlash-like hysteresis exist in this system. In order to simplify the identification procedure, the dead zone is compensated by a dead zone inverse model-based compensator. Thus, in the identification, only the effect of backlash-like hysteresis existing in the ball-screw-nut mechanism will

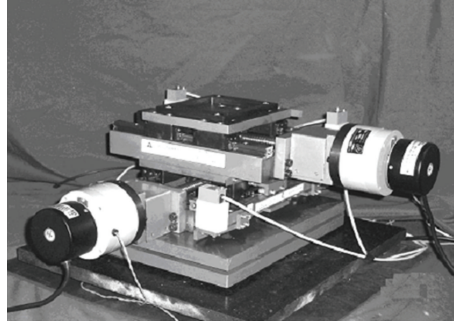


Figure 4: The X-Y moving positioning stage.

be considered. Therefore, the identified system is actually a typical sandwich system with backlash-like hysteresis. In this section, only the identification procedure of axis A will be presented due to the limited space. The corresponding models used to describe the behavior of axis A are shown as follows:

(1) the input linear model (L_1):

$$x(k) = -a_1x(k-1) - a_2x(k-2) + b_0u(k-1), \quad (6.1)$$

(2) the model of the backlash-like hysteresis:

$$v(k) = \begin{cases} m_1(x(k) - D_1), & x(k) > \frac{v(k-1)}{m_1} + D_1, \quad x(k) > x(k-1), \\ v(k-1), & \frac{v(k-1)}{m_2} - D_2 \leq x(k) \leq \frac{v(k-1)}{m_1} + D_1, \\ m_2(x(k) + D_2), & x(k) < \frac{v(k-1)}{m_2} - D_2, \quad x(k) < x(k-1), \end{cases} \quad (6.2)$$

(3) the output linear model (L_2):

$$y(k) = -c_1y(k-1) - c_2y(k-2) + d_0v(k) + d_1v(k-1), \quad (6.3)$$

where $y(k)$ is the moving speed of the work platform.

Based on the operating requirement, a sequence of square wave plus sinusoidal wave is used to excite the system within the operating range. The corresponding amplitude of the input varies in the range between -1.09 V and 1.05 V, and the sample period is 0.5 ms.

In this model, both b_0 and d_0 are set to one. The initial values of the other parameters are chosen as $\eta(0) = 0.00116$, $\mu = 1$, $\theta_0 = [0, 0, 0, 0, 1, 1, 0.001, 0.001, 0]^T$, and $\beta(k_0) = 0$. After 6700 steps, the convergence of the estimation is achieved. Figure 5 illustrates the corresponding procedure of the parameter estimation. It shows that the estimate procedure converges quickly. Figure 6 shows the corresponding mean square error (MSE) of the parameter estimation. We can see that the MSE is decreased sharply in the beginning, at the 180th step, and a local minimum can be found. After that, the algorithm jumps out of the local minimum, and the corresponding MSE gradually converges to a constant of about 0.4.

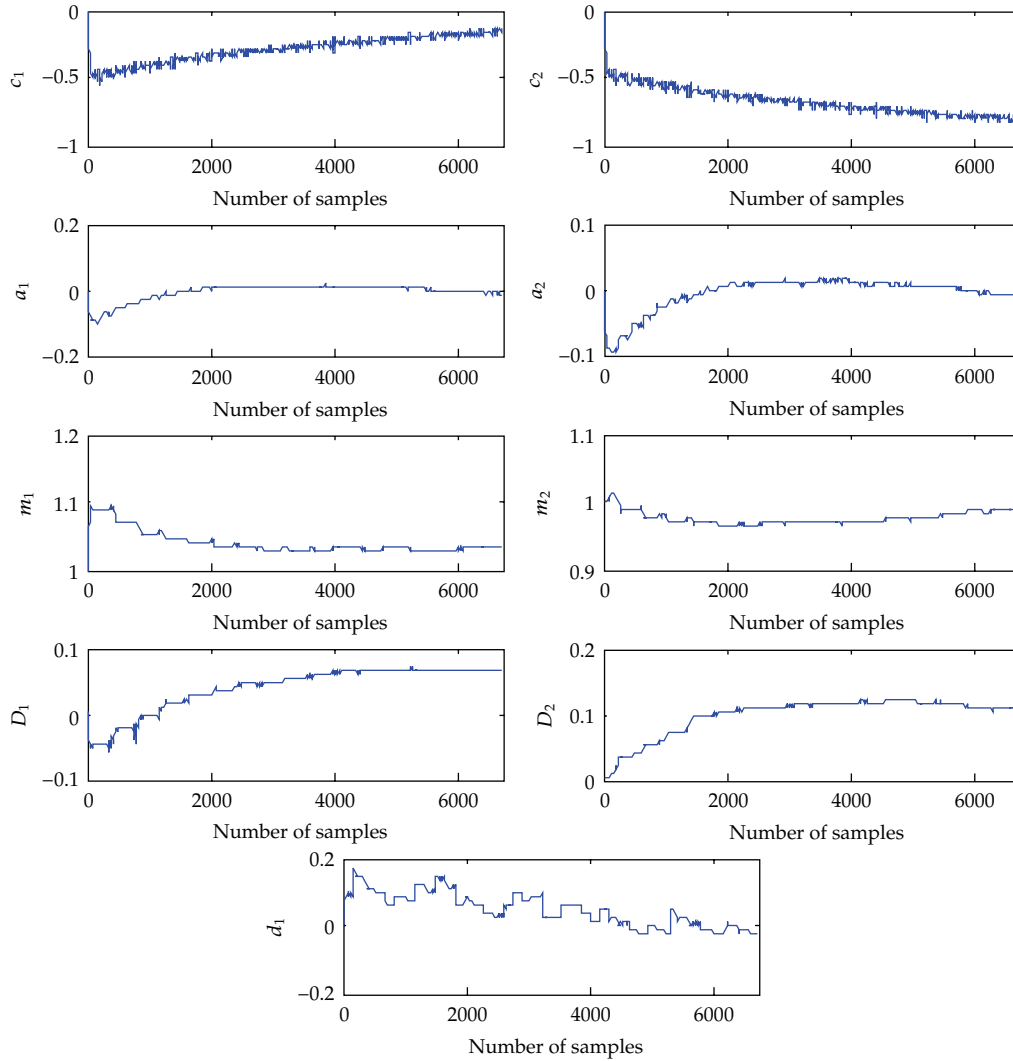


Figure 5: Convergence of parameters of the model of the X-Y moving positioning stage.

Then, the corresponding model validation result is shown in Figure 7(a), while Figure 7(b) shows the comparison of the input-output plots between the proposed model and the real data. The maximum relative modeling error is less than 11%. Moreover, it is obvious that the obtained model can accurately approximate the behavior of the X-Y moving positioning stage. Hence, it can be concluded that the proposed identification method is rather promising in engineering application.

7. Conclusion

In this paper, a recursive gradient-based identification algorithm for the sandwich system with backlash-like hysteresis is proposed. The subgradient is applied to the search of gradient direction at the nonsmooth points of the system. In order to find the proper search

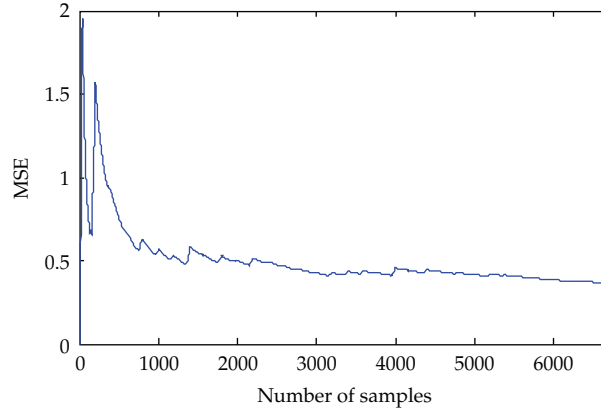


Figure 6: The MSE plot.

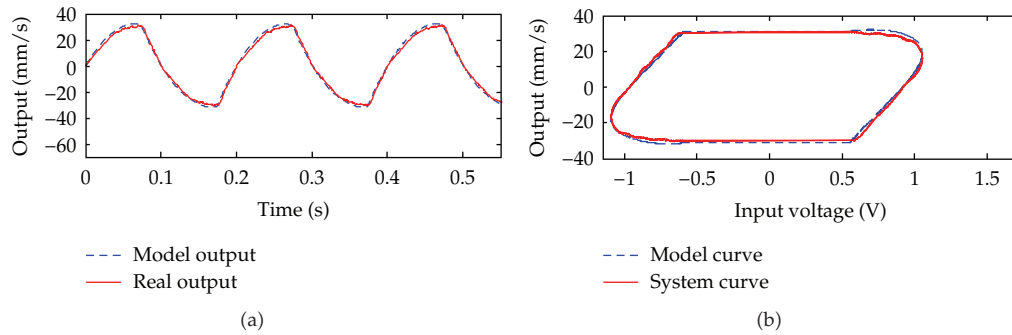


Figure 7: Model validation result.

direction at the nonsmooth points, the technique of so-called bundle method is utilized. Simulation results have shown that the proposed algorithm has provided us with an option for identification of nonsmooth dynamic systems, and it provides a novel method to identify the more complicated nonsmooth systems. The experimental results of X-Y stage also show that the proposed method has potential in engineering applications.

Appendix

Based on (3.5) and Step 3 of the algorithm, it is obtained

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \eta(k)e(k, \hat{\theta}(k))\bar{h}(k). \quad (\text{A.1})$$

Subtracting the local optimal value θ_1 from both sides of (A.1), it leads to

$$\hat{\theta}(k+1) - \theta_1 = \hat{\theta}(k) - \theta_1 + \eta(k)e(k, \hat{\theta}(k))\bar{h}(k). \quad (\text{A.2})$$

Rewrite (A.2) as

$$\tilde{\boldsymbol{\theta}}(k+1) = \tilde{\boldsymbol{\theta}}(k) + \eta(k)e(k, \hat{\boldsymbol{\theta}}(k))\bar{\mathbf{h}}(k), \quad (\text{A.3})$$

where $\tilde{\boldsymbol{\theta}}(k+1) = \hat{\boldsymbol{\theta}}(k+1) - \boldsymbol{\theta}_1$.

Choose the quadratic function as

$$L(k+1) = \tilde{\boldsymbol{\theta}}^T(k+1)\tilde{\boldsymbol{\theta}}(k+1) + e^2(k, \hat{\boldsymbol{\theta}}(k+1)). \quad (\text{A.4})$$

According to (A.3), it leads to

$$\begin{aligned} \tilde{\boldsymbol{\theta}}^T(k+1)\tilde{\boldsymbol{\theta}}(k+1) - \tilde{\boldsymbol{\theta}}^T(k)\tilde{\boldsymbol{\theta}}(k) &= 2\eta(k)e(k, \hat{\boldsymbol{\theta}}(k))\tilde{\boldsymbol{\theta}}^T(k)\bar{\mathbf{h}}(k) \\ &\quad + \eta^2(k)e^2(k, \hat{\boldsymbol{\theta}}(k))\bar{\mathbf{h}}^T(k)\bar{\mathbf{h}}(k). \end{aligned} \quad (\text{A.5})$$

As $f(k, \boldsymbol{\theta}) = [y(k) - \hat{y}(k, \boldsymbol{\theta})]^2$, if (3.5) holds, then the cutting-plane model is

$$\hat{e}_1(k, \boldsymbol{\theta}) = \max \left\{ \left[y(k) - \hat{y}(k, \hat{\boldsymbol{\theta}}(k)) \right]^2 + 2e(k, \hat{\boldsymbol{\theta}}(k)) \left(\hat{\boldsymbol{\theta}}^T(k) - \boldsymbol{\theta}^T \right) \bar{\mathbf{h}}(k) - \beta_j(k) \right\}. \quad (\text{A.6})$$

Based on the definitions of $\beta_j(k)$ and $\alpha_j(k)$, as well as the idea of bundle method, we know that $\hat{e}_1(k, \boldsymbol{\theta}) \leq f(k, \hat{\boldsymbol{\theta}}(k))$. Thus, considering (A.6) yields

$$2e(k, \hat{\boldsymbol{\theta}}(k)) \left(\hat{\boldsymbol{\theta}}^T(k) - \boldsymbol{\theta}^T \right) \bar{\mathbf{h}}(k) \leq \beta_j(k). \quad (\text{A.7})$$

Choosing $\boldsymbol{\theta} = \boldsymbol{\theta}_1$ and holding (A.7) yield

$$2e(k, \hat{\boldsymbol{\theta}}(k)) \left(\hat{\boldsymbol{\theta}}^T(k) - \boldsymbol{\theta}_1^T \right) \bar{\mathbf{h}}(k) \leq \beta_j(k). \quad (\text{A.8})$$

Based on (A.5) and (A.8), we obtain

$$\tilde{\boldsymbol{\theta}}(k+1)\tilde{\boldsymbol{\theta}}^T(k+1) - \tilde{\boldsymbol{\theta}}(k)\tilde{\boldsymbol{\theta}}^T(k) \leq \eta(k)\beta_j(k) + \eta^2(k)e^2(k, \hat{\boldsymbol{\theta}}(k))\bar{\mathbf{h}}^T(k)\bar{\mathbf{h}}(k). \quad (\text{A.9})$$

According to the idea of the gradient algorithm and [16], we know that

$$e(k, \hat{\boldsymbol{\theta}}(k+1)) = e(k, \hat{\boldsymbol{\theta}}(k)) + \Delta e(k, \hat{\boldsymbol{\theta}}(k)). \quad (\text{A.10})$$

Hence, the change of $e(k)$ is written as

$$\Delta e(k, \hat{\theta}(k)) = \left[\frac{\partial e(k, \hat{\theta}(k))}{\partial \theta(k)} \right]^T \Delta \theta(k) = -\bar{\mathbf{h}}(k) \Delta \theta(k). \quad (\text{A.11})$$

According to (3.5) and Step 3 of the algorithm, it yields:

$$\Delta \theta(k) = \eta(k) \mathbf{t}(k) = \eta(k) e(k, \hat{\theta}(k)) \bar{\mathbf{h}}^T(k). \quad (\text{A.12})$$

From (A.11) and (A.12), we get

$$\Delta e(k, \hat{\theta}(k)) = -\eta(k) e(k, \hat{\theta}(k)) \bar{\mathbf{h}}(k) \bar{\mathbf{h}}^T(k). \quad (\text{A.13})$$

According to (A.10) and (A.13), we obtain

$$e^2(k, \hat{\theta}(k+1)) - e^2(k, \hat{\theta}(k)) = -2\eta(k) e^2(k, \hat{\theta}(k)) \bar{\mathbf{h}}(k) \bar{\mathbf{h}}^T(k) + \eta^2(k) e^2(k, \hat{\theta}(k)) \left[\bar{\mathbf{h}}(k) \bar{\mathbf{h}}^T(k) \right]^2. \quad (\text{A.14})$$

Based on (A.4), (A.9), and (A.14), it leads to the following:

$$\begin{aligned} L(k+1) - L(k) &\leq \eta(k) \beta_j(k) \mid \eta^2(k) e^2(k, \hat{\theta}(k)) \bar{\mathbf{h}}^T(k) \bar{\mathbf{h}}(k) \left(1 + \left[\bar{\mathbf{h}}(k) \bar{\mathbf{h}}^T(k) \right] \right) \\ &\quad - 2\eta(k) e^2(k, \hat{\theta}(k)) \bar{\mathbf{h}}(k) \bar{\mathbf{h}}^T(k) \end{aligned} \quad (\text{A.15})$$

if $0 < \eta(k) \leq (2e^2(k, \hat{\theta}(k)) \bar{\mathbf{h}}(k) \bar{\mathbf{h}}^T(k) - \beta_j(k)) / e^2(k, \hat{\theta}(k)) \bar{\mathbf{h}}^T(k) \bar{\mathbf{h}}(k) (1 + [\bar{\mathbf{h}}(k) \bar{\mathbf{h}}^T(k)])$ and $\beta_j(k) < 2e^2(k, \hat{\theta}(k)) [\bar{\mathbf{h}}(k) \bar{\mathbf{h}}^T(k)]$, we have

$$L(k+1) - L(k) \leq 0. \quad (\text{A.16})$$

Hence, the parameters θ can be convergent to a local optimal value.

Acknowledgments

This work was supported by the projects of Shanghai Normal University (DZL811, DRL904, and DYL201005); the projects of Shanghai Education Commission (11YZ92); the project of NSFC (Grant nos. 60971004 and 61171088); the projects of Science and Technology Commission of Shanghai (09220503000, 10JC1412200, and 09ZR1423400).

References

- [1] N. J. Bershad, P. Celka, and S. McLaughlin, "Analysis of stochastic gradient identification of Wiener-Hammerstein systems for nonlinearities with Hermite polynomial expansions," *IEEE Transactions on Signal Processing*, vol. 49, no. 5, pp. 1060–1071, 2001.
- [2] A. H. Tan and K. Godfrey, "Identification of Wiener-Hammerstein models using linear interpolation in the frequency domain (LIFRED)," *IEEE Transactions on Instrumentation and Measurement*, vol. 51, no. 3, pp. 509–521, 2002.
- [3] M. Boutayeb and M. Darouach, "Recursive identification method for MISO Wiener-Hammerstein model," *IEEE Transactions on Automatic Control*, vol. 40, no. 2, pp. 287–291, 1995.
- [4] V. Cerone and D. Regruto, "Bounding the parameters of linear systems with input backlash-like hysteresis," in *Proceedings of the American Control Conference*, pp. 4476–4481, Portland, Ore, USA, June 2005.
- [5] E.-W. Bai, "Identification of linear systems with hard input nonlinearities of known structure," *Automatica*, vol. 38, no. 5, pp. 853–860, 2002.
- [6] F. Giri, Y. Rochdi, F. Z. Chaoui, and A. Brouri, "Identification of Hammerstein systems in presence of hysteresis-backlash and hysteresis-relay nonlinearities," *Automatica*, vol. 44, no. 3, pp. 767–775, 2008.
- [7] R. Dong, Y. Tan, and H. Chen, "Recursive identification for dynamic systems with backlash," *Asian Journal of Control*, vol. 12, no. 1, pp. 26–38, 2010.
- [8] R. Dong, Q. Tan, and Y. Tan, "Recursive identification for dynamic systems with output backlash-like hysteresis and its convergence," *International Journal of Applied Mathematics and Computer Science*, vol. 19, no. 4, pp. 631–638, 2009.
- [9] R. Dong and R. Tan, "Online identification algorithm and convergence analysis for sandwich systems with backlash," *International Journal of Control, Automation and Systems*, vol. 9, no. 3, pp. 1–7, 2011.
- [10] C. Eitzinger, "Nonsmooth training of fuzzy neural networks," *Soft Computing*, vol. 8, pp. 443–448.
- [11] M. M. Mäkelä, M. Miettinen, L. Lukšan, and J. Vlček, "Comparing nonsmooth nonconvex bundle methods in solving hemivariational inequalities," *Journal of Global Optimization*, vol. 14, no. 2, pp. 117–135, 1999.
- [12] S. A. Miller, *An inexact bundle method for solving large structured linear matrix inequality [Doctoral dissertation]*, University of California, Santa Barbara, Calif, USA, 2001.
- [13] C. Lemaréchal, "Nondifferentiable optimization," in *Optimization*, G. L. Nemhauser, A. H. G. Rinnooy Kan, and M. J. Todd, Eds., vol. 1, pp. 529–572, North-Holland, Amsterdam, The Netherlands, 1989.
- [14] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, Mass, USA, 1998.
- [15] R. Fletcher, *Practical Methods of Optimization*, John Wiley & Sons, Chichester, 2nd edition, 1987.
- [16] C.-C. Ku and K. Y. Lee, "Diagonal recurrent neural networks for dynamic systems control," *IEEE Transactions on Neural Networks*, vol. 6, no. 1, pp. 144–156, 1995.

Research Article

The Hybrid Adaptive Control of T-S Fuzzy System Based on Niche

Tong Zhao,^{1,2} Guo-ping Lu,³ Yun-li Hao,⁴ and Yi-min Li⁴

¹ Department of Mathematics, East China Normal University, Shanghai 200241, China

² Department of Mathematics, Nantong Shipping College, Jiangsu, Nantong 226010, China

³ Department of Mathematics, Jiangsu College of Information Technology, Jiangsu, Wuxi 214400, China

⁴ Faculty of Science, Jiangsu University, Jiangsu, Zhenjiang 212013, China

Correspondence should be addressed to Tong Zhao, zhaotong@ntsc.edu.cn

Received 31 March 2012; Revised 14 May 2012; Accepted 14 May 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 Tong Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Based on the niche characteristics, a hybrid adaptive fuzzy control method with the function of continuous supervisory control is proposed in this paper. Considering the close degree of Niche as the consequent of adaptive T-S fuzzy control system, the hybrid control law is designed by tracking, continuous supervisory, and adaptive compensation. Adaptive compensator is used in the controller to compensate the approximation error of fuzzy logic system and the effect of the external disturbance. The adaptive law of consequent parameters, which is achieved in this paper, embodies system adaptability as biological individual. It is proved that all signals in the closed-loop system are bounded and tracking error converges to zero by Lyapunov stability theory. The effectiveness of the approach is demonstrated by the simulation results.

1. Introduction

In recent years, fuzzy technique has gained rapid development in complex nonlinear plants [1]. Fuzzy logic offers human reasoning capabilities to capture nonlinearities and uncertainties, which cannot be described by precise mathematical models. Theoretical proofs of fuzzy models as universal approximators have been presented in the last decade [2, 3]. Fuzzy adaptive control methodologies have emerged in recent years as promising ways to approach nonlinear control problems. Fuzzy control, in particular, has had an impact in the control community because of the simple approach it provides to use heuristic control knowledge for nonlinear control problems. Recently, fuzzy systems have aroused a great resurgence of interest from part of the control community on the ground that they may be applied to model ill-defined complex systems. Direct and indirect schemes are two staple configurations for adaptive fuzzy controls. It has been established that indirect and direct

controls (DAFC and IAFC) are able to incorporate plant knowledge and control knowledge, respectively, to yield stable and robust control systems [4]. The last two decades or so have witnessed a large quantity of results on indirect [5, 6] or direct [7–11] adaptive control (IAC/DAC) using fuzzy systems. However, the fuzzy adaptive controller proposed in [6, 7] ensure that the tracking error asymptotically convergence to zero (or a neighborhood of the origin) if the minimum approximation error is squared integrable along the state trajectory. And those algorithms are just confined to the linear and nonlinear systems whose state variables are assumed to be available for measurement. In many complicated cases, all state variables are not measurable such that output feedback or observer-based adaptive control techniques have to be applied. In [12–15], observer-based IAC and DAC algorithms are proposed for nonlinear systems, respectively. But [6, 12, 13] proposed adaptive control gain is only applicable to nonlinear systems with unknown constant gain. So a hybrid adaptive fuzzy control is needed.

In order to exploit the relative advantages of indirect and direct adaptive configurations at the same time, some researchers have developed several hybrid IAC and DAC algorithms [16–18] where a weighting factor, which can be adjusted by the tradeoff between knowledge of the plant and knowledge of the control, is adopted to sum together the control efforts from both the IAC and DAC. However, those schemes have their limitations. Above all, they take full-state feedback, which can be unsuitable for nonlinear systems without state variables available. Moreover, the conventional adaptive controller proposed in [19] combines indirect, direct, and variable structure methods; nonetheless its plants are assumed to be linear systems only. The author of [6] has developed a hybrid adaptive fuzzy control (HAFC) for nonlinear systems. However, input gains are required to be a constant 1 and control gain is an unknown constant. The HAFC algorithm of [20] can just be applied to robot manipulators with bounds estimation whereas that of [21] derives an unsupported HAFC scheme from a faulty Lyapunov derivative. On the other hand, a certain observer-based combined direct and indirect adaptive fuzzy neural controller is developed in [22]. Sensor fault may be in any form, even unbounded, which will make the system failure unavoidably. Paper [23] proposed a reliable observer-based controller, which makes the system work well no matter whether sensor faults occur or not.

Among various kinds of fuzzy models, there is a very important class of Takagi-Sugeno (T-S) fuzzy models [24] which have recently gained much popularity because of their special rule consequent structure and success in a functional approximation [6, 25]. In the recent two years, with the stability theory of T-S fuzzy system drawing lots of researchers attention, Yeh et al. [25] proposed stability analysis of interconnected and robustness design of time-delay fuzzy systems using fuzzy Lyapunov method. Moreover, T-S fuzzy model plays an important role in dealing with practical problems, such as oceanic structure [26], Vehicle occupant classification [27], and engineering systems [28], in which T-S fuzzy systems are applied to sensor fault estimation.

Upon all the fuzzy control method mentioned above, the consequent rules only contain the mathematical expressions without the practical significances. In this paper, the self-adaptability of the ecosystem is introduced into the fuzzy control method. Ecosystem is a complex large system, the interaction between individuals within the system, "the interaction" between the system and the outside world, made the internal dynamics of ecological system extremely complex. Biological evolution in complex ecosystems displays a strong nonlinearity [29], and increasing complexity is an indication of biological evolution, which makes it easier for organisms to adapt to environmental changes, and higher adaptability is the performance of a higher degree of evolution, that is, the so-called "survival

of the fittest" [30]. Because of the redundancy and stability of ecological system, individuals acquire their own living ability and evolve to the advantage direction. This is why the ecological system could keep balance. The stability of the system is closely related with individuals' ecological niche. Through such feature, Wang and Yi-min [31] discussed a method of fuzzy control based on niche model. The basic concept and methods of niche are introduced. Using the similarity measures, the capabilities of the living's self-learning and self-organization are joined in fuzzy system, gaining the fuzzy system which has the meaning of niche. Considering the high-order nonlinearity, complexity, and biological individuals' adaptability of ecological systems, we combine biological characteristics with adaptive fuzzy control method and raise a new hybrid T-S fuzzy adaptive control way, which contains great biological significance itself. The consequent of fuzzy system is parameter's ecological niche close degree functions for T-S model. In condition of immeasurable biological individuals' ecological factors, in [17, 29], the author designs controller and obtains adaptive law with consequent parameter which takes good tracking effect to changing error target function. Fuzzy T-S system of ecological niche reflects the degree of biological exploiting and usage and develops to its favorable way. During this course, the system moves forward by mutual changes, coordination with environment and themselves. This fuzzy adaptive control of niche reflects that the biology of individual organisms in the new environment has a strong adaptability, strong in the biological and practical significance, but also provided a reference value for the artificial cultivation of rare species. Therefore, fuzzy control background gets its biological meaning in this way.

2. Problem Formulation

Consider the n th-order nonlinear system of the following form:

$$\begin{aligned} x^{(n)} &= f(x, \dot{x}, \dots, x^{(n-1)}) + g(x, \dot{x}, \dots, x^{(n-1)})u + d(t) \\ y &= x, \end{aligned} \quad (2.1)$$

where f and g are unknown but bounded functions, $u, y \in R$ are the input, output of the system, respectively, and y and $d(t)$ is the external bounded disturbance. $X = (x_1, \dots, x_n)^T = (x, \dot{x}, \dots, x^{(n-1)})^T \in R^n$ is the state vector where not all x_i are assumed to be available for measurement. In order for the system (2.1) to be controllable, it is required that $g(X) \neq 0$ for X in a certain controllability region $U_c \subset R^n$. Since $g(X)$ is continuous, we assume $0 < g(X) \leq \infty$ for $X \in U_c$. The control objective is to design a combined controller, tune the correlation parameters of adaptive law, and force the system output y to follow a bounded reference signal $y_m(t)$ under the constraint that all signals involved must be bounded. To begin with, the signal vector $\bar{y}_m(t)$, the tracking error vector e , and estimation error vector \bar{e} are defined as $\bar{y}_m = (y_m, \dot{y}_m, \dots, \dot{y}_m^{(n-1)})^T \in R^n$, $e = y_m - X = (e, \dot{e}, \dots, e^{(n-1)})^T \in R^n$, $\bar{e} = \bar{y}_m - \bar{X} = (\bar{e}, \dot{\bar{e}}, \dots, \dot{\bar{e}}^{(n-1)})^T \in R^n$, where \bar{X} and \bar{e} denote the estimates of X and e , respectively. If the functions $f(X)$ and $g(X)$ are known and the system is free of external disturbance $d(t)$, then we can choose the controller u^* to cancel the nonlinearity and design controller. Let $k_0 = (k_{01}, \dots, k_{0n})^T \in R^n$ be chosen such that all roots of the polynomial $s^n + k_{0n}s^{n-1} + \dots + k_{01}$

are in the open left half-plane and control law of the certainty equivalent controller is obtained as

$$u^* = \frac{1}{g(X)} \left[-f(X) + y_m^{(n)} + k_0^T e \right]. \quad (2.2)$$

Substituting (2.2) to (2.1), we get the main objective of the control is $\lim_{t \rightarrow \infty} e(t) = 0$. However $f(X)$ and $g(X)$ are unknown, the ideal controller (2.2) cannot be implemented, and not all system states X can be measured. So we have to design an observer to estimate the state vector X in the following.

3. T-S Fuzzy System Based on Niche

3.1. Description of T-S Fuzzy System Based on Niche

Because $f(X)$ and $g(X)$ are unknown, T-S fuzzy systems are used to approximate them. The basic configuration of T-S [24] system is expressed as

$$\begin{aligned} R_f^l : & \text{ if } x_1 \text{ is } A_1^l, \dots, x_n \text{ is } A_n^l, \text{ then } \hat{f}_l(X) = \frac{3}{2} - \phi \left[\frac{\lambda_f^l - \lambda}{\sigma_f^l + \sigma} \right], \\ R_g^l : & \text{ if } x_1 \text{ is } G_1^l, \dots, x_n \text{ is } G_n^l, \text{ then } \hat{g}_l(X) = \frac{3}{2} - \phi \left[\frac{\lambda_g^l - \lambda}{\sigma_g^l + \sigma} \right], \\ R_u^l : & \text{ if } x_1 \text{ is } U_1^l, \dots, x_n \text{ is } U_n^l, \text{ then } \hat{u}_l(X) = \frac{3}{2} - \phi \left[\frac{\lambda_u^l - \lambda}{\sigma_u^l + \sigma} \right] \quad (l = 1, \dots, m). \end{aligned} \quad (3.1)$$

Here, A_i^l, G_i^l, U_i^l are fuzzy sets, $\lambda^l = (x_1, x_2, \dots, x_n)$ represents the real ecological factor of ecologic niche, and $\lambda = (x_1^*, x_2^*, \dots, x_n^*)$ is the ideal ecological factor. Consequent indicates the difference between real state of ecologic niche and ideal one. we choose Gaussian membership function, which satisfies $x_i^l - x_j^l \gg \delta$, where x_i^l, x_j^l are the centers and σ is the variance of the functions. Then difference could be showed by approach functions as $H(A, B) = 3/2 - \phi[(\lambda_f^l - \lambda)/(\sigma_f^l + \sigma)]$, where $\phi(x) = (1/\sqrt{2\pi}) \int_{-\infty}^x e^{-t^2/2} dt$. From approach functions above, we could know consequent is a zero-order T-S mode. Here, we use a_0^l as $3/2 - \phi[(\lambda_f^l - \lambda)/(\sigma_f^l + \sigma)]$ for convenience. Let m be the number of systems with central average defuzzifier, and product inference and singleton fuzzifier can be expressed as

$$\hat{f}(X | \theta_f) = \frac{\sum_{l=1}^m a_0^l \left(\prod_{i=1}^n \exp \left(- \left((x_i - \bar{x}_i^l) / \sigma_i^l \right)^2 \right) \right) \tilde{X}}{\sum_{l=1}^m \prod_{i=1}^n \exp \left(- \left((x_i - \bar{x}_i^l) / \sigma_i^l \right)^2 \right)} = \sum_{l=1}^m \xi^l(X) \theta_f^l = \xi(X)^T \theta_f, \quad (3.2)$$

where $\xi^l(X) = \prod_{i=1}^n \exp \left(- \left((x_i - \bar{x}_i^l) / \sigma_i^l \right)^2 \right) \tilde{X} / \sum_{l=1}^m \prod_{i=1}^n \exp \left(- \left((x_i - \bar{x}_i^l) / \sigma_i^l \right)^2 \right)$, $\theta_f^l = (a_0^l, 0, \dots, 0) \in R^{1 \times (n+1)}$, $\tilde{X} = [1, x_1, \dots, x_n]^T$; $\xi(X) = (\xi^1(X), \dots, \xi^m(X)) \in R^{1 \times m}$,

$\theta_f = \begin{pmatrix} \theta_f^1 \\ \vdots \\ \theta_f^m \end{pmatrix} \in R^{m \times (n+1)}$; $\widehat{g}(X \mid \theta_g) = \xi(X)^T \theta_g$ and $u_D(X \mid \theta_D) = \eta(X)^T \theta_D$ can be expressed similarly.

3.2. Niche T-S Model Parameter Optimization Method

Niche fuzzy T-S model based on parameter optimization of backpropagation algorithm known input and output data (x_0^p, y_0^p) , $x_0^p \in U \in R^n$, $y_0^p \in V \in R$, the task is to determine the form (3.2) Niche T-S model parameters, To fitting error $e^p = (1/2)[\widehat{f}(x_0^p) - y_0^p]^2$ the minimum. Assume M is known, by adjusting $\lambda_f^l, \sigma_f^l, \bar{x}_i^l, \sigma_i^l$, and let e^p be the minimum. To facilitate the discussion, use $\widehat{e}, \widehat{f}, \widehat{y}$ to represent $e^p, \widehat{f}(x_0^p)$, and y_0^p . $b = \sum_{l=1}^m Z^l$, $Z^l = \prod_{i=1}^n \exp(-((x_i - \bar{x}_i^l)/\sigma_i^l)^2)$, $a = \sum_{l=1}^m (a_0^l Z^l)$.

Using gradient descent to determines the parameters [6]

$$\lambda_f^l(q+1) = \lambda_f^l(q) - \gamma \left. \frac{\partial \widehat{e}}{\partial \lambda_f^l} \right|_q, \quad l = 1, 2, \dots, m, \quad q = 0, 1, 2, \dots, \gamma \quad (3.3)$$

determines the steps.

Then

$$\frac{\partial \widehat{e}}{\partial \lambda_f^l} = \frac{\partial \widehat{e}}{\partial \widehat{f}} \frac{\partial \widehat{f}}{\partial a} \frac{\partial a}{\partial a_0^l} \frac{\partial a_0^l}{\partial \lambda_f^l} = (\widehat{f} - \widehat{y}) \cdot \frac{1}{\sqrt{2\pi}} e^{-((\lambda_f^l(q) - \lambda)/(\sigma_f^l(q) + \sigma))^2/2} \cdot \frac{1}{\sigma_f^l + \sigma} \cdot \frac{1}{b} \cdot Z^l X. \quad (3.4)$$

Then,

$$\lambda_f^l(q+1) = \lambda_f^l(q) - \gamma (\widehat{f} - \widehat{y}) \cdot \frac{1}{\sqrt{2\pi}} e^{-((\lambda_f^l(q) - \lambda)/(\sigma_f^l(q) + \sigma))^2/2} \cdot \frac{1}{\sigma_f^l(q) + \sigma} \cdot \frac{1}{b} \cdot Z^l X. \quad (3.5)$$

Equally,

$$\sigma_f^l(q+1) = \sigma_f^l(q) - \gamma (\widehat{y} - \widehat{f}) \cdot \frac{1}{\sqrt{2\pi}} e^{-((\lambda_f^l(q) - \lambda)/(\sigma_f^l(q) + \sigma))^2/2} \cdot \frac{\lambda_f^l(q) - \lambda}{(\sigma_f^l(q) + \sigma)^2} \cdot \frac{1}{b} \cdot Z^l X \quad (3.6)$$

adjustable parameters \bar{x}_i^l, σ_i^l with the same way:

$$\begin{aligned}\bar{x}_i^l(q+1) &= \bar{x}_i^l(q) - \gamma \left. \frac{\partial \hat{e}}{\partial \bar{x}_i^l} \right|_q = \bar{x}_i^l(q) - \gamma (\hat{f} - \hat{y}) \left(\hat{f} - \frac{N}{b} \right) \cdot \frac{2(x_i - \bar{x}_i^l(q))}{(\sigma_i^l(q))^2} X \\ \sigma_i^l(q+1) &= \sigma_i^l(q) - \gamma \left. \frac{\partial \hat{e}}{\partial \sigma_i^l} \right|_q = \sigma_i^l(q) - \gamma (\hat{f} - \hat{y}) \left(\hat{f} - \frac{N}{b} \right) \cdot \frac{2(x_i - \bar{x}_i^l(q))^2}{(\sigma_i^l(q))^3} X,\end{aligned}\quad (3.7)$$

where $N = \sum_{l=1}^m (3/2 - \phi((\lambda_f^l - \lambda)/(\sigma_f^l + \sigma)))$, $\lambda_f^l, \sigma_f^l, \lambda_u^l$ and σ_u^l can be optimized directly.

4. Hybrid Controller with Supervisory and Compensation Control Scheme

The overall control law is constructed as

$$u = \alpha u_I(\bar{X}) + (1 - \alpha) u_D(\bar{X} | \theta) + u_s + u_c, \quad (4.1)$$

where u_I is an indirect controller, u_D is the output of the T-S controller, u_s is the supervisory control to force the state within the constraint set, $\alpha \in [0, 1]$ is a weighting factor, and u_c is the compensate controller of adaptive control. Since X cannot be available and $f(X)$ and $g(X)$ are unknown, we replace the functions $f(X)$, $g(X)$, and error vector e by estimation functions $\hat{f}(\bar{X})$, $\hat{g}(\bar{X})$, and \bar{e} . The certainty equivalent controller can be rewritten as

$$u^* = \frac{1}{\hat{g}(\bar{X})} \left[-\hat{f}(\bar{X}) + y_m^{(n)} + k_0^T \bar{e} \right]. \quad (4.2)$$

The indirect control law is written as $u_I = (1/\hat{g}(\bar{X}))[-\hat{f}(\bar{X}) + y_m^{(n)} + k_0^T \bar{e}]$. Applying (4.2), (4.1) to (2.1), the error dynamic equation is

$$\begin{aligned}\dot{e} &= Ae - Bk_0^T \bar{e} + B \left\{ \alpha \left[\hat{f}(\bar{X}) - f(X) + (\hat{g}(\bar{X}) - g(X)) u_I \right] \right\} \\ &\quad + (1 - \alpha) g(X)(u^* - u_D) + Bg(X)(u_s + u_c) - Bd,\end{aligned}\quad (4.3)$$

defining $k_c = (k_{cn}, \dots, k_{c1})$ as the observer vector, the observation errors are defined as $\tilde{e} = e - \bar{e}$ from (4.2) and [24] and we get

$$\begin{aligned}\dot{\tilde{e}} &= \Lambda \tilde{e} + B \left\{ \alpha \left[\hat{f}(\bar{X}) - f(X) + \hat{g}(\bar{X}) - g(X) u_I \right] + (1 - \alpha) g(X)(u^* - u_D) \right\} \\ &\quad - Bg(X)(u_s + u_c) - Bd,\end{aligned}\quad (4.4)$$

where

$$\Lambda = A - k_c C^T = \begin{bmatrix} -k_{cn} & 1 & 0 & 0 & \cdots & 0 & 0 \\ -k_{c(n-1)} & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -k_{c2} & 0 & 0 & 0 & \cdots & 0 & 1 \\ -k_{c1} & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad (4.5)$$

$$B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}.$$

Let $\bar{A} = A - Bk_0^T$ be strictly Hurwitz matrix; so there exists a positive definite symmetric $n \times n$ matrix P, \bar{P} which satisfies the Lyapunov equation $\Lambda^T P + P\Lambda = -Q$, $\bar{A}^T \bar{P} + \bar{P}\bar{A} = -\bar{Q}$, where \bar{Q} and Q are arbitrary $n \times n$ definite symmetric matrix. Let $V_{\bar{e}} = (1/2)\bar{e}^T P \bar{e}$. Since \bar{Q} and k_c are designer by the designer, we can choose \bar{Q} and k_c , such that $\dot{V}_{\bar{e}} \leq 0$. Hence, $V_{\bar{e}}$ is a bounded function and there exists a constant value $\bar{V}_{\bar{e}}$, such that $V_{\bar{e}} \leq \bar{V}_{\bar{e}}$.

5. Hybrid Adaptive Control of Niche

We will develop the hybrid adaptive control such that the closed-loop output $y(t)$ follows $y_m(t)$. Let us replace $\hat{f}(\bar{X}), \hat{g}(\bar{X})$, and $u_D(\bar{X})$ by $\hat{f}(\bar{X} \mid \theta_f), \hat{g}(\bar{X} \mid \theta_g)$, and $u_D(\bar{X} \mid \theta_D)$, respectively. Therefore, the error dynamics (4.3) can be rewritten as

$$\begin{aligned} \tilde{e} = \Lambda \tilde{e} + B \Big\{ \alpha \Big[\hat{f}(\bar{X} \mid \theta_f) - f(X) + \hat{g}(\bar{X} \mid \theta_g) - g(X)u_I \Big] \\ + (1 - \alpha)g(X) \Big(u^* - u_D(\bar{X} \mid \theta_D) \Big) \Big\} + Bg(X)(u_s + u_c) - Bd. \end{aligned} \quad (5.1)$$

Let $V_{\tilde{e}} = (1/2)\tilde{e}^T P \tilde{e}$ then using (5.1), we have

$$\begin{aligned} \dot{V}_{\tilde{e}} \leq -\frac{1}{2}\tilde{e}^T Q \tilde{e} + \left| \tilde{e}^T P B \right| \Big\{ \alpha \Big[\left| \hat{f}(\bar{X} \mid \theta_f) \right| + |f(X)| + \left| \hat{g}(\bar{X} \mid \theta_g)u_I \right| + |g(X)u_I| \Big] \\ + \left| (1 - \alpha)g(X)u^* \right| + \left| \left(g(X)u_D(\bar{X} \mid \theta_D) \right) \right| \Big\} \\ + \left| \tilde{e}^T P B \right| \left[(1 - \alpha)|g(X)\alpha + u_c| \right] - \tilde{e}^T P B g(X)u_s. \end{aligned} \quad (5.2)$$

In order to design u_s so that $\dot{V}_{\tilde{e}} \leq 0$, we need the following assumption.

Assumption 5.1. We could find three functions as $f^u(X), g^u(X), g_l(X)$ and get $|f(X)| \leq f^u(X) \approx f^u(\bar{X})$ and $0 < g_l(\bar{X}) \approx g_l(X) \leq g(X) \leq g^u(X) \approx g^u(\bar{X})$, in which $X \in U_c$ and

$f^u(X) \approx f^u(\bar{X}) < \infty$, $g^u(X) \approx g^u(\bar{X}) < \infty$. This is due to the fact that we can choose k_c to let $X \approx \bar{X}$. Also external disturbance is bounded. We design $|d(t)| \leq d_m$.

From Assumption 5.1, we choose the supervisory control u_s as

$$u_s = \left(\frac{V_{\tilde{e}}}{\bar{V}}\right)^p \text{sgn}(\tilde{e}^T PB) \left[\frac{\alpha}{g_l(X)} |\hat{f}(\bar{X} | \theta_f)| + f^u(|\hat{g}(\bar{X} | \theta_g)| + g^u|u_l|) + (f^u(X) + |y_m^{(n)}| + |k^T e|) \right] \\ + \left(\frac{V_{\tilde{e}}}{\bar{V}}\right)^p \text{sgn}(\tilde{e}^T PB) (1 - \alpha) \left([|u_d|] + \frac{1}{g_l(X)} (f^u(X) + |y_m^{(n)}| + |k_0^T \tilde{e}|) + g^u|u_c| + d_m \right). \quad (5.3)$$

And we choose $u_c = c_o \text{sgn}(\tilde{e}^T PB)$, and c_o is a nonnegative constant. Considering the case $V_{\tilde{e}} > \bar{V}$ and substitute (5.3) into (5.2), we obtain $\dot{V}_{\tilde{e}} \leq -(1/2)\tilde{e}^T Q \tilde{e} \leq 0$. Therefore, if the closed-loop system with the fuzzy controller u as

$$u = \alpha \hat{g}^{-1}(\bar{X} | \theta_g) \left[-\hat{f}(\bar{X} | \theta_f) + |y_m^{(n)}| + |k_0^T \tilde{e}| \right] + (1 - \alpha) u_D(\bar{X} | \theta_D) + u_s + u_c \quad (5.4)$$

works well in the sense that the error is not too large, if $V_{\tilde{e}} \leq \bar{V}$, then the supervisory control u_s is zero. If the system tends to diverge, that is, $V_{\tilde{e}} > \bar{V}$, then be u_s gins to force $V_{\tilde{e}} \leq \bar{V}$.

6. Design of Adaptive Law

In order to adjust the parameters in the fuzzy logic system, we have to derive adaptive laws. Hence, the optimal parameters estimates θ_f^* , θ_g^* , and θ_D^* are defined as

$$\theta_f^* = \arg \min_{\theta_f \in \Omega_f} \left[\sup_{X \in \Omega_X, \bar{X} \in \Omega_{\bar{X}}} |\hat{f}(\bar{X} | \theta_f) - f(X)| \right], \\ \theta_g^* = \arg \min_{\theta_g \in \Omega_g} \left[\sup_{X \in \Omega_X, \bar{X} \in \Omega_{\bar{X}}} |\hat{g}(\bar{X} | \theta_g) - g(X)| \right], \quad (6.1) \\ \theta_D^* = \arg \min_{\theta_D \in \Omega_D} \left[\sup_{X \in \Omega_X, \bar{X} \in \Omega_{\bar{X}}} |u^*(X) - u(\bar{X} | \theta_D)| \right],$$

where $\Omega_f, \Omega_g, \Omega_D, \Omega_X$, and $\Omega_{\bar{X}}$, are compact sets of suitable bounds on $\theta_f, \theta_g, \theta_D, \bar{X}$, and X , respectively, and they are defined as

$$\Omega_f = \{\theta_f : \|\theta_f\| \leq M_f\}, \quad \Omega_g = \{\theta_g : \|\theta_g\| \leq M_g\}, \quad \Omega_D = \{\theta_D : \|\theta_D\| \leq M_D\}, \\ \Omega_{\bar{X}} = \{\bar{X} : \|\bar{X}\| \leq M_{\bar{X}}\}, \quad \Omega_X = \{X : \|X\| \leq M_X\}, \quad (6.2)$$

where $M_f, M_g, M_D, M_{\bar{X}}$ and M_X , are positive constants. Define the minimum approximation errors as

$$\omega = \alpha \left[\hat{f}(\bar{X} \mid \theta_f^*) - f(X) + \hat{g}(\bar{X} \mid \theta_g^*) - g(X) u_I \right] + (1 - \alpha) g(X) \left(u^*(X) - u(\bar{X} \mid \theta_D^*) \right) - d(t). \quad (6.3)$$

Now consider the Lyapunov function

$$\begin{aligned} V = & \frac{1}{2} \tilde{e}^T P \tilde{e} + \frac{1}{2\Gamma_f} (\theta_f - \theta_f^*)^T (\theta_f - \theta_f^*) + \frac{\alpha}{2\Gamma_g} (\theta_g - \theta_g^*)^T (\theta_g - \theta_g^*) \\ & + \frac{(1 - \alpha)}{2\Gamma_D} (\theta_D - \theta_D^*)^T (\theta_D - \theta_D^*), \end{aligned} \quad (6.4)$$

where $\Gamma_f = \begin{bmatrix} \tau_{f0} & & \\ & \ddots & \\ & & \tau_{fn} \end{bmatrix}_{(n+1)(n+1)}$ and Γ_g and Γ_D is expressed similarly.

We get the adaptive law as

$$\dot{\theta}_f = -\Gamma_f \xi(\bar{X}) B^T P \tilde{e}, \quad \dot{\theta}_g = -\Gamma_g \xi(\bar{X}) B^T P \tilde{e} u_I, \quad \dot{\theta}_D = -\Gamma_D \eta(\bar{X}) g(X) B^T P \tilde{e}. \quad (6.5)$$

So

$$(\dot{\theta}_f)^T = \begin{bmatrix} \dot{a}_0^l \\ \dot{a}_1^l \\ \vdots \\ \dot{a}_n^l \end{bmatrix} = \xi(\bar{X}) B^T P \tilde{e} \begin{bmatrix} \tau_{f0} \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (6.6)$$

Furthermore, the adaptive law of niche factors is derived as follows. We derived the adaptive law of real ecologic factors, see Formula (6.7), (6.8), which represented the real niche always develop towards the ideal one. It reflected the compensation of the control system to the external disturbance:

$$\begin{aligned} \dot{\lambda}_f(q+1) &= -\frac{1}{\sqrt{2\pi}} e^{-((\lambda_f(q)+\lambda)/(\sigma_f(q)+\sigma))^2/2} \frac{1}{\sigma_f(q)+\sigma} \tau_{f0} \xi(\bar{X}) B^T P \tilde{e}, \\ \dot{\sigma}_f(q+1) &= \frac{1}{\sqrt{2\pi}} e^{-((\lambda_f(q)+\lambda)/(\sigma_f(q)+\sigma))^2/2} \frac{\lambda_f(q)+\lambda}{(\sigma_f(q)+\sigma)^2} \tau_{f0} \xi(\bar{X}) B^T P \tilde{e}. \end{aligned} \quad (6.7)$$

The same way we have

$$\begin{aligned}
\dot{\lambda}_g(q+1) &= -\frac{1}{\sqrt{2\pi}} e^{-((\lambda_g(q)+\lambda)/(\sigma_g(q)+\sigma))^2/2} \frac{1}{\sigma_g(q)+\sigma} \tau_{g0} B^T P \tilde{e} u_I, \\
\dot{\sigma}_g(q+1) &= \frac{1}{\sqrt{2\pi}} e^{-((\lambda_g(q)+\lambda)/(\sigma_g(q)+\sigma))^2/2} \frac{\lambda_g(q)+\lambda}{(\sigma_g(q)+\sigma)^2} \tau_{g0} B^T P \tilde{e} u_I, \\
\dot{\lambda}_D(q+1) &= -\frac{1}{\sqrt{2\pi}} e^{-((\lambda_D(q)+\lambda)/(\sigma_D(q)+\sigma))^2/2} \frac{1}{\sigma_D(q)+\sigma} \tau_{D0} \eta(\bar{X}) g(X) B^T P \tilde{e}, \\
\dot{\sigma}_D(q+1) &= \frac{1}{\sqrt{2\pi}} e^{-((\lambda_D(q)+\lambda)/(\sigma_D(q)+\sigma))^2/2} \frac{\lambda_D(q)+\lambda}{(\sigma_D(q)+\sigma)^2} \tau_{D0} \eta(\bar{X}) g(X) B^T P \tilde{e}.
\end{aligned} \tag{6.8}$$

Applying u_s we have $\dot{V} \leq -(1/2)\tilde{e}^T Q \tilde{e} + \tilde{e}^T P B \omega$.

Design parameters of, vector adaptive law of θ_f , θ_g , θ_D , stability, and performance analysis are similar to those of [18]; here we omit it.

7. Simulation

Form combined fuzzy adaptive control of two-dimensional predation system [32] as

$$\begin{aligned}
\frac{dx_1}{dt} &= x_1 \left(a - cx_1 - bx_1^2 - h(x) \right) - x_2 \frac{\alpha x_1^2}{1 + \beta x_1^2} \\
\frac{dx_2}{dt} &= x_2 \left(-q + \frac{k \alpha x_1^2}{1 + \beta x_1^2} \right),
\end{aligned} \tag{7.1}$$

where $x_1(t)$ means total number of food at the time of t and $x_2(t)$ denotes the total number of predators; $a, b, c, q, k, \alpha, \beta$ are the regular numbering ecology, and k is transforming factor, q represents death ratio of predators. $h(x)$ is the function of Holling's functional responses, and $x_2(\alpha x_1^2/(1 + \beta x_1^2))$ is the third kind of Holling's functional responses. The demonstration of the two-dimensional predator system without the controller is shown in Figure 1.

In order to reach an ideal ecologic balance in this two-dimension predators system, we get a way to control it, where $h(x) = 0.0012x_1^3$, $a = 1.0$, $b = 0.0014$, $c = 0.06$, $q = 0.2$, $k = 0.08$, $\alpha = 5.0$, $\beta = 1.2$.

Then

$$\begin{aligned}
\frac{dx_1}{dt} &= x_1 \left[\left(1 + 1.2x_1^2 \right) \left(1 - 0.06x_1 - 0.0014x_1^2 - 0.0012x_1^3 \right) - x_1x_2 \right] \\
\frac{dx_2}{dt} &= x_2 \left(0.16x_1^2 - 0.2 \right) + u \\
y &= x_1.
\end{aligned} \tag{7.2}$$

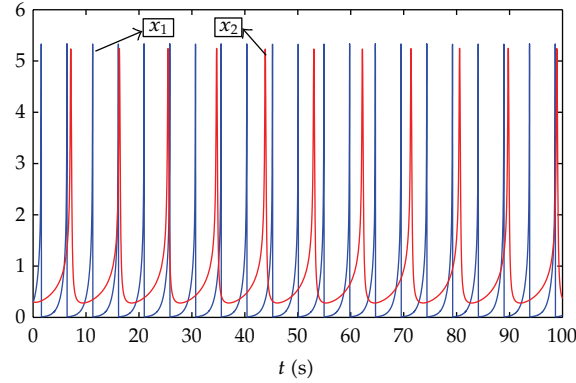


Figure 1: Two-dimension predator system without the controller.

In order to establish the direct relation between output y and controller u , we need the derivation of y . After derivating twice, we get

$$\begin{aligned} \ddot{y} = \ddot{x}_1 = & x_1 \left[(1 + 1.2x_1^2) (0.8 - 0.06x_1 - 0.0014x_1^2 - 0.0012x_1^3) - x_1x_2 \right]^2 \\ & - x_1^2x_2 (0.8 - 0.06x_1 + 1.3586x_1^2 - 0.0732x_1^3 - 0.00168x_1^4 - 0.00144x_1^5 - x_1x_2) \\ & + 2.3972x_1^2 - 0.2196x_1^3 - 0.00672x_1^4 - 0.0072x_1^5 - 0.06x_1 - x_1^2u. \end{aligned} \quad (7.3)$$

Here, we command

$$\begin{aligned} f(X) = & x_1 \left[(1 + 1.2x_1^2) (0.8 - 0.06x_1 - 0.0014x_1^2 - 0.0012x_1^3) - x_1x_2 \right]^2 \\ & - x_1^2x_2 (0.8 - 0.06x_1 + 1.3586x_1^2 - 0.0732x_1^3 - 0.00168x_1^4 - 0.00144x_1^5 - x_1x_2) \\ & + 2.3972x_1^2 - 0.2196x_1^3 - 0.00672x_1^4 - 0.0072x_1^5 - 0.06x_1 \\ g(X) = & -x_1^2. \end{aligned} \quad (7.4)$$

Then (7.2) and (7.3) could be shown as

$$\ddot{y} = f(X) + g(X)u. \quad (7.5)$$

If hoping to apply adaptive fuzzy controlling system here, we have to firstly confirm the boundary of f^u , g^u , and g_l . From [29], we know that $0.2 < |x_1| < 5.5, 1.5 < |x_2| < 5$, then

$$\begin{aligned} |f(x_1, x_2)| & \leq 2.7205e + 0.03 \approx 2.7205e + 0.03 = f^u(\bar{x}_1, \bar{x}_2) \\ 0 < g_l(\bar{x}_1, \bar{x}_2) & \approx g_l(x_1, x_2) = 0.04 \leq |g(x_1, x_2)| \leq 30.25 = g^u(x_1, x_2) \approx g^u(\bar{x}_1, \bar{x}_2). \end{aligned} \quad (7.6)$$

According to the above, we find out that the scope of $f(x_1, x_2)$ is larger than $g(x_1, x_2)$. Therefore, we choose $\tau_{f0} = 50, \tau_{g0} = 1, \tau_{D0} = 15$. Then we select $M_f = 7.7, M_g = 31, M_D =$

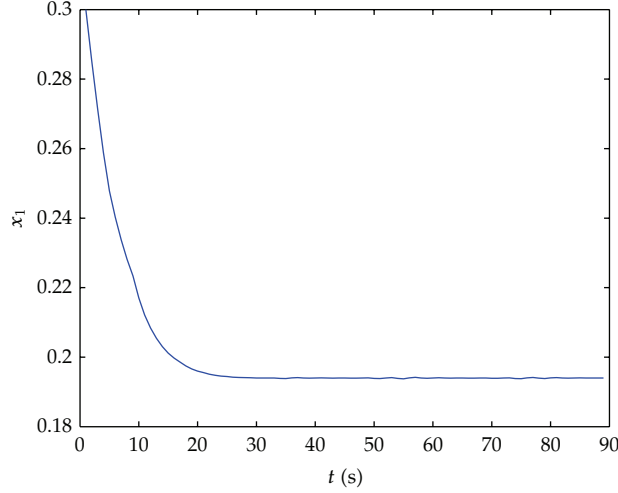


Figure 2: The trend of the prey numbers with T-S controller.

31, $\varepsilon = 0.5$, $|u| \leq 5$. The proposal method of this paper does not need to predefine the reference signal, but to reach the balance of each individual self-adaptively. Conveniently, we suppose $k_c = [89 \ 184]$ (in this way, $s^2 + k_1s + k_2$ could be stable.) and $k_0 = [4 \ 4]$, we select $Q = \begin{bmatrix} 10 & 13 \\ 13 & 28 \end{bmatrix}$, from $\Lambda_c^T P + P \Lambda_c = -Q$, then we solve $P = \begin{bmatrix} 29 & -14 \\ -14 & 7 \end{bmatrix}$. \hat{Q} as $\begin{bmatrix} 40 & 25 \\ 25 & 30 \end{bmatrix}$, and $\hat{A} = \begin{bmatrix} 0 & 1 \\ -4 & -4 \end{bmatrix}$ so we have $\hat{P} = \begin{bmatrix} 15 & 5 \\ 5 & 5 \end{bmatrix}$; when letting $\lambda_{\min} = 2.93$, we could get $\bar{V} = 2\lambda_{\min} = 5.86$. In this way, we get the ideal $= 0$, $\sigma = 1$. The adaptive law comes out as follows:

$$\begin{aligned}
 \dot{\lambda}_f(q+1) &= -\frac{50}{\sqrt{2\pi}} e^{-(\lambda_f(q)/(\sigma_f(q)+1))^2/2} \frac{1}{\sigma_f(q)+1} \xi(\bar{X}) B^T P \tilde{e} \\
 \dot{\sigma}_f(q+1) &= \frac{50}{\sqrt{2\pi}} e^{-(\lambda_f(q)/(\sigma_f(q)+1))^2/2} \frac{\lambda_f(q)}{(\sigma_f(q)+1)^2} \xi(\bar{X}) B^T P \tilde{e} \\
 \dot{\lambda}_g(q+1) &= -\frac{1}{\sqrt{2\pi}} e^{-(\lambda_g(q)/(\sigma_g(q)+1))^2/2} \frac{1}{\sigma_g(q)+1} B^T P \tilde{e} u_I \\
 \dot{\sigma}_g(q+1) &= \frac{1}{\sqrt{2\pi}} e^{-(\lambda_g(q)+0)/(\sigma_g(q)+1))^2/2} \frac{\lambda_g(q)}{(\sigma_g(q)+1)^2} B^T P \tilde{e} u_I \\
 \dot{\lambda}_D(q+1) &= -\frac{15}{\sqrt{2\pi}} e^{-(\lambda_D(q)/(\sigma_D(q)+1))^2/2} \frac{1}{\sigma_D(q)+1} \eta(\bar{X}) g(X) B^T P \tilde{e} \\
 \dot{\sigma}_D(q+1) &= \frac{15}{\sqrt{2\pi}} e^{-(\lambda_D(q)/(\sigma_D(q)+1))^2/2} \frac{\lambda_D(q)}{(\sigma_D(q)+1)^2} \eta(\bar{X}) g(X) B^T P \tilde{e}.
 \end{aligned} \tag{7.7}$$

We choose $x_1(0) = 0.3$, $x_2(0) = 1.0$, then the trend of the prey and predator numbers with T-S controller show as in Figures 2 and 3. The trend of the T-S controller is shown in Figure 4.

From the simulation figures, we can see that the prey and predator numbers reach a stable status in a short period of time under the control of the proposed method in this

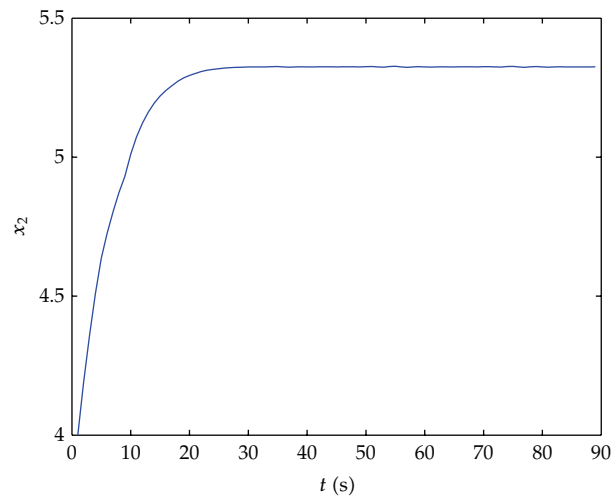


Figure 3: The trend of the predator numbers with T-S controller.

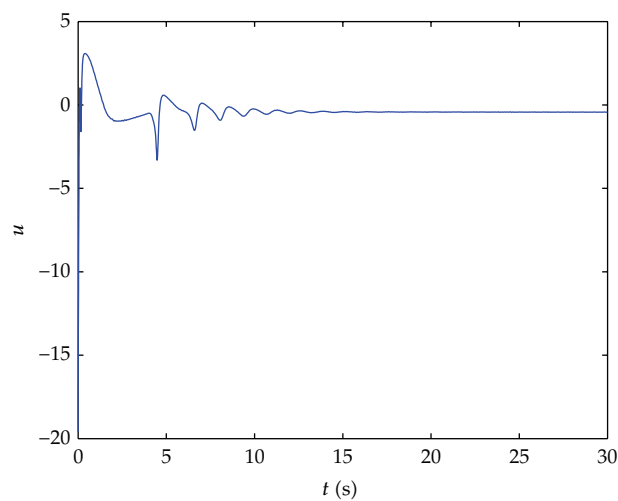


Figure 4: The trend of the T-S controller.

paper. The creature individuals show a characteristic of self-adaptation according to outside changes.

From the figure of the trend of the T-S controller, it also reaches an ideal status to maintain the overall balance of the two-dimension predator system. The ecological system is optimized with the use of this control method.

8. Conclusion

For the ecological niche, a hybrid adaptive fuzzy control method with the function of continuous supervisory control is proposed in this paper. Let the close degree of Niche which contains parameters as the consequent of adaptive T-S fuzzy control system, then designs

the hybrid control law by tracking, continuous supervisory and adaptive compensation. Using gradient descent to optimize the parameters, we get the adaptive law of consequent parameters, embodying biological individual's ability of adaptability. Based on Lyapunov stability theory, it is proved that all signals in the closed-loop system are bounded and tracking error converges to zero. This paper shows that the fuzzy methods provide good results in practical engineering problems. The performance of the developed approach is illustrated by simulation, on two-dimension predation system model.

Acknowledgment

Work supported by national science foundation(11072090).

References

- [1] C. W. Chen, K. Yeh, and K. F. R. Liu, "Adaptive fuzzy sliding mode control for seismically excited bridges with lead rubber bearing isolation," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 17, no. 5, pp. 705–727, 2009.
- [2] B. Kosko, "Fuzzy systems as universal approximators," *IEEE Transactions on Computers*, vol. 43, no. 11, pp. 1329–1333, 1994.
- [3] R. Rovatti, "Fuzzy piecewise multilinear and piecewise linear systems as universal approximators in sobolev norms," *IEEE Transactions on Fuzzy Systems*, vol. 6, no. 2, pp. 235–249, 1998.
- [4] L. X. Wang, *A Course in Fuzzy Systems and Control*, Prentice Hall, Englewood Cliffs, NJ, USA, 1997.
- [5] R. Shahnazi and M. R. Akbarzadeh-T, "PI adaptive fuzzy control with large and fast disturbance rejection for a class of uncertain nonlinear systems," *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 1, pp. 187–197, 2008.
- [6] L. X. Wang, "Stable adaptive fuzzy control of nonlinear systems," *IEEE Transactions on Fuzzy Systems*, vol. 1, no. 2, pp. 146–155, 1993.
- [7] P. A. Phan and T. J. Gale, "Direct adaptive fuzzy control with a self-structuring algorithm," *Fuzzy Sets and Systems*, vol. 159, no. 8, pp. 871–899, 2008.
- [8] S. Labiod and T. M. Guerra, "Adaptive fuzzy control of a class of SISO nonaffine nonlinear systems," *Fuzzy Sets and Systems*, vol. 158, no. 10, pp. 1126–1137, 2007.
- [9] J. H. Park, S. H. Huh, S. H. Kim, S. J. Seo, and G. T. Park, "Direct adaptive controller for nonaffine nonlinear systems using self-structuring neural networks," *IEEE Transactions on Neural Networks*, vol. 16, no. 2, pp. 414–422, 2005.
- [10] M. Wang, B. Chen, and S.-L. Dai, "Direct adaptive fuzzy tracking control for a class of perturbed strict-feedback nonlinear systems," *Fuzzy Sets and Systems*, vol. 158, no. 24, pp. 2655–2670, 2007.
- [11] S. S. Ge and T. T. Han, "Semiglobal ISpS disturbance attenuation with output tracking via direct adaptive design," *IEEE Transactions on Neural Networks*, vol. 18, no. 4, pp. 1129–1148, 2007.
- [12] R. Qi and M. A. Brdys, "Stable indirect adaptive control based on discrete-time T-S fuzzy model," *Fuzzy Sets and Systems*, vol. 159, no. 8, pp. 900–925, 2008.
- [13] C.-H. Hyun, C.-W. Park, and S. Kim, "Takagi-Sugeno fuzzy model based indirect adaptive fuzzy observer and controller design," *Information Sciences*, vol. 180, no. 11, pp. 2314–2327, 2010.
- [14] S. Tong, H.-X. Li, and W. Wang, "Observer-based adaptive fuzzy control for SISO nonlinear systems," *Fuzzy Sets and Systems*, vol. 148, no. 3, pp. 355–376, 2004.
- [15] Y. G. Leu, W. Y. Wang, and T. T. Lee, "Observer-based direct adaptive fuzzy-neural control for nonaffine nonlinear systems," *IEEE Transactions on Neural Networks*, vol. 16, no. 4, pp. 853–861, 2005.
- [16] M. K. Ciliz, "Combined direct and indirect adaptive control for a class of nonlinear systems," *IET Control Theory & Applications*, vol. 3, no. 1, pp. 151–159, 2009.
- [17] Y. Q. Zheng, Y. J. Liu, S. C. Tong, and T. S. Li, "Combined adaptive fuzzy control for uncertain MIMO nonlinear systems," in *Proceedings of the American Control Conference (ACC '09)*, pp. 4266–4271, St. Louis, Mo, USA, June 2009.
- [18] Q. Ding, H. Chen, C. Jiang, and Z. Chen, "Combined indirect and direct method for adaptive fuzzy output feedback control of nonlinear system," *Journal of Systems Engineering and Electronics*, vol. 18, no. 1, pp. 120–124, 2007.

- [19] X. Ye and J. Huang, "Decentralized adaptive output regulation for a class of large-scale nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 276–281, 2003.
- [20] N. Hovakimyan, E. Lavretsky, B. J. Yang, and A. J. Calise, "Coordinated decentralized adaptive output feedback control of interconnected systems," *IEEE Transactions on Neural Networks*, vol. 16, no. 1, pp. 185–194, 2005.
- [21] S. Tong, H. X. Li, and G. Chen, "Adaptive fuzzy decentralized control for a class of large-scale nonlinear systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 34, no. 1, pp. 770–775, 2004.
- [22] S. S. Stanković, D. M. Stipanović, and D. D. Šiljak, "Decentralized dynamic output feedback for robust stabilization of a class of nonlinear interconnected systems," *Automatica*, vol. 43, no. 5, pp. 861–867, 2007.
- [23] Z. Gao, T. Breikin, and H. Wang, "Reliable observer-based control against sensor failures for systems with time delays in both state and input," *IEEE Transactions on Systems, Man, and Cybernetics Part A*, vol. 38, no. 5, pp. 1018–1029, 2008.
- [24] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 15, no. 1, pp. 116–132, 1985.
- [25] K. Yeh, C.-Y. Chen, and C.-W. Chen, "Robustness design of time-delay fuzzy systems using fuzzy Lyapunov method," *Applied Mathematics and Computation*, vol. 205, no. 2, pp. 568–577, 2008.
- [26] C.-W. Chen, "The stability of an oceanic structure with T-S fuzzy models," *Mathematics and Computers in Simulation*, vol. 80, no. 2, pp. 402–426, 2009.
- [27] Z. Gao and Y. Zhao, "Vehicle occupant classification algorithm based on T-S fuzzy model," *Procedia Engineering*, vol. 24, pp. 500–504, 2011.
- [28] Z. Gao, X. Shi, and S. X. Ding, "Fuzzy state/disturbance observer design for T-S fuzzy systems with application to sensor fault estimation," *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 38, no. 3, pp. 875–880, 2008.
- [29] A. L. Goldberger, "Fractal mechanisms in the electrophysiology of the heart," *IEEE Engineering in Medicine and Biology*, vol. 11, no. 2, pp. 47–52, 1992.
- [30] A. Babloyantz and A. Destexhe, "Is the normal heart a periodic oscillator?" *Biological Cybernetics*, vol. 58, no. 3, pp. 203–211, 1988.
- [31] X.-M. Wang and L. I. Yi-min, "Kind of new niche fuzzy control method," *Science Technology and Engineering*, vol. 24, pp. 6318–6321, 2007.
- [32] S. P. Li, "Dynamical behavior analysis and control of two functional ecosystem," *Jiang Su University*, pp. 33–43, 2005.

Research Article

Modeling and Analysis of Epidemic Diffusion within Small-World Network

Ming Liu and Yihong Xiao

Department of Management Science and Engineering, Nanjing University of Science and Technology, Jiangsu 210094, China

Correspondence should be addressed to Ming Liu, liumingsu@126.com

Received 12 March 2012; Revised 29 April 2012; Accepted 29 April 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 M. Liu and Y. Xiao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To depict the rule of epidemic diffusion, two different models, the Susceptible-Exposure-Infected-Recovered-Susceptible (SEIRS) model and the Susceptible-Exposure-Infected-Quarantine-Recovered-Susceptible (SEIQRS) model, are proposed and analyzed within small-world network in this paper. Firstly, the epidemic diffusion models are constructed with mean-field theory, and condition for the occurrence of disease diffusion is explored. Then, the existence and global stability of the disease-free equilibrium and the endemic equilibrium for these two complex epidemic systems are proved by differential equations knowledge and Routh-Hurwitz theory. At last, a numerical example which includes key parameters analysis and critical topic discussion is presented to test how well the proposed two models may be applied in practice. These works may provide some guidelines for decision makers when coping with epidemic diffusion controlling problems.

1. Introduction

Disastrous epidemics such as SARS and H1N1 can significantly impact people's life. The outbreak of infections in Europe last year is another recent example. The infection, from a strain of *Escherichia coli*, can lead to kidney failure and death and is difficult to treat with antibiotics. It is now widely recognized that a large-scale epidemic diffusion can conceivably cause many deaths and more people of permanent sequelae, which presents a severe challenge to a local or regional health-care systems.

After an epidemic outbreak, public officials are faced with many critical and complex issues, the most important of which is to make certain how the epidemic diffuses. Actually, many recent research efforts have been devoted to understanding the prevention and control of epidemics, such as those of Wein et al. [1], Wein et al. [2], Craft et al. [3], Kaplan et al. [4, 5], and Matsuura et al. [6]. A very recent research by Shi and Dong [7] formulates and discusses

models for the spread of infectious diseases with variable population sizes and vaccinations on the susceptible individuals. Various mathematical models, such as susceptible-infected (SI), susceptible-infected-recovered (SIR), susceptible-infected-susceptible (SIS), susceptible-infected-recovered-susceptible (SIRS), susceptible-exposure-infective-recovered (SEIR), and susceptible-infective-quarantine-recovered-susceptible (SIQRS), are proposed to analyze and study the general characteristics of epidemic. It is worth mentioning that the major purpose of these papers is to compare the performance of the following two strategies, the traced vaccination (TV) strategy and the mass vaccination (MV) strategy. Furthermore, epidemic diffusion models they adopted are based on traditional compartment models, while complex topological structure of social contact network has not been considered.

As it is well known, a class of network with a topology interpolating between that of lattices and random graphs is proposed by Watts and Strogatz [8]. In these models, a fraction of the links of the lattice is randomized by connecting nodes, with probability p , with any other node. For a range of p , the network exhibits “small world” behavior, where a local neighborhood (as in lattices) coexists with a short average path length (as in random graphs). Analysis of real networks reveals the existence of small worlds in many interaction networks, including networks of social contacts [9]. Recently, attention has been focused on the impact of network topology on the dynamics of the processes running on it with emphasis on the spreading of infectious diseases. For many infectious diseases, a small-world network on an underlying regular lattice is a suitable simplified model for the contact structure of the host population. It is well known that the contact network plays an important role both in the short-term and in the long-term dynamics of epidemic spread [10]. Thus, one of the major motivations for studying the small world network in this paper is to better understand the structure of social contact network, where there is a natural link between the epidemiological modeling and the science of small world network.

Saramäki and Kaski [11] proposed an SIR model for modeling the spreading process of randomly contagious diseases, such as influenza, based on a dynamic small-world network. Masuda and Konno [12] presented a multistate epidemic process based on a complex network. They analyzed the steady states of various multistate disease propagation models with heterogeneous contact rates. Xu et al. [13] presented a modified SIS model based on complex networks, small world and scale free, to study the spread of an epidemic by considering the effect of time delay. Based on two-dimension small-world networks, a susceptible-infected (SI) model with epidemic alert is proposed by Han [14]. This model indicates that the broadcasting of a timely epidemic alert is helpful and necessary in the control of epidemic spreading, and this is in agreement with the general view of epidemic alert. Stone et al. [15] studied the relative effects of vaccinations and avoidance of infected individuals in a susceptible-infected-recovered (SIR) epidemic model on a dynamic small-world network. They derived the critical mobility required for an outbreak to occur as a function of the disease’s infectivity, recovery rate, avoidance rate, and vaccination rate. Hsu and Shih [16] focused on the human-to-human transmission of influenza- and investigated the effects of air travel activities on an influenza pandemic in a small-world network. This study also investigated how the small-world properties of an air transportation network facilitate the spread of influenza around the globe. The results show that, as soon as the influenza is spread to the top 50 global airports, the transmission is greatly accelerated. It is worth mentioning that majority of the existing studies relies on different kinds of differential equations. For instance, first-order partial differential equations are used to integrate the age structures; second-order partial differential equations are suitable when a diffusion term

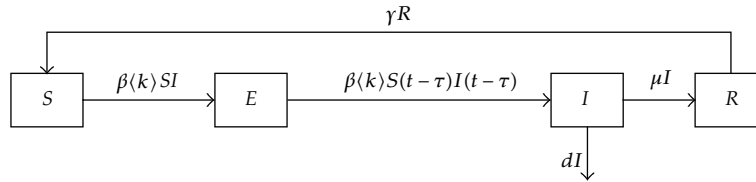


Figure 1: Framework of Model I.

exists; and integral differential equations or differential equations are often used when time delay or delay factors are considered [17].

In this paper, two different models (SEIRS, SEIQRS) based on small-world network are formulated for the spread of infectious diseases. The existence and global stability of the disease-free equilibrium and the endemic equilibrium for these two epidemic systems are proved by differential equations knowledge and Routh-Hurwitz theory. A numerical example, which includes key parameters analysis and critical topic discussion (e.g., medicine resources demand forecasting) is presented to test how well the proposed two models may be applied in practice.

The remainder of the paper is organized as follows. Section 2 presents the SEIRS model considering small-world network effect. Section 3 introduces the SEIQRS model considering small world network effect. Section 4 shows the numerical experiment and discussions. Finally, Section 5 proposes the conclusions.

2. SEIRS Model in Small-World Network (Model I)

2.1. Basic Introduction

For the compartment model of epidemic diffusion in a mature theory, herein, we omit the verbose introduction of the framework process. Readers who interest in such a topic can find more nutriment in [7, 11, 13, 14]. In this section, we consider the situation that infected person will not be quarantined, and we divide people in epidemic area into four groups: susceptible people (S), exposed people (E), infected people (I), and recovered people (R). A survey by Tham [18] shows that part of recovered people who are discharged from the healthcare department will be reinfected again. Thus, considering the small-world network of the social contact, the structure of susceptible-exposure-infective-recovered-susceptible (SEIRS) model is shown as Figure 1.

Notations used in the following sections are specified as follows:

N : population size in epidemic area,

$S(t)$: number of susceptible people, $s(t) = S(t)/N$,

$E(t)$: number of exposed people, $e(t) = E(t)/N$,

$I(t)$: number of infected people, $i(t) = I(t)/N$,

$R(t)$: number of recovered people, $r(t) = R(t)/N$,

$\langle k \rangle$: average degree distribution of small world network,

β : propagation coefficient of the epidemic,

γ : reinfected rate of recovered people,

μ : recovered rate,

τ : incubation period of the epidemic,

d : death rate of infected people.

Intuitively, we have the first two equations:

$$\begin{aligned} S(t) + E(t) + I(t) + R(t) &= N, \\ s(t) + e(i) + i(t) + r(t) &= 1. \end{aligned} \quad (2.1)$$

Based on mean-field theory [19], the time-based parameter $s(t)$ meets the following equation from time t to $t + \Delta t$:

$$s(t + \Delta t) - s(t) = -\beta \langle k \rangle s(t) i(t) \Delta t + \gamma r(t) \Delta t. \quad (2.2)$$

Thus, we get

$$\frac{s(t + \Delta t) - s(t)}{\Delta t} = -\beta \langle k \rangle s(t) i(t) + \gamma r(t). \quad (2.3)$$

It can be rewritten as

$$\frac{ds(t)}{dt} = -\beta \langle k \rangle s(t) i(t) + \gamma r(t). \quad (2.4)$$

Similarly, we have the other three ordinary differential equations as follows:

$$\begin{aligned} \frac{de(t)}{dt} &= \beta \langle k \rangle s(t) i(t) - \beta \langle k \rangle s(t - \tau) i(t - \tau), \\ \frac{di(t)}{dt} &= \beta \langle k \rangle s(t - \tau) i(t - \tau) - di(t) - \mu i(t), \\ \frac{dr(t)}{dt} &= \mu i(t) - \gamma r(t). \end{aligned} \quad (2.5)$$

Thus, the SEIRS epidemic diffusion model which considers small-world network effect can be formulated as follows (Model I):

$$\begin{aligned} \frac{ds(t)}{dt} &= -\beta \langle k \rangle s(t) i(t) + \gamma r(t), \\ \frac{de(t)}{dt} &= \beta \langle k \rangle s(t) i(t) - \beta \langle k \rangle s(t - \tau) i(t - \tau), \\ \frac{di(t)}{dt} &= \beta \langle k \rangle s(t - \tau) i(t - \tau) - di(t) - \mu i(t), \\ \frac{dr(t)}{dt} &= \mu i(t) - \gamma r(t). \end{aligned} \quad (2.6)$$

Here, $\beta, \langle k \rangle, \gamma, \mu, d, \tau > 0$. Initial conditions for this epidemic diffusion model are demonstrated as follows:

$$\begin{aligned} i(0) &= i_0 \ll 1, \\ e(0) &= \langle k \rangle i(0), \\ s(0) &= 1 - e_0 - i_0, \\ r(0) &= 0. \end{aligned} \tag{2.7}$$

2.2. Analysis of the SEIRS Model

As to Model I, while such an epidemic diffusion system is stable, the number of people in different groups will be unchanged. Hence, we have $s(t) = s(t - \tau)$, $i(t) = i(t - \tau)$, and we get

$$\frac{de(t)}{dt} = \beta \langle k \rangle s(t) i(t) - \beta \langle k \rangle s(t - \tau) i(t - \tau) = 0, \tag{2.8}$$

$$\frac{di(t)}{dt} = \beta \langle k \rangle s(t) i(t) - di(t) - \mu i(t). \tag{2.9}$$

Equation (2.8) means that the number of exposed people is constant when epidemic diffusion system is stable. As we all know, if an epidemic is wide spread, it should satisfy the following condition:

$$\left. \frac{di(t)}{dt} \right|_{t=0} > 0. \tag{2.10}$$

Together, this equation with (2.9), we can get

$$s_0 > \frac{d + \mu}{\beta \langle k \rangle}. \tag{2.11}$$

Equation (2.11) shows that the spread of epidemic outbreaks only when s_0 meets the above condition. As $s(t) + e(t) + i(t) + r(t) = 1$, combined with (2.6), we get

$$\begin{aligned} \frac{ds(t)}{dt} &= -\beta \langle k \rangle s(t) i(t) + \gamma(1 - s(t) - e(t) - i(t)), \\ \frac{di(t)}{dt} &= \beta \langle k \rangle s(t) i(t) - di(t) - \mu i(t). \end{aligned} \tag{2.12}$$

Let $ds(t)/dt = 0$ and $di(t)/dt = 0$, we can get an obvious equilibrium point for such an epidemic diffusion model as follows:

$$P_1 = (s, i) = (1, 0). \tag{2.13}$$

As (2.13) shows, the number of infected people is zero, which indicates that spread of epidemics in such an area does not happen. All people are susceptible individuals. Herein, we refer to such a point as the disease-free equilibrium point.

In the other side, according to (2.8), the number of exposed people is constant. Thus, combined with (2.12), we can get another equilibrium result as follows:

$$P_2 = (s, i) = \left(\frac{d + \mu}{\beta \langle k \rangle}, \frac{\gamma [\beta \langle k \rangle (1 - e) - (d + \mu)]}{\beta \langle k \rangle (\gamma + d + \mu)} \right). \quad (2.14)$$

Such a result shows that, when epidemic diffusion system is stable, a certain amount of infected people exist in disaster area. Herein, we refer it as the endemic equilibrium point.

Lemma 2.1. *Disease-free equilibrium point P_1 is stable only when $\beta < (d + \mu) / \langle k \rangle$.*

Proof. As $P_1 = (s, i) = (1, 0)$, we can obtain the Jacobi matrix of (2.12) as follows:

$$J_{P_1} = \begin{bmatrix} \frac{\partial \Pi_1}{\partial s} & \frac{\partial \Pi_1}{\partial i} \\ \frac{\partial \Pi_2}{\partial s} & \frac{\partial \Pi_2}{\partial i} \end{bmatrix} = \begin{bmatrix} -\gamma & -\beta \langle k \rangle - \gamma \\ 0 & \beta \langle k \rangle - d - \mu \end{bmatrix}. \quad (2.15)$$

Here, Π_1 and Π_2 are the two differential equations in (2.12). The secular equation for the Jacobi matrix is

$$(\lambda + \gamma)(\lambda - \beta \langle k \rangle + d + \mu) = 0. \quad (2.16)$$

It is easy to get the two characteristic roots for this secular equation, which are $-\gamma$ and $\beta \langle k \rangle - d - \mu$. Based on Routh-Hurwitz stability criterion, when $\beta < (d + \mu) / \langle k \rangle$, real parts of the two characteristic roots are negative. Thus, the disease-free equilibrium point $P_1 = (s, i) = (1, 0)$ is stable only when $\beta < (d + \mu) / \langle k \rangle$. \square

Lemma 2.2. *Endemic equilibrium point P_2 is stable only when $\beta > (d + \mu) / \langle k \rangle (1 - e)$.*

Proof. Similarly as Lemma 2.1, coupling with (2.14), the Jacobi matrix of (2.12) can be rewritten as follows:

$$J_{P_2} = \begin{bmatrix} \frac{\gamma [-\beta \langle k \rangle (1 - e) - \gamma]}{(\gamma + d + \mu)} & -d - \mu - \gamma \\ \frac{\gamma [\beta \langle k \rangle (1 - e) - (d + \mu)]}{(\gamma + d + \mu)} & 0 \end{bmatrix}. \quad (2.17)$$

The secular equation for (2.17) can be expressed as follows:

$$a\lambda^2 + b\lambda + c = 0. \quad (2.18)$$

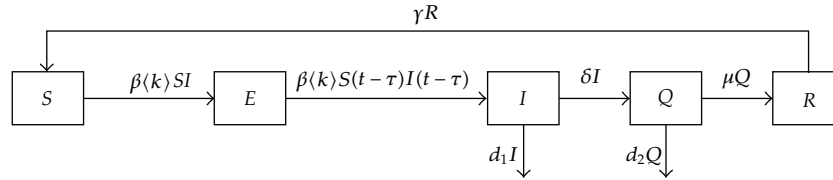


Figure 2: Framework of Model II.

Herein, $a = 1$, $b = \gamma[\beta\langle k\rangle(1-e) + \gamma]/(\gamma + d + \mu)$, and $c = \gamma[\beta\langle k\rangle(1-e) - (d + \mu)]$. Based on the quadratic equation theory, such a secular equation contains two characteristic roots λ_1 and λ_2 , and satisfies

$$\lambda_1 + \lambda_2 = -\frac{b}{a} = -\frac{\gamma[\beta\langle k\rangle(1-e) + \gamma]}{(\gamma + d + \mu)} < 0, \quad (2.19)$$

$$\lambda_1 \cdot \lambda_2 = \frac{c}{a} = \gamma[\beta\langle k\rangle(1-e) - (d + \mu)]. \quad (2.20)$$

According to Routh-Hurwitz stability criterion, if we want to get two negative characteristic roots λ_1 and λ_2 again, the (2.20) should be constant greater than zero, which means that $\beta > (d + \mu)/\langle k\rangle(1-e)$ should be satisfied. Thus, only when $\beta > (d + \mu)/\langle k\rangle(1-e)$, the endemic equilibrium point P_2 is stable. \square

3. SEIQRS Model Considering Small-World Network Effect (Model II)

3.1. Basic Introduction

In this section, the quarantine measure are considered in disaster area, and people in epidemic area are divided into five groups here: susceptible people (S), exposed people (E), infected people (I), quarantined people (Q), and recovered people (R). Framework of epidemic diffusion in this section is illustrated as Figure 2.

Likewise, the SEIQRS epidemic diffusion model which considers small-world network effect can be formulated as follow (Model II):

$$\begin{aligned} \frac{ds(t)}{dt} &= -\beta\langle k\rangle s(t)i(t) + \gamma r(t), \\ \frac{de(t)}{dt} &= \beta\langle k\rangle s(t)i(t) - \beta\langle k\rangle s(t-\tau)i(t-\tau), \\ \frac{di(t)}{dt} &= \beta\langle k\rangle s(t-\tau)i(t-\tau) - d_1 i(t) - \delta i(t), \\ \frac{dq(t)}{dt} &= \delta i(t) - d_2 q(t) - \mu q(t), \\ \frac{dr(t)}{dt} &= \mu q(t) - \gamma r(t). \end{aligned} \quad (3.1)$$

Here, $Q(t)$ stands for number of quarantined people, $q(t) = Q(t)/N$. $s(t) + e(t) + i(t) + q(t) + r(t) = 1$. d_1 is death rate of infected people, d_2 is death rate of quarantined people, and δ is quarantine rate of infected people. Moreover, $\beta, \langle k \rangle, \gamma, \delta, \mu, d_1, d_2, \tau > 0$. Initial conditions for Model II are presented as follows:

$$\begin{aligned} i(0) &= i_0 \ll 1, \\ e(0) &= \langle k \rangle i(0), \\ s(0) &= 1 - e_0 - i_0, \\ q(0) &= 0, \\ r(0) &= 0. \end{aligned} \tag{3.2}$$

3.2. Analysis of the Epidemic Diffusion Model

Similar as Section 2.2, combined with (2.8) and (3.2), the (3.1) can be converted as follows:

$$\begin{aligned} \frac{ds(t)}{dt} &= -\beta \langle k \rangle s(t)i(t) + \gamma [1 - s(t) - e(t) - i(t) - q(t)], \\ \frac{di(t)}{dt} &= \beta \langle k \rangle s(t)i(t) - (d_1 + \delta)i(t), \\ \frac{dq(t)}{dt} &= \delta i(t) - (d_2 + \mu)q(t). \end{aligned} \tag{3.3}$$

Let $ds(t)/dt = 0$, $di(t)/dt = 0$ and $dq(t)/dt = 0$, two equilibrium points for Model II can be solved and shown as follows:

$$\begin{aligned} P_3 &= (s, i, q) = (1 - e, 0, 0), \\ P_4 &= (s, i, q) = \left(\frac{d_1 + \delta}{\beta \langle k \rangle}, B, \frac{\delta}{d_2 + \mu} B \right). \end{aligned} \tag{3.4}$$

Herein, $B = \gamma [\beta \langle k \rangle (1 - e) - (d_1 + \delta)] (d_2 + \mu) / \beta \langle k \rangle [(d_1 + \delta + \gamma)(d_2 + \mu) + \gamma \delta]$, P_3 is the disease-free equilibrium point, and P_4 is the endemic equilibrium point.

Lemma 3.1. Disease-free equilibrium point P_3 is stable only when $\beta < (d_1 + \delta) / \langle k \rangle (1 - e)$.

Lemma 3.2. Endemic equilibrium point P_4 is stable only when $\beta > (d_1 + \delta) / \langle k \rangle (1 - e)$.

The proof procedure for Lemmas 3.1 and 3.2 is similar as in Section 2. Thus, it is trivial to do the work again.

Remark 3.3. From the above four lemmas, we can get the first conclusion that threshold of the epidemic diffusion depends on some key parameters, such as average degree distribution of the small-world network, recovered rate, death rate of infected people, and also a number of exposed people when the system is stable.

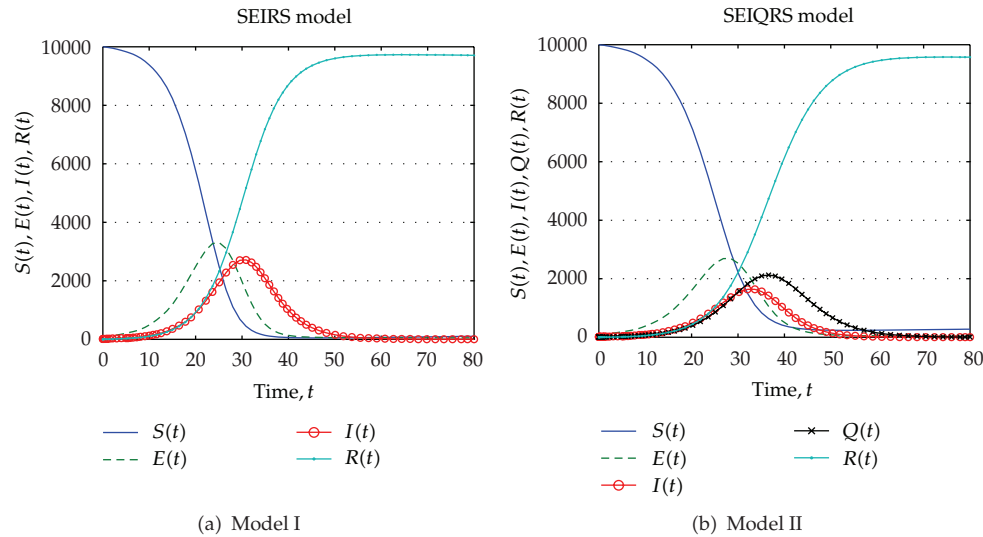


Figure 3: Numerical simulation for Model I and Model II.

4. Numerical Simulation and Discussion

4.1. Numerical Experiment

In this section, we take a numerical simulation to test how well the proposed two models may be applied in practice. The initial values of parameters in the proposed epidemic diffusion models are acquired from [20] and by interviews with public health care administrative personnel, which are given as follows: $\beta = 2 \times 10^{-5}$, $\langle k \rangle = 6$, $\gamma = 2 \times 10^{-4}$, $\delta = 0.3$, $\mu = 0.2$, $d = d_1 = 5 \times 10^{-3}$, $d_2 = 1 \times 10^{-3}$, $\tau = 5$ (day), $N = 10^4$, and $i(0) = 1 \times 10^{-3}$. We use MATLAB 7.0 mathematical solver together with Runge-Kutta method to simulate these two models. The tests are performed on an Intel(R) Core(TM) i3 CPU 2.4 GHz with 2 GB RAM under Microsoft Windows XP. Figure 3 is the numerical simulation of these two epidemic models. The curves respectively represent the different groups of people over time.

As Figure 3 shows, both infected curves $I(t)$ in Model I and Model II exhibit similar trends, namely, the number of infected people will first increase along with the spreading the epidemic, and then it will decrease after the epidemic is brought under control. Threshold value of epidemic diffusion exists in both Model I and Model II. The rush of $I(t)$ is around on the 32-33 day. It seems that the quarantine measure does not work in our example. On the contrary, by comparing Figure 3(a) with Figure 3(b), one can observe that the number of infected people in Model II is way below it in Model I, suggesting that the quarantine measure will significantly reduce the infectivity in the disaster area. Similarly, trend can be found in $E(t)$ curve. We conclude all these result owe to the quarantine measure in Model II.

Based on the theory analysis in Sections 2 and 3, one can observe that some factors, such as β and $\langle k \rangle$, are key parameters in epidemic diffusion system. Herein, we present a short sensitivity analysis for them. Holding all the other parameters fixed as given in the numerical example, except that β takes on four different values ranging from $\beta = 2 \times 10^{-5}$ to 8×10^{-5} with an increment of 2×10^{-5} , Figure 4 shows that number of infected people changed over time. As Figure 4 shows, one can observe that, no matter in Model I or Model II, there

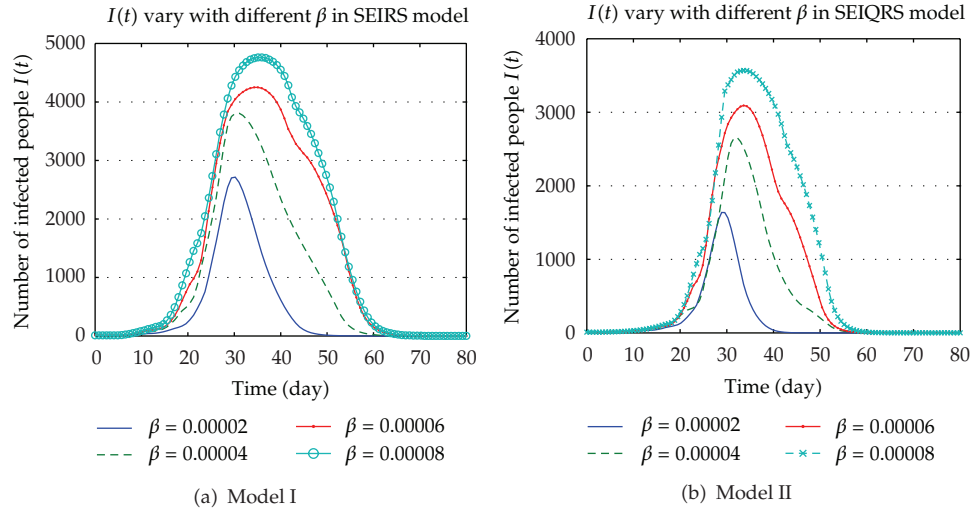


Figure 4: Number of infected people with different β .

almost have no differences among these curves in the first 15–20 days. However, differences are obvious in the following days. The larger initial size of β is, the faster increments speed is. Figure 4 tells us that controlling the propagation coefficient is an effective way to prevent the spread of the epidemic. By comparison with Figures 4(a) and 4(b), a number of infected people in Model II is also way below it in Model I, such result enlightens us again that the quarantine measure is an important factor in epidemic controlling.

It holds all the other parameters are fixed as given in the numerical example, except that $\langle k \rangle$ takes on four different values ranging from 4 to 10 with an increment of 2. Figure 5 shows that the number of infected people is changed over time. Similarly, no matter in Model I or in Model II, a number of infected people shows a positive proportional to the parameter $\langle k \rangle$. Figure 5 transfers an important information, that is, self-quarantine and decreasing the contact with people around are effective strategies for controlling epidemic diffusion. Such a conclusion explains that why Chinese government implements a series of strict quarantine measures during the SARS period.

4.2. A Critical Topic Discussion

The purpose of dynamic analysis of epidemic diffusion models considering small-world network effect in this paper is to depict the epidemic diffusion rule and provide guidance for the emergency management practice. One critical topic of emergency management practice is to forecast the medicine resources demand. However, it is often difficult to predict the actual demand based on historical data (for many events, the historical data may not even exist). Moreover, before 2003, when SARS hit Asia, such an operation in some parts of China has always been done unsystematically (based on the decision maker's experience), which leads to stock-out or surplus phenomenon happening occasionally.

Based on dynamic analysis of epidemic diffusion models in the above sections, herein, we are going to discuss how to forecast the time-varying demand in disaster area. Let $D(t)$ represents demand for medicine resources in disaster area at time t . Obviously, the more

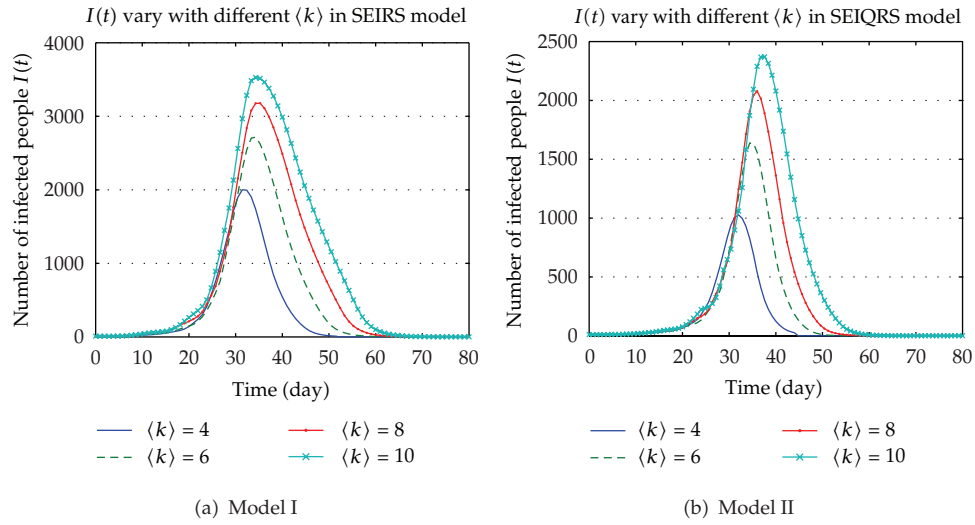


Figure 5: Number of infected people with different $\langle k \rangle$.

people infected, the more resources would be demanded, which can be rewritten as

$$D(t) \propto f[I(t) + Q(t)]. \quad (4.1)$$

Furthermore, each infected person or quarantined person should be cured for certain time (the cure cycle), for example, 30 days. During the cure cycle, demand of medicine resources for each person may shows a property of nonlinearity. Here, we use a function $\varphi(t)$ to represent it. Hence, the total demand of medicine resources for each infected/quarantined person is

$$\varphi = \int_0^c \varphi(t) dt, \quad (4.2)$$

where c is the cure cycle.

To Model I, average demand for medicine resources at time t can be formulated as

$$D_I(t) = \frac{I(t) \cdot \varphi}{c} = \frac{N}{c} \int_0^c \varphi(t) dt \int_0^t \beta \langle k \rangle s(t-\tau) i(t-\tau) - di(t) - \mu i(t) dt. \quad (4.3)$$

Hence, we get

$$\frac{dD_I(t)}{dt} = \Lambda [\beta \langle k \rangle s(t-\tau) i(t-\tau) - di(t) - \mu i(t)], \quad (4.4)$$

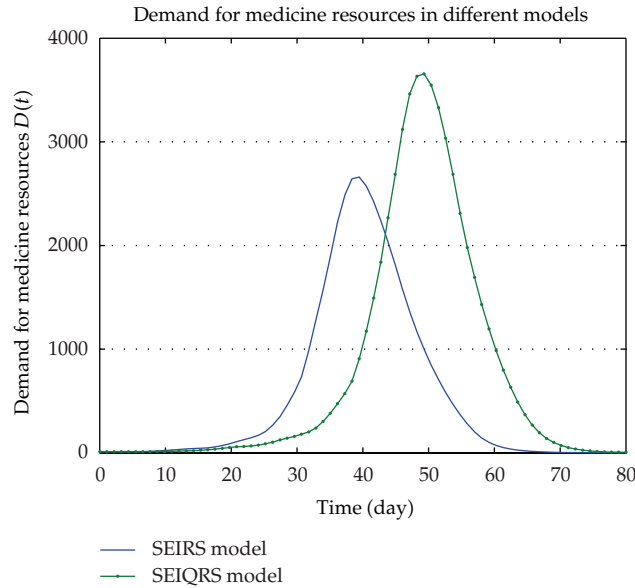


Figure 6: Demand for medicine resources in different models.

where Λ is a constant. Similarly, to Model II, we have

$$\begin{aligned}
 D_{II}(t) &= \frac{[I(t) + Q(t)] \cdot \varphi}{c} \\
 &= \frac{N}{c} \int_0^c \varphi(t) dt \left(\int_0^t \beta \langle k \rangle s(t-\tau) i(t-\tau) - d_1 i(t) - \delta i(t) dt + \int_0^t \delta i(t) - d_2 q(t) - \mu q(t) dt \right),
 \end{aligned} \tag{4.5}$$

$$\frac{dD_{II}(t)}{dt} = \Lambda [\beta \langle k \rangle s(t-\tau) i(t-\tau) - d_1 i(t) - d_2 q(t) - \mu q(t)]. \tag{4.6}$$

To depict the variation trend of demand for medicine resources, here, we let $\Lambda = 1$. Such an operation will not affect the final result. According to (4.4) and (4.6), holding all the parameters fixed as given in the above numerical example, one can get the demand for medicine resources by these two models as Figure 6 shows.

As Figure 6 shows, variation trend of demand for medicine resources presents an obvious three different stages. At the first stage (e.g., 0–15 days), epidemic has just outbreak, and it has not yet caused a widespread diffusion. Such a period is the best rescue time. Demand for medicine resources in this period are kept in low level. Hence, medicine resources inventory in local health departments should be distributed to infected people as quickly as possible. If the rescue opportunities in first stage are missed, epidemic will cause a widespread diffusion in the following time, which brings us to the second stage. In such a stage (e.g., 15–70 days), demand for medicine resources is time varying, resources distribution program in such a stage should be varied over time correspondingly. All these changes make the emergency work much more trouble. At the third stage (e.g., 70–more days), demand for medicine resources shows stable again. The inventory of medicine

resources in local health departments should be replenished at this period; meanwhile, a fraction of medicine resources should be allocated to the remaining infected areas.

5. Conclusions

In this paper, we analyze two different epidemic models (SEIRS, SEIQRS) considering small-world network effect, and we prove the global stability of the disease-free equilibrium and the endemic equilibrium for them. A numerical example, which includes key parameters analysis and critical topic discussion, is presented to test how well the proposed two models may be applied in practice. The novelty of our model against the existing works in literature is characterized by the following aspects

(1) While most research on epidemic diffusion studies a compartment model taking no consideration the complexity of social contact network, the models proposed in this paper address the small world network effect in match of the course of an epidemic diffusion.

(2) Based on the theory analysis and numerical simulation, epidemic diffusion rules are depicted, and a series of suggestions for emergency management practice is presented. These methods remarkably outperform the traditional measurements and will be much more suitable for real operations.

As the limitations of the models, they are developed in geographic area where an epidemic disease has been spreading and does not consider possible cross area diffusion between two or more geographic areas. We assume that once an epidemic outbreaks, the government has effective means to separate the epidemic areas so that cross-area spread can be basically prevented. However, this cannot always be guaranteed in reality.

Acknowledgments

This work has been partially supported by the MOE (Ministry of Education in China) Project of Humanities and Social Sciences (Project no. 11YJCZH109), and by the Project of the Nanjing University of Science and Technology (2011YBXM96), (2011XQTR10), (AE88072), and (JGQN1102).

References

- [1] L. M. Wein, D. L. Craft, and E. H. Kaplan, "Emergency response to an anthrax attack," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 7, pp. 4346–4351, 2003.
- [2] L. M. Wein, Y. Liu, and T. J. Leighton, "HEPA/vaccine plan for indoor anthrax remediation," *Emerging Infectious Diseases*, vol. 11, no. 1, pp. 69–76, 2005.
- [3] D. L. Craft, L. M. Wein, and A. H. Wilkins, "Analyzing bioterror response logistics: the case of anthrax," *Management Science*, vol. 51, no. 5, pp. 679–694, 2005.
- [4] E. H. Kaplan, D. L. Craft, and L. M. Wein, "Emergency response to a smallpox attack: the case for mass vaccination," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 16, pp. 10935–10940, 2002.
- [5] E. H. Kaplan, D. L. Craft, and L. M. Wein, "Analyzing bioterror response logistics: the case of smallpox," *Mathematical Biosciences*, vol. 185, no. 1, pp. 33–72, 2003.
- [6] H. Matsuura, K. Koide, N. Noda, T. Nemoto, M. Nakano, and K. I. Makino, "Stochastic dynamics in biological system and information," *International Journal of Innovative Computing, Information and Control*, vol. 4, no. 2, pp. 233–248, 2008.
- [7] P. L. Shi and L. Z. Dong, "Dynamical models for infectious diseases with varying population size and vaccinations," *Journal of Applied Mathematics*, vol. 2012, Article ID 824192, 20 pages, 2012.

- [8] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [9] S. Eubank, H. Guclu, V. S. A. Kumar et al., "Modelling disease outbreaks in realistic urban social networks," *Nature*, vol. 429, no. 6988, pp. 180–184, 2004.
- [10] M. M. Telo Da Gama and A. Nunes, "Epidemics in small world networks," *European Physical Journal B*, vol. 50, no. 1-2, pp. 205–208, 2006.
- [11] J. Saramäki and K. Kaski, "Modelling development of epidemics with dynamic small-world networks," *Journal of Theoretical Biology*, vol. 234, no. 3, pp. 413–421, 2005.
- [12] N. Masuda and N. Konno, "Multi-state epidemic processes on complex networks," *Journal of Theoretical Biology*, vol. 243, no. 1, pp. 64–75, 2006.
- [13] X. J. Xu, H. O. Peng, X. M. Wang, and Y. H. Wang, "Epidemic spreading with time delay in complex networks," *Physica A*, vol. 367, pp. 525–530, 2006.
- [14] X. P. Han, "Disease spreading with epidemic alert on small-world networks," *Physics Letters, Section A*, vol. 365, no. 1-2, pp. 1–5, 2007.
- [15] T. E. Stone, M. M. Jones, and S. R. McKay, "Comparative effects of avoidance and vaccination in disease spread on a dynamic small-world network," *Physica A*, vol. 389, no. 23, pp. 5515–5520, 2010.
- [16] C. I. Hsu and H. H. Shih, "Transmission and control of an emerging influenza pandemic in a small-world airline network," *Accident Analysis and Prevention*, vol. 42, no. 1, pp. 93–100, 2010.
- [17] H. Y. Wang, X. P. Wang, and A. Z. Zeng, "Optimal material distribution decisions based on epidemic diffusion rule and stochastic latent period for emergency rescue," *International Journal of Mathematics in Operational Research*, vol. 1, no. 1-2, pp. 76–96, 2009.
- [18] K. Y. Tham, "An emergency department response to severe acute respiratory syndrome: a prototype response to bioterrorism," *Annals of Emergency Medicine*, vol. 43, no. 1, pp. 6–14, 2004.
- [19] J. Marro and R. Dickman, *Nonequilibrium Phase Transitions in Lattice Models*, Collection Aléa-Saclay: Monographs and Texts in Statistical Physics, Cambridge University Press, Cambridge, UK, 1999.
- [20] M. Liu and L. Zhao, "Analysis for epidemic diffusion and emergency demand in an anti-bioterrorism system," *International Journal of Mathematical Modelling and Numerical Optimisation*, vol. 2, no. 1, pp. 51–68, 2011.

Research Article

Stability and Limit Oscillations of a Control Event-Based Sampling Criterion

M. De la Sen,¹ J. C. Soto,² and A. Ibeas³

¹ Institute for Research and Development of Processes, University of Basque Country, Campus of Leioa (Bizkaia), Apartado 644, Bilbao, Leioa 48090, Spain

² Department of Applied Mathematics, University of Basque Country, Campus of Bilbao (Bizkaia), Apartado 644, Bilbao 48012, Spain

³ Department of Telecommunications and Systems Engineering, Autonomous University of Barcelona, Bellaterra 08193, Barcelona, Spain

Correspondence should be addressed to M. De la Sen, manuel.delasen@ehu.es

Received 24 February 2012; Revised 20 April 2012; Accepted 26 April 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 M. De la Sen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper investigates the presence of limit oscillations in an adaptive sampling system. The basic sampling criterion operates in the sense that each next sampling occurs when the absolute difference of the signal amplitude with respect to its currently sampled signal equalizes a prescribed threshold amplitude. The sampling criterion is extended involving a prescribed set of amplitudes. The limit oscillations might be interpreted through the equivalence of the adaptive sampling and hold device with a nonlinear one consisting of a relay with multiple hysteresis whose parameterization is, in general, dependent on the initial conditions of the dynamic system. The performed study is performed on the time domain.

1. Introduction

Nonperiodic sampling theory opens a set of new technical possibilities compared with the classical sampling with periodic sampling period [1–17]. Those possibilities are as follows:

- (1) to adapt the sampled signals to get better performances [1–3, 6, 8, 9, 16]: for instance, if that signal varies rapidly, then the sampling period is made smaller and vice versa. In general, some constraints in terms of bandwidth, stability, and technical requirements related to circuitry or computing should be respected so that the sampling rate has to belong to some appropriate admissibility domain;
- (2) transmission errors from data to results in algebraic problems like controllability, and observability, might be reduced by a judicious selection of the sampling

instants when choosing a fixed sampling period. The reason is that the condition number of the matrix defining the problem depends on such a choice and one can convert a one-parameter optimization problem (a fixed sampling period) into a multiple one (the whole set of distinct sampling periods). In particular, the smaller the condition number of the coefficient matrix is, the smaller are the relative transmission errors from the data to the results depending on each particular problem dealt with [13, 14, 16]. The technique might be used by its “ad-hoc” implementation in a great variety of problems like biology measurements, economics, control theory and engineering, [16], statistics, random sampling [18–22];

- (3) to improve the adaptation transients in recursive identification or adaptive control of both classical or hybrid systems by combining the estimation algorithm with the signal adaptation, [2, 3, 9, 16]. Related adaptive sampling techniques can be used in the context of expert systems to improve the performances under supervisory rules (see, e.g., [17] and references therein).

Nonperiodic sampling being updated under certain adaptive sampling laws can often be interpreted as event-driven [23, 24], since, although sampling occurs through time, most of sampling rules involve signal comparison rules related to their immediate previous sampled values or involve certain performance tests. There are a set of background interesting papers, available in the literature, in which sampling is considered either state-dependent, random, or based in stochastic considerations, in general, and used in a number of applications. See, for instance, [19–22, 25–27] and references therein. The constant difference of amplitudes sampling criterion consists of keeping constant the absolute increment of the signal being sampled inbetween each two consecutive sampling instants. The sampling criterion together with its associate sampling and zero-order-hold device is equivalent to a separate nonlinearity which is fully equivalent to a multiple relay with hysteresis (i.e., a multiple bang-bang device with hysteresis). See [1, 15, 18, 28–32] and some references there in. In particular, the sampling criterion based on constant difference of amplitudes was generalized in [1] to the use of several threshold amplitudes the initial sampling criterion proposal of [30] based on a single constant difference of amplitudes. This equivalence motivates that the discretized system exhibits some properties being commonly associated with certain nonlinear systems, like for instance, the potential existence of limit oscillations. A close nonlinear model was proposed in [21] for feedback-based stabilization by triggering the plant output samples through the crossings, with hysteresis, of the signal through its quantization levels. In [22], a close problem related to saturating quantized measurements is focused on. It is well known that limit cycles are highly unsuitable in applications where the objective is to get a zero asymptotic tracking errors. However, they are pursued as an objective for the design of oscillators in some applications as in the design of tank circuits for tuning a suited frequency in radio or TV. There are unified sampling formulations available in the background literature including the presence of sampling constraints in [33, 34] and references there in, and work is also in progress to extend results to the presence of internal delays [35].

This paper characterizes and formalizes mathematically in the time domain the above sampling criterion by extending some previous background results in [1, 18], where the study of oscillations was only approximate and made in a first-harmonic approximation in the frequency domain by using the describing function approach, while the stability properties were not investigated. The results are obtained in the time domain rather than in the frequency one. This allows not to necessarily assume in the problem statement that the linear

dynamics exhibits low-pass filtering properties so as to justify the use of a first-harmonic approximation method, as it was done in [1, 15, 18] which was an important limitation in those papers. In this way, such an assumption is no longer needed in the subsequent study which is performed with an exact analytic treatment rather than involving an approximate one. Also, a set of difference amplitudes, rather than a constant fixed one, are allowed in a generalized version of that sampling criterion in order to generalize the problem and to improve its potential applicability. If there is just a single amplitude available to be used as adaptive sampling threshold, then the sampling criterion is referred to as *constant amplitude difference sampling criterion (CADSC)*. If several amplitudes are used, then the sampling criterion is referred to as *sampling-dependent amplitude difference sampling criterion (SDADSC)*. Note that the model obtained in [4] for nonperiodically sampled systems is basically a linear time-varying difference equation. This model is useful to describe discretized systems under varying sampling periods. The time-varying coefficients of the discrete equation depend on both the sequence of sampling periods and the continuous-time parameters. Thus, it may be applied also to the criterion of constant difference of amplitudes. However, some properties like, for instance, the ability of generating limit oscillations are not easily discovered from an earlier inspection when using such a time-varying linear equation. The analysis method for stability and limit oscillations consists basically of the following steps:

- (a) describe the linear uncontrolled continuous-time system by an ordinary differential equation of n th order submitted to a piecewise constant control input which varies at a set of sampling instants with, in general, time-varying sampling periods. The “ad hoc” control device for this purpose is referred to as a sampling and hold device. The solution of such a differential system is referred to as the “output” of the system;
- (b) discretize the equivalent differential system of n th order at generic sampling instants. Since the input is piecewise constant with discontinuities at such time instants, the solution of the differential equation for any given initial conditions coincides with that of the discretized system at sampling instants. The feedback law for a regulator with unity feedback is introduced so that the piecewise constant feedback control takes the minus values of the output at sampling instants;
- (c) define the generic sampling instants as those generated by the event-driven law of constant absolute difference of amplitudes of the feedback error inbetween each two consecutive sampling instants. This is generalized for a set of prescribed amplitudes in a more general sampling criterion. The amplitude, or the set of amplitudes, parameterize the solution together with the parameters of the continuous-time differential equation. It is seen that the zero-order and hold device together with the sampling criterion is equivalent to a relay with a multiple hysteresis. This suggests that limit cycles of the solution can potentially exist;
- (d) limit cycles are found by investigating double points of the solution in the time domain.

The dynamic system studied in this paper is complex in the sense that a continuous-time dynamic system is controlled by a feedback law consisting of an adaptive sampling criterion which is based on the use of a set of threshold amplitudes to calculate the sequence of sampling instants. For a second-order case study given in Section 5, it is shown that the zero-order hold used for discretization plus the adaptive sampling criterion itself are jointly equivalent to a relay device with multiple hysteresis. The whole feedback type is

hybrid since it consist of a continuous-time system under nonlinear feedback and, in this sense, the whole system is a complex dynamic system. The equivalent multiple relay with hysteresis nonlinearity in the feedback-loop allows to interpret the presence of sustained limit oscillations as an asymptotic solution of the state-space trajectory of the closed-loop system.

2. Some Preliminary Framework and Basic Results

Notation. \mathbf{R} is the set of real numbers, $\mathbf{R}_{0+} := \{\mathbf{R} \ni z \geq 0\}$ and $\mathbf{R}_+ := \{\mathbf{R} \ni z > 0\} \equiv \mathbf{R}_{0+} \setminus \{0\}$:

- (i) \mathbf{N} the set of natural numbers, $\mathbf{N}_0 = \mathbf{N} \cup \{0\}$ and $\bar{k} := \{1, 2, \dots, k\} \subset \mathbf{N}$ is the set of natural numbers ranging from 1 to k ;
- (ii) $\text{PC}(\mathbf{R}_{0+}, \mathbf{R})$ is the set of piecewise continuous functions on \mathbf{R}_{0+} ;
- (iii) $\text{PC}^{(n-1)}([0, T_{\text{per}}]; \mathbf{R})$ is the set of real almost everywhere piecewise $(n-1)$ th continuous-time differentiable functions on the definition domain $[0, T_{\text{per}}]$;
- (iv) I_n is the n th order identity matrix;
- (v) the disjunction logic rule (spelled “or”) and the conjunction logic rule (spelled “and”) are denoted by the symbols \vee and \wedge , respectively;
- (vi) the ℓ_2 (or spectral) vector norm of $z \in \mathbf{R}^q$ is defined as $\|z\|_2 = \sqrt{z^T z}$ (with the superscript “ T ” standing for transposition. The ℓ_2 -vector norm coincides with the Froebenius or Euclidean vector norm;
- (vii) for a real matrix $M \in \mathbf{R}^{p \times q}$, its ℓ_2 -induced matrix norm is

$$\begin{aligned} \|M\|_2 &:= \max \left(\frac{\|Mz\|_2}{\|z\|_2} : 0 < \|z\|_2 \leq 1 \right) = \max \left(\sqrt{z^T z} : \|z\|_2 = 1 \right) \\ &= \max \left(\left| \lambda_i(M^T M) \right|^{1/2} : \lambda_i \in \sigma(M^T M); \forall i \in \bar{n}_\sigma \right), \end{aligned} \quad (2.1)$$

where $\sigma(M^T M)$ is the spectrum of the square matrix $M^T M$ consisting of $1 \leq n_\sigma \leq q$ distinct real eigenvalues $\lambda_i; i \in \bar{n}_\sigma$. The above positive real maximum defining the spectral $\|M\|_2$ will be denoted by $\lambda_{\max}(M^T M)$. If $q = p$, then $\|M\|_2^2 = \|M^T M\|_2 = \lambda_{\max}(M^T M) = |\lambda_{\max}(M)|^2$;

- (viii) $f \in C_T(\mathbf{R}^p \times [t_k, t_{k+1}]; \mathbf{R}^q)$ is a testing real vector function $f: \mathbf{R}^p \times [t_k, t_{k+1}) \rightarrow \mathbf{R}^q$ within a testing class C_T being of the form $f(x_\tau, \tau)$, where “ s ” stands for cartesian product of sets, with x_τ being a real p -dimensional strip on $[t_k, t_{k+1})$ where t_k and t_{k+1} are two consecutive sampling instants from some sampling criterion SC. Thus, f is a piecewise real vector function from \mathbf{R}^p to \mathbf{R}^q on $[t_k, t_{k+1})$ valued at some argument vector function $x: \mathbf{R}^p \times [t_k, t_{k+1}) \rightarrow \mathbf{R}^p$.

Consider the ordinary linear time-invariant differential equation:

$$A(D)y(t) = B(D)u(t), \quad D^i y(0) = y^{(i)}(0) \in \mathbf{R}, \quad (2.2)$$

under a piecewise continuous control input $u \in \text{PC}(\mathbf{R}_{0+}, \mathbf{R})$ to be specified later on, where $\mathbf{R}_{0+} := \{\mathbf{R} \ni z \geq 0\}$ and the polynomials $A(D)$ and $B(D)$ of real coefficients in the time-

derivative operator $D := d/dt$ (subject to $D^0 = 1$ and $D^i = DD^{i-1}$; for all $i \in \mathbf{N}$), which is formally equivalent to the Laplace transform argument “ s ,” are

$$A(D) = \sum_{i=0}^n a_i D^{n-i}, \quad B(D) = \sum_{i=0}^m b_i D^{m-i}, \quad (2.3)$$

where $a_0 \neq 0$, $b_0 \neq 0$ and $n := \deg(A(D)) \geq m := \deg(B(D))$ so that the transfer function $G(s) = B(s)/A(s)$ is realizable, where the Laplace argument “ s ” is formally equivalent to the time derivative operator “ $D = d/dt$.” It is assumed with no loss in generality that the polynomial $A(D)$ is monic, that is, $a_0 = 1$, since any other nonzero value $a_0 \neq 1$ can also lead to the differential equation (2.2) after normalization by a_0 of all the remaining coefficients of $A(D)$ and $B(D)$. It is also assumed that any potential zero cancellations in those polynomials, if any, are stable. This guarantees that the state-space realization is either minimal (i.e., no such cancellations exist) or, otherwise, any existing uncontrollable/unobservable mode is stable so that it does not contribute to the asymptotic solution as time tends to infinity. It is well known that the differential equation (2.2) can be described by a n th order differential system of first-order differential equations. Through the paper, it will be assumed that the differential system (2.2)-(2.3) will be controlled by a unity feedback control using, in general, nonperiodically samples $y(t_i)$; $i \in \mathbf{N}_0$ of the solution $y(t)$. The unity feedback control law is from a zero-order sampling and hold device by

$$u(t) = e(t_k) := y(t_k) - r(t_k); \quad \forall t \in [t_k, t_{k+1}), \quad (2.4)$$

where $r \in \text{PC}(\mathbf{R}_{0+}, \mathbf{R})$ and $e \in \text{PC}(\mathbf{R}_{0+}, \mathbf{R})$ are, the reference function and the error feedback respectively, and

$$\begin{aligned} \text{SI} &:= \left\{ t_k \in \mathbf{R}_{0+} : t_0 \in \mathbf{R}_{0+}, \infty \geq \bar{t} \geq t_{k+1} > t_k, \forall k \in \text{ID} \subset \mathbf{N}_0 \right\}, \\ \text{SP} &:= \left\{ T_k \in \mathbf{R}_+ : T_k := t_{k+1} - t_k \leq \bar{T} \leq \infty, \forall k \in \text{ID} \subset \mathbf{N}_0 \right\}, \end{aligned} \quad (2.5)$$

are, the totally ordered set of sampling instants of indicator set $\text{ID} \subset \mathbf{N}_0$ and the associated set of sampling periods with the same indicator set, respectively. Note that $u \in \text{PC}(\mathbf{R}_{0+}, \mathbf{R})$ and it is, in particular, piecewise constant. The following simple sampling process consistency result holds directly from (2.5).

Lemma 2.1. *Assume by convention, and with no loss in generality, that the first sampling instant $t_0 = 0$ and that $k \in \mathbf{N} \cap \text{ID} \Rightarrow (k-1) \in \text{ID}$. Then,*

$$\begin{aligned} \bar{t} < \infty &\iff \left[(\text{Card}(\text{ID}) = \text{Card}(\text{SI}) = \text{Card}(\text{SP}) = k+1 < \infty) \right. \\ &\quad \left. \wedge \left(\exists t_k := \max_{j \in \text{ID}}(t_j) \leq \bar{t} < \infty \right) \wedge \left(\bar{T} = \limsup_{j \rightarrow \infty} T_j = T_k = \infty \right) \right], \\ \bar{t} = \infty &\iff \left[(\text{Card}(\text{ID}) = \text{Card}(\text{SI}) = \text{Card}(\text{SP}) = \aleph_0) \wedge \left(\neg \exists t_k := \max_{j \in \text{ID}}(t_j) < \infty \right) \wedge \left(\bar{T} < \infty \right) \right]. \end{aligned} \quad (2.6)$$

The convention $t_0 = 0$ does not imply loss in generality and it is adopted to simplify the exposition. The convention $k \in \mathbf{N} \cap \text{ID} \Rightarrow (k-1) \in \text{ID}$ means that no natural number is missed inbetween any two consecutive ones in the enumeration of the members of SI and SP. The first part of Lemma 2.1 related to $\bar{t} < \infty$ means that the sampling process stops in finite time so that there is a maximum and last finite sampling instant and a last unbounded sampling period (therefore, the sequence of sampling periods is unbounded with infinite superior limit), and also that the number of sampling instants and periods is finite. The part of Lemma 2.1 for $\bar{t} = \infty$ states that the sampling process never ends so that there are infinitely many sampling instants and periods belonging to their respective numerable sets. Therefore, the cardinal of those sets is denoted by \aleph_0 related to infinite cardinals of numerable sets while the ∞ symbol is usually applied to cardinals of nonnumerable sets of infinitely many elements.

2.1. General Sampling Criterion and a Particular Sampling Criterion of Interest

A general sampling criterion SC is defined as an iterative procedure for some given testing function of the error and/or some of its time derivatives on a next tentative sampling period:

$$\begin{aligned}
 t_{k+1} \in \text{SI} = \text{SI}(\text{SC}) \text{ is generated by the sampling criterion SC; } \forall k \in \mathbf{N}_0 \text{ if} \\
 t_{k+1} := \text{Arg min} \left(\mathbf{R}_{0+} \ni t > t_k : f \in C_T \left(D^i e \times [t_k, t) \subset \mathbf{R}^J \times [t_k, t), \text{ some } i \in J \subseteq \overline{n-1} \cup \{0\}; \mathbf{R} \right) \right. \\
 \left. \text{satisfies SC; } t_k \in \text{SI} \right) \in \text{SI},
 \end{aligned} \tag{2.7}$$

where $t_0 \in \text{SI}$. Therefore, given a set of sampling instants $t_j \in \text{SI}$ for all $j \in \overline{k} \cup \{0\}$, then $t_{k+1} \in \text{SI}$ or $T_k = \infty$ if the sampling criterion ends such that $\text{Card}(\text{SI}) = k + 1 < \infty$. Sampling criteria through testing functions have been obtained in [1–3, 9, 14]. Some of them generate sampling periods in-between consecutive sampling instants as being inversely proportional to the time derivative of the sampled function, or to a combination of consecutive time derivatives, between a maximum admissibility interval (chosen from engineering requirements as such a stability or suited bandwidth). Other types of sampling criteria are chosen through integral criteria over the current sampling period of a quadratic, or some higher even power, of the error time integral between the sampled function and its previous sampled value. A very important one is the so-called criterion of constant difference of amplitudes, firstly proposed in [30], and then generalized formally in [1], and intuitively focused on in [18], to the use of a set of amplitudes which are thresholds of the variation of the sampled signal for each next sampling process. The whole element consisting of the sampling and hold device plus the CADSC (or the more general SDADSC) is, equivalently, modelled with a multiple-hysteresis relay, [1, 15, 18]. This curious nonlinearity in the control law allows an easy interpretation about why sustained oscillations can appear even when the main forward dynamics is linear. If a tracker is being designed, then the use of multiple amplitudes as signal sampling thresholds allows to decrease the amplitude of eventual sustained oscillations and then to improve the tracking servo from a control engineering point of view. Those sampling criteria have the important property, that they are able to generate sustained oscillations of great interest in oscillator design but unsuitable in tracking control problems since a permanent

error between the tracked reference and the governed output signals always exists (see also [16, 17, 23]). In particular, some more general sampling criteria are obtained in [17] which include as particular cases many of those ones existing by that date in the background literature. Some of the results in this paper apply to generic sampling criteria (2.7) irrespective of each particular SC. Other specific results are mainly concerned with a particular sampling criterion, the so-called, SDADSC [1], which is defined implicitly as follows:

$$t_{k+1} = \text{Arg min}(\mathbf{R}_{0+} \ni t > t_k : |e(t) - e(t_k)| = \delta_k \in \mathbf{R}_+, t_k \in \text{SI}) \in \text{SI}, \quad (2.8)$$

for some given sampling set $\text{ST}\delta := \{\delta_k \in \mathbf{R}_+ : 0 < \underline{\delta} \leq \delta_k \leq \bar{\delta} < \infty, \text{ for all } k \in \text{ID}\}$ of amplitude thresholds of the SC. Note from (2.2)–(2.4) that the solution of (2.2) is unique for each set of initial conditions from Picard-Lindelöf theorem for existence and uniqueness of systems of differential equations from continuity and complete induction arguments as follows. Provided that a unique solution exists on $[t_0, t_k)$ for given initial conditions $D^i y(0) = y^{(i)}(0) \in \mathbf{R}$, a continuous and time-differentiable solution also exists and it is unique on $[t_0, t_k]$, and since the input is piecewise constant on $[t_k, t_{k+1})$, it is also continuous and time-differentiable on $[t_k, t_{k+1})$ for all $t_k \in \text{SI}$. Furthermore, the solution is everywhere continuously time-differentiable if $n-m \geq 2$. This follows from the uniqueness of the solutions of ordinary differential equations (ODE) for each given set of initial conditions. The following consistency lemma follows. If $\delta_k = \delta \in \mathbf{R}_+$ for all $k \in \mathbf{N}_0$ in (2.8), then the sampling criterion becomes, in particular, the CADSC [1].

Lemma 2.2. $t_0 \in \text{SI} \Rightarrow t_k \in \text{SI}$ for all $k \in \mathbf{N}$, via the sampling rule (2.7), irrespective of the sampling set of amplitudes $\text{ST}\delta$.

Proof. Proceed by complete induction by assuming that $t_j \in \text{SI}$ for all $j \in \overline{k-1} \cup \{0\}$ so that from (2.8):

$$\begin{aligned} t_{j+1} &= \text{Arg}(t > t_j : |e(t) - e(t_j)| = \delta_j \in \mathbf{R}_+, t_j \in \text{SI}) \in \text{SI}, \quad \forall j \in \overline{k-1} \cup \{0\} \subset \text{ID} \\ &\iff [(t_{j+1} = \text{Arg}(t > t_j : |e(t) - e(t_j)| \\ &= \delta_j \in \mathbf{R}_+, \forall t_j (\leq t_{k-1}) \in \text{SI}) \in \text{SI}]; \quad \forall j \in \overline{k-1} \cup \{0\} \subset \text{ID} \\ &\wedge (t_{k+1} = \text{Arg}(t > t_k : |e(t) - e(t_k)| = \delta_k \in \mathbf{R}_+, t_k \in \text{SI}) \in \text{SI}) \\ &\iff t_{j+1} = \text{Arg}(t > t_j : |e(t) - e(t_j)| = \delta_j \in \mathbf{R}_+, \forall t_j (\leq t_k) \in \text{SI}) \in \text{SI}; \quad \forall j \in \overline{k} \cup \{0\} \subset \text{ID}. \end{aligned} \quad (2.9)$$

□

The sampling criterion (2.8) and its particular version for constant amplitude is of major theoretical interest because the study of the dynamics it generates combines properties of discrete-time systems with some properties of nonlinear systems since, in particular, limit cycles appear in the solution of (2.2)–(2.4) for both the SDADSC, [1, 18]. If $\delta_k = \delta$ is a positive real constant then the sampling criterion is the CADSC, [1, 15–18]. It has been also used in some practical applications, in particular, for tuning PID controllers, [1]. Since the generation of each next sampling period is given by an implicit function in such sampling criteria, the whole control scheme might be considered in the framework of event-driven processes. It is known that for any nonsingular real matrix $\mathbf{T} \in \mathbf{R}^{n \times n}$, a time-differentiable real-state vector

function $x : \mathbf{R}_{0+} \rightarrow \mathbf{R}^n$ satisfying $x(t) = T(y(t), Dy(t), \dots, D^{n-1}y(t))^T$ may be defined so that (2.2)–(2.4) is equivalent to the n th order dynamic feedback system:

$$\dot{x}(t) = Ax(t) + b(r(t_k) - y(t_k)), \quad y(t) = c^T x(t), \quad \forall t \in [t_k, t_{k+1}), \quad (2.10)$$

where $A \in \mathbf{R}^{n \times n}$ and $b, c \in \mathbf{R}^n$ are the matrix of dynamics and the control and output vectors of the continuous-time system, which depend on the coefficients of the polynomials $A(D)$ and $B(D)$ and on the entries to the matrix \mathbf{T} , subject to initial conditions $x(0) = \mathbf{T}(y(0), Dy(0), \dots, D^{n-1}y(0))^T$ at $t = 0$. Equation (2.10) holds for any set of sampling instants SI independent of the particular sampling criterion (2.7). The solution of the first equation in (2.10) within $[t_k, t_{k+1})$ yields directly, again irrespective of SC, the following discrete-time system:

$$\begin{aligned} x(t_{k+1}) &= \Phi(T_k)x(t_k) + \Gamma(T_k)u(t_k) = \Psi(T_k)x(t_k) + \Gamma(T_k)r(t_k), \\ &= \prod_{i=0}^k [\Psi(T_i)]x(0) + \sum_{i=0}^k \prod_{j=i+1}^k [\Psi(T_j)]\Gamma(T_i)r(t_i), \end{aligned} \quad (2.11a)$$

$$y(t_{k+1}) = c^T x(t_{k+1}), \quad (2.11b)$$

where

$$\Psi(T_k) := \Phi(T_k) - \Gamma(T_k)c^T = e^{AT_k} \left(I_n - \left(\int_0^{T_k} e^{A(T_k-\tau)} d\tau \right) bc^T \right) \quad (2.12a)$$

$$\Phi(T_k) := e^{AT_k}, \quad \Gamma(T_k) := \left(\int_0^{T_k} e^{A(T_k-\tau)} d\tau \right) b, \quad (2.12b)$$

where $\Phi(T_k)$ and $\Psi(T_k)$ are the open-loop (i.e., control-free) and closed-loop (i.e., controlled) matrices of dynamics, respectively, and $\Gamma(T_k)$ and c are the control and output vectors, respectively.

2.2. Basic Stability Results

The global BIBO (bounded-input bounded-output) stability of the controlled closed-loop system is discussed provided that the uncontrolled transfer function: $G(s) := c^T(sI_n - A)^{-1}b$ is stable and it possesses a sufficiently small static gain related to the admissible variation domain of the time-varying sampling periods. In the regulation case (i.e., the case of identically zero reference signal $r(t)$), the closed-loop system is globally asymptotically Lyapunov stable. Note that the static gain of $|G(s)|$, $|G(0)| = |c^T A^{-1}b|$, varies linearly with $|b^T c|$ since $\det(A) \neq 0$ if A is a stability matrix. Note that the assumption of smallness of the static gain of the open-loop transfer function is always achievable via incorporation of an amplifier of sufficiently small gain K to the forward loop provided that such a condition is not directly satisfied by the given transfer function so that $|Kc^T b|$ is as sufficiently small as requested. It is proven in the next result that if the maximum allowable time-varying

sampling period T_{\max} increases, then the allowed $|G(0)|$ being compatible with stability decreases correspondingly. If the minimum allowable sampling period T_{\min} increases then such a gain may increase while keeping the stability. The inequality useful for stability in Theorem 2.3 below is $|b^T c| \leq (1 - \varepsilon - e^{-rT_{\min}})r/T_{\max}(1 - e^{-rT_{\max}})$, with $\varepsilon < 1 - e^{-rT_{\min}}$ and $\max(\operatorname{Re} \lambda_i(A) < -r < 0; \lambda_i(A) \in \sigma(A)$ for all $i \in \bar{n}_{\sigma A}$). The stability abscissa of the system matrix is also relevant in the sense that the gain is allowed to increase as such an abscissa increases. The subsequent result is concerned with such considerations.

Theorem 2.3. *Assume that there is an admissibility bounded interval $[T_{\min}, T_{\max}]$ such that $T_k \in [T_{\min}, T_{\max}]$ for all $T_k \in \text{SP}$ for some given sampling criterion SC. Assume also that A is a stability matrix (i.e., $G(s) := c^T(sI_n - A)^{-1}b$ is a stable transfer function). Then, if $|b^T c|$ is sufficiently small according to an explicit trade-off related to the size of $[T_{\min}, T_{\max}]$ and the stability abscissa of the matrix A , then the closed-loop system is BIBO stable. Furthermore, it is globally asymptotically Lyapunov stable in the regulation case without any extra assumptions on the uncontrolled transfer function.*

Proof. Direct calculations with (2.12a) and (2.12b) yield:

$$\begin{aligned} \|\Psi(T_k)\|_2 &\leq \|e^{AT_k}\|_2 \left\| I_n - \left(\int_0^{T_k} e^{-A\tau} d\tau \right) b c^T \right\|_2 \leq e^{-rT_k} + \frac{1 - e^{-rT_k}}{r} T_k |b^T c| \\ &\leq e^{-rT_{\min}} + \frac{1 - e^{-rT_{\max}}}{r} T_{\max} |b^T c| \leq 1 - \varepsilon < 1 \end{aligned} \quad (2.13)$$

provided that $|b^T c| \leq (1 - \varepsilon - e^{-rT_{\min}})r/T_{\max}(1 - e^{-rT_{\max}})$, for some prefixed real constant $\varepsilon \in (0, 1)$ satisfying $0 < \varepsilon < 1 - e^{-rT_{\min}}$, $\max(\operatorname{Re} \lambda_i(A) < -r < 0; \lambda_i(A) \in \sigma(A)$ for all $i \in \bar{n}_{\sigma A}$) (for some $1 \leq n_{\sigma A} \leq n$ being the number of distinct eigenvalues of A), and $T_k \in [T_{\min}, T_{\max}]$ for all $T_k \in \text{SP}$ and any given sampling criterion SC. Proceeding recursively and taking ℓ_2 -vector and matrix norms in (2.11a) and (2.11b), one gets since

$$\begin{aligned} \|\Psi(T_k)\|_2 &\leq \rho := 1 - \varepsilon < 1, \\ \|\Gamma(T_k)\|_2 &\leq \frac{(1 - e^{-rT_k})}{r} |b^T c| \leq \frac{(1 - \varepsilon - e^{-rT_{\min}})}{T_{\max}}, \end{aligned} \quad (2.14)$$

$$\begin{aligned} \|x(t_{k+1})\|_2 &\leq \left\| \prod_{i=0}^k [\Psi(T_i)] \right\|_2 \|x(0)\|_2 + \sum_{i=0}^k \left\| \prod_{j=i+1}^k [\Psi(T_j)] \right\|_2 \|\Gamma(T_i)\|_2 |r(t_i)| \\ &\leq \varepsilon^k \|x(0)\|_2 + \frac{1 - \varepsilon - e^{-rT_{\min}}}{T_{\max}} \left(\sum_{i=0}^k \varepsilon^{k-i} \right) \max_{i \in \bar{k} \cup \{0\}} (|r(t_i)|) \\ &\leq \varepsilon^k \|x(0)\|_2 + \frac{1 - \varepsilon - e^{-rT_{\min}}}{(1 - \varepsilon)T_{\max}} \max_{i \in \mathbb{N}_0} (|r(t_i)|) \\ &\leq K_x(\|x(0)\|_2) + K_{r0}K_r \\ &< \infty, \quad \forall t_{k+1} \in \text{SI}, \end{aligned} \quad (2.15)$$

where $K_{r0} := (1 - \varepsilon - e^{-rT_{\min}})/(1 - \varepsilon)T_{\max}$, $K_r := \max_{i \in \mathbb{N}_0} (\|r(t_i)\|_2)$, and then

$$0 \leq \limsup_{k \rightarrow \infty} \|x(t_{k+1})\|_2 \leq K_{r0} \max_{i \in \mathbb{N}_0} (\|r(t_i)\|_2) \leq K_{r0} K_r < \infty, \quad (2.16)$$

since $\varepsilon^k \rightarrow 0$ as $k \rightarrow \infty$. Equations (2.11a) and (2.11b)-(2.12a) and (2.12b) are replaced within the inter-sample time intervals with:

$$x(t_k + \tau) = \Psi(\tau)x(t_k) + \Gamma(\tau)r(t_k), \quad \forall \tau \in [0, T_k), \quad (2.17)$$

$$\begin{aligned} \Psi(\tau) &:= \Phi(\tau) - \Gamma(\tau)c^T = e^{A\tau} \left(I_n - \left(\int_0^\tau e^{A(\tau-\tau')} d\tau' \right) b c^T \right), \quad \forall \tau \in [0, T_k), \\ \Phi(\tau) &:= e^{A\tau}, \quad \Gamma(\tau) := \left(\int_0^\tau e^{A(\tau-\tau')} d\tau' \right) b, \quad \forall \tau \in [0, T_k), \forall T_k \in \text{SP}, \end{aligned} \quad (2.18)$$

so that, one gets from (2.17)-(2.18) by using (2.15)-(2.16):

$$\begin{aligned} \|x(t_k + \tau)\|_2 &\leq \|\Psi(\tau)\|_2 \|x(t_k)\|_2 + \|\Gamma(\tau)\|_2 |r(t_k)| \\ &\leq K' \varepsilon^k (\|x(0)\|_2) + \left(K_{r0} + r^{-1} |b^T c| \right) K_r < \infty, \quad \forall \tau \in [0, T_k), \forall T_k \in \text{SP}, \end{aligned} \quad (2.19)$$

$$0 \leq \limsup_{k \rightarrow \infty} \|x(t_k + \tau)\|_2 \leq \left(K_{r0} + r^{-1} |b^T c| \right) K_r < \infty, \quad \forall \tau \in [0, T_k), \forall T_k \in \text{SP}. \quad (2.20)$$

Thus, the dynamic system (2.11a) and (2.11b)-(2.12a) and (2.12b) is bounded-input-bounded-output (BIBO) stable for any uniformly bounded reference $r(t)$. If the reference is identically zero (regulation), then $K_r = 0$ so that, one gets from (2.16)–(2.20), that $\lim_{k \rightarrow \infty} x(t_k + \tau) = 0$ for all $\tau \in [0, T_k)$ for all $T_k \in \text{SP}$, irrespective of the initial conditions so that the dynamic system is globally asymptotically Lyapunov stable. The proof is complete. \square

It is obvious that $\Psi(T_k)$ is a convergent matrix (i.e., a stability matrix in the discrete sense then all its eigenvalues have modulus less than unity) under the conditions of Theorem 2.3. Note that, otherwise, the state at sampling instants would be at least critically stable and would diverge for certain bounded inputs which could be fixed by construction. Since the system is globally Lyapunov stable then it also exhibits ultimate boundedness in the usual Lyapunov sense as direct conclusion from Theorem 2.3. A more general ultimate boundedness results is now derived without invoking a sufficiently small static gain of the uncontrolled system for a string of consecutive products of the matrix $\Psi(T_k)$ being convergent.

Theorem 2.4. *The following properties hold:*

- (i) assume that for a given SC and each $k \in \text{ID}$, there exists $1 \leq i = i(k) \leq \bar{i} < \infty$ such that $\|\bar{\Psi}(t_k, t_{k+i})\|_2 \leq \varepsilon_1 < 1$, where $\bar{\Psi}(t_k, t_{k+i}) := \prod_{j=k}^{k+i-1} [\Psi(T_j)]$. Thus, the system is BIBO-stable for any bounded initial state and possesses the ultimate boundedness property for any bounded reference sequence $r(t_k)$ for all $t_k \in \text{SI}$;
- (ii) if $\|\bar{\Psi}(t_k, t_{k+\ell})\|_2 \leq K(k, \ell) \leq K_{\bar{\Psi}} < \infty$ (which holds in particular if $\|\Psi(T_k)\|_2 \leq 1$ for all $t_k \in \text{SP}$) and $\|\sum_{i=0}^k \prod_{j=i+1}^k [\Psi(T_j)] \Gamma(T_i) r(t_i)\|_2 < \infty$ for all $k \in \mathbf{N}$, then the system is BIBO-stable for any bounded initial state.

Proof. For any finite $k \in \text{ID}$, define the nonnegative scalar function $v(\|x(t_k)\|_2) := x^T(t_k)x(t_k)$. Then, for some finite $1 \leq i = i(k) \leq \bar{i}$:

$$\begin{aligned}
 v(\|x(t_{k+i})\|_2) - v(\|x(t_k)\|_2) &= x^T(t_{k+i})x(t_{k+i}) - x^T(t_k)x(t_k) \\
 &\leq \left[\sum_{\ell=i}^{k+i-1} \left(\left\| \prod_{j=k+1}^{k+i-1} [\Psi(T_j)] \right\|_2 \|\Gamma(T_\ell)\|_2 |r(T_\ell)| \right) \right]^2 \\
 &\quad + \left(2 \left\| \prod_{j=k}^{k+i-1} [\Psi(T_j)] \right\|_2 \sum_{\ell=i}^{k+i-1} \left(\left\| \prod_{j=k+1}^{k+i-1} [\Psi(T_j)] \right\|_2 \|\Gamma(T_\ell)\|_2 |r(T_\ell)| \right) \right. \\
 &\quad \left. - (1 - \varepsilon_1^2) \left\| \prod_{j=k}^{k+i-1} [\Psi(T_j)] \right\|_2^2 \|x(t_k)\|_2 \right) \\
 &\quad \times \|x(t_k)\|_2 v(\|x(t_{k+i})\|_2) - v(\|x(t_k)\|_2) \\
 &= x^T(t_{k+i})x(t_{k+i}) - x^T(t_k)x(t_k) \\
 &\leq K_{kr}^2(t_k, t_{k+i}) + (2 K_{k\Psi}(t_k, t_{k+i}) K_{kr}(t_k, t_{k+i}) \\
 &\quad - (1 - \varepsilon_1^2) K_{k\Psi}^2(t_k, t_{k+i}) \|x(t_k)\|_2) \|x(t_k)\|_2 \\
 &\leq \bar{K}_{kr}^2 + (2 \bar{K}_{k\Psi} \bar{K}_{kr} - (1 - \varepsilon_1^2) K_{k\Psi}^2(t_k, t_{k+i}) \|x(t_k)\|_2) \|x(t_k)\|_2,
 \end{aligned} \tag{2.21}$$

where

$$\begin{aligned}
 0 \leq K_{k\Psi}(t_k, t_{k+i}) &:= \left\| \prod_{j=k}^{k+i-1} [\Psi(T_j)] \right\|_2 \leq \bar{K}_{\Psi} < \infty, \\
 0 \leq K_{rk}(t_k, t_{k+i}) &:= \sum_{\ell=i}^{k+i-1} \left(\left\| \prod_{j=k+1}^{k+i-1} [\Psi(T_j)] \right\|_2 \|\Gamma(T_\ell)\|_2 |r(T_\ell)| \right) \leq \bar{K}_r < \infty,
 \end{aligned} \tag{2.22}$$

with $\bar{K}_{\Psi} := \max_{k \leq j \leq k+\bar{i}} (K_{k\Psi}(t_k, t_{k+i}) : k \in \text{ID})$ and $\bar{K}_r := \max_{k \leq j \leq k+\bar{i}} (K_{rk}(t_k, t_{k+i}) : k \in \text{ID})$ being independent of k . Now, proceed by contradiction by assuming that $\{\|x(t_\ell)\|_2\}_{\ell \in \text{ID}}$ is unbounded. Thus, there exists a subsequence $\{\|x(t_k)\|_2\}_{k \in \text{ID}_0 \subseteq \text{ID}}$ which diverges for some numerable subset ID_0 of ID so that $v(\|x(t_{k+i})\|_2) > v(\|x(t_k)\|_2)$ and $\lim_{k \rightarrow \infty} v(\|x(t_k)\|_2) = \infty$.

Choose $k \in \text{ID0}$ being arbitrarily large but finite, $k + i(k) \in \text{ID0}$ so that $\|x(t_k)\|_2 \geq \max((2\bar{K}_{k\Psi}\bar{K}_{kr} + \varepsilon_2)/(1 - \varepsilon_1^2)K_{k\Psi}^2(t_k, t_{k+i}), M_k)$ with M_k being an arbitrarily large positive real number depending on k and $\varepsilon_2 \in \mathbf{R}_+$. This is always possible since $\{\|x(t_k)\|_2\}_{k \in \text{ID0}}$ diverges and $0 \leq \varepsilon_1 < 1$. From (2.21), one gets:

$$\begin{aligned} 0 &< v(\|x(t_{k+i})\|_2) - v(\|x(t_k)\|_2) \\ &\leq \bar{K}_{kr}^2 + \left(2\bar{K}_{k\Psi}\bar{K}_{kr} - (1 - \varepsilon_1^2)K_{k\Psi}^2(t_k, t_{k+i})\|x(t_k)\|_2\right)\|x(t_k)\|_2 \\ &\leq \bar{K}_{kr}^2 - \varepsilon_3 M_k \leq 0, \end{aligned} \quad (2.23)$$

if $M_k \geq \bar{K}_{kr}^2/\varepsilon_3$ which leads to a contradiction. As a result, $\{\|x(t_k)\|_2\}_{k \in \text{ID}}$ is bounded from above if there exists $1 \leq i = i(k) \leq \bar{i} < \infty$ such that $\|\bar{\Psi}(t_k, t_{k+i})\|_2 \leq \varepsilon_1 < 1$, and the reference sequence is uniformly bounded. Using a similar technique as in Theorem 2.3, it may be proved that the state is bounded within any intersample period on $[t_k, \infty)$, for all $k \in \text{ID}$ and finite and any sampling criterion SC (i.e., for any $t_k \in \text{SI}$ generated from any SC). Therefore, the system possesses the property of ultimate boundedness. On the other hand, the recursive equations (2.15) lead to a bounded solution sequence on any interval $[0, t_k)$ of finite measure if the reference sequence is uniformly bounded since all the matrices of parameters of the discrete-time dynamic system are bounded for finite time irrespective of further considerations. Therefore, ultimate boundedness implies BIBO stability and global stability of the unforced discrete-time system. The proof of Property (i) is complete. Property (ii) follows directly by taking upper bounds via the use of norms in (2.12b). \square

The following result parallel to Theorem 2.4 is concerned with instability:

Theorem 2.5. *Assume that for a given SC and each $k \in \text{ID}$, there exists $\infty > i = i(k) \geq j \in \mathbf{N}$ such that $\|\bar{\Psi}(t_k, t_{k+i})\|_2 \geq \varepsilon_1 > 1$. Thus, the discrete-time (2.17)-(2.18) system is unstable.*

Proof. Take the set of sampling instants $t_j \in \text{SI}$ for the given SC and zero reference input. Now, take initial conditions $x(t_k)$ at a finite $t_k \in \text{SI}$ which are a nonzero eigenvector of $\bar{\Psi}(t_k, t_{k+j\ell})$ so that $\lim_{\ell \rightarrow \infty} \inf \|\bar{\Psi}(t_k, t_{k+j\ell})x(t_k)\|_2 \geq \lim_{\ell \rightarrow \infty} \inf \varepsilon_1^\ell \|x(t_k)\|_2 = \infty$. Then, the system is unstable. \square

Note that the stability condition in terms of the modulus of eigenvalues being less than unity is equivalent in terms of positive definiteness of the matrix of dynamics to:

$$\varepsilon_0^2 I_n \leq \bar{\Psi}^T(t_k, t_{k+i})\bar{\Psi}(t_k, t_{k+i}) \leq \varepsilon_1^2 I_n, \quad (2.24)$$

where $0 \leq \varepsilon_0 \leq \varepsilon_1 < 1$, which could be alternative used in both the statement and proof of Theorem 2.4.

3. Oscillations and Periodic Oscillations

Concerning the discrete-time system (2.17)-(2.18), whose expression at sampling instants are (2.11a) and (2.11b)-(2.12a) and (2.12b) for any sampling criterion SC, the following definitions for weak and strong oscillatory solutions will apply.

Definition 3.1. The discrete-time system (2.11a) and (2.11b)-(2.12a) and (2.12b) has a weak oscillatory output solution for a given sampling criterion SC and some initial condition $x(0) \in \mathbf{R}^n$ if for any given $t \in \mathbf{R}_{0+}$, such that $y(t) \neq 0$, there exist finite real numbers $\alpha(t) \geq \varepsilon_\alpha$ and $\beta(t) \geq \varepsilon_\beta$, being in general dependent on t , for some $\varepsilon_\alpha, \varepsilon_\beta \in \mathbf{R}_+$, such that $\text{sign}(\delta_y(t, t + \alpha(t))\delta_y(t_k, t + \alpha(t) + \beta(t))) \leq 0$, where $\delta_y(t, t') := y(t') - y(t)$.

Definition 3.2. The discrete-time system (2.11a) and (2.11b)-(2.12a) and (2.12b) has a strong oscillatory output solution for some initial condition $x(0) \in \mathbf{R}^n$ if for any given $t \in \mathbf{R}_{0+}$, such that $y(t) \neq 0$, $\text{sign}(\delta_y(t, t + \alpha(t))\delta_y(t_k, t + \alpha(t) + \beta(t))) < 0$ and $y(t + \alpha(t))$ and $y(t + \alpha(t) + \beta(t))$ are not both zero.

Definition 3.3. The discrete-time system (2.17)-(2.18) has a periodic weak oscillatory output solution of oscillation period $T_{\text{per}} \in \mathbf{R}_+$, for some initial condition $x(0) \in \mathbf{R}^n$, if $\text{sign}(\delta_y(t, t + T_{\text{per}}/2)\delta_y(t, t + T_{\text{per}})) \leq 0$, $y(t, t + T_{\text{per}}) = y(t)$, for all $t \in \mathbf{R}_+$.

Definition 3.4. The discrete-time system (2.17)-(2.18) has a periodic strong oscillatory output solution of oscillation period T_{per} for some initial condition $x(0) \in \mathbf{R}^n$ if it has a periodic weak oscillatory output for such a period and, furthermore, $\text{sign}(\delta_y(t, t + T_{\text{per}}/2)\delta_y(t, t + T_{\text{per}})) < 0$ if $y(t, t + T_{\text{per}}) = y(t) = 0$, for all $t \in \mathbf{R}_+$.

Note that a solution may be oscillatory (Definitions 3.1-3.2) without being periodic (Definitions 3.3-3.4) when there are changes in the sign of the incremental output along intervals of finite duration. A weak oscillation compared to a strong oscillation allows positive or negative increments of the output at finite intervals always of the same sign. The above Definitions 3.1-3.4 might also be referred to, in general, to nonsymmetric oscillations related to their deviations from zero. Note that trivial solutions, that is, those being identically zero are not periodic solutions according to the given definitions. Note also that periodic solutions can possess an oscillation period which is not the sum of any fixed set of consecutive sampling periods even for such a set obeying a rule implying some repetitive sequence of periods. It turns out that the concepts of oscillation and periodic oscillation may be extended to any of the components of the state vector. The next result establishes clear implications among Definitions 3.1-3.4.

Theorem 3.5. *If an output solution is strongly oscillatory, then it is also weakly oscillatory.*

If an output solution is strongly periodic oscillatory then it is also weakly periodic oscillatory.

If an output solution is weakly (strongly) periodic oscillatory, then it is also weakly (strongly) oscillatory.

Note that oscillations are not always detectable for any given sampling criterion at arbitrary sampling instants since hidden oscillations can exist which cannot be detected at sampling instants. However, sufficient conditions for existence of oscillations can be formulated at sampling instants as stated in the subsequent results, whose proofs are direct conclusions of Definitions 3.1-3.2.

Theorem 3.6. *The discrete-time system (2.11a) and (2.11b)-(2.12a) and (2.12b) exhibits a weak oscillatory output at sampling instants for a given sampling criterion SC and some initial condition $x(0) \in \mathbf{R}^n$ if for any $t_k \in \text{SI}$, such that $y(t_k) \neq 0$, there exist finite natural numbers $k_1(k)$ and $k_2(k)$, being in general dependent on $k \in \text{ID}$, such that $\text{sign}(\delta_y(t_k, t_{k+k_1(k)})\delta_y(t_k, t_{k+k_1(k)+k_2(k)})) \leq 0$, where $\delta_y(t_k, t_j) := y(t_j) - y(t_k)$.*

Theorem 3.7. *The discrete-time system (2.11a) and (2.11b)-(2.12a) and (2.12b) has a strong oscillatory output at sampling instants for some initial condition $x(0) \in \mathbf{R}^n$ if $\text{sign}(\delta_y(t_k, t_{k+k_1(k)})\delta_y(t_k, t_{k+k_1(k)+k_2(k)})) < 0$ but $y(t_{k+k_1(k)})$ and $y(t_{k+k_1(k)+k_2(k)})$ are not both zero.*

Remark 3.8. The existence of weak and strong oscillations under the sufficient conditions of Theorems 3.6 and 3.7, respectively, may be investigated explicitly by the use of the state evolution over a finite number of consecutive sampling instants through (2.11b) together with the output expression at sampling instants in the second formula of (2.11a).

Note that the detection of periodic oscillations involving sampling instants only is not feasible even in terms of sufficient-type conditions since the period of such oscillations is not necessarily the exact sum of a consecutive number of limit sampling periods. See, for instance, [1, 15, 16], for SDADSC and CADSC, respectively. The following result states that stable uncontrolled systems which are closed-loop stable under unity feedback, fulfil the conditions of Theorem 2.4 and which do not have stable equilibrium points exhibit oscillatory responses.

Theorem 3.9. *Assume that the closed-loop discrete-time system has no stable equilibrium point, while the uncontrolled system is stable under the conditions of Theorem 2.4, and the sampling criterion also fulfils the conditions of Theorem 2.4. Then, any solution of the discrete-time closed-loop system is at least weakly oscillatory and bounded.*

Proof. Since any state solution is bounded for bounded initial conditions and do not converge to a constant equilibrium point, it follows that all the state components verify the incremental changes of sign of Definition 4.1, since no one can either converge to a constant or to be unbounded. \square

A direct related result which follows from Theorem 2.5 is now stated by simple inspection without a formal proof.

Theorem 3.10. *Assume that the closed-loop discrete-time system has no stable equilibrium point while the discrete-time system is unstable under the conditions of Theorem 2.5 for some sampling criterion SC. Then, no solution of the discrete-time closed-loop system can be bounded, while it can be weakly oscillatory and unbounded.*

Limit cycles are asymptotic isolated limit periodic oscillations in certain nonlinear systems which are usually independent of the initial conditions (as, e.g., the well-known Van der Pol equation). Since some nonlinear systems can also posses oscillations which depend on initial conditions, as for instance, the also well-known Duffing equation modelling certain nonlinear strings with combined linear and cubic effects, no difference is made at the moment between both situations. More precisely, a limit cycle on a plane or a n th-dimensional manifold is a closed trajectory having the property that at least one another trajectory spirals into either as time tends to infinity (stable limit cycles or a self-sustained periodic oscillations) or as time tends to minus infinity (unstable limit cycles). It turns out that a limit cycle exist in the dynamic system of Section 2 only if for a given sampling criterion SC:

$$\lim_{k \rightarrow \infty} D^i y(t_k + \tau) = D^i y^*(\tau), \quad \forall i \in \overline{n-1} \cup \{0\}, \quad (3.1)$$

for some periodic function $y^* \in \text{PC}^{(n-1)}([0, T_{\text{per}}]; \mathbf{R})$ of period $T_{\text{per}} > 0$ such that:

- (1) all its time derivatives until order $(n - 1)$ exist and are almost everywhere continuous except at the sampling instants;
- (2) $D^n y^*(t)$ exists everywhere on its definition domain, but it is not required to be continuous in-between sampling instants, so that it is not required for the limit cycle to satisfy $y^* \in PC^{(n)}([0, T_{\text{per}}]; \mathbf{R})$;
- (3) $D^n y^*(0^+) = D^{(n)} y^*(T_{\text{per}}^+)$ for all $\tau \in [0, T_{\text{per}})$ such that $t_k + \tau \notin \text{SI}$ and for $(t_k + \tau)^+ = t_{k+i}^+$ if $t_{k+i} \in \text{SI}$ for some $i \in \mathbf{N}$.

4. Limit Oscillations under Sampling Criteria

Note from (2.11a) and (2.11b)-(2.12a) and (2.12b) and (2.17)-(2.18) that for any SC:

$$\begin{aligned}
 x(t_{k+\ell} + \tau) &= \Psi(\tau)(\Psi(T_k)x(t_k) + \Gamma(T_k)r(t_k)) + \Gamma(\tau)r(t_{k+\ell}) \\
 &= \Psi(\tau) \left(\prod_{i=k}^{k+\ell-1} [\Psi(T_i)]x(t_k) + \sum_{i=k}^{k+\ell-1} \prod_{j=i+1}^{k+\ell-1} [\Psi(T_j)]\Gamma(T_i)r(t_i) \right) \\
 &\quad + \Gamma(\tau)r(t_{k+\ell})y(t_{k+\ell} + \tau) = c^T x(t_k + \tau) \\
 &= c^T \Psi(\tau) \left(\prod_{i=k}^{k+\ell-1} [\Psi(T_i)]x(t_k) + \sum_{i=k}^{k+\ell-1} \prod_{j=i+1}^{k+\ell-1} [\Psi(T_j)]\Gamma(T_i)r(t_i) \right) + \Gamma(\tau)r(t_{k+\ell}),
 \end{aligned} \tag{4.1}$$

for all $t_k \in \text{SI}$, for all $\tau \in [0, T_{k+\ell})$, for all $T_k \in \text{SP}$, and for all $k \in \text{ID}$, for all $\ell \in \mathbf{N}$, subject to the parameterizations (2.18) becoming (2.12a) and (2.12b) at sampling instants. If a limit oscillation exists then, one gets for (4.1):

$$\begin{aligned}
 &\exists \lim_{\text{SI} \ni t_{k+\ell} \rightarrow \infty} x(t_{k+\ell} + \tau) \\
 &= \lim_{\text{SP} \ni T_{k+i}, \text{ID} \ni k \rightarrow \infty} x^* \left(\sum_{i=k}^{k+\ell-1} T_{k+i} + \tau \right) \\
 &= \lim_{\text{SI} \ni t_k \rightarrow \infty, T_i \in \text{SP}} \left(\Psi(\tau) \left(\prod_{i=k}^{k+\ell-1} [\Psi(T_i)]x(t_k) + \sum_{i=k}^{k+\ell-1} \prod_{j=i+1}^{k+\ell-1} [\Psi(T_j)]\Gamma(T_i)r(t_i) \right) + \Gamma(\tau)r(t_{k+\ell}) \right) \\
 &= \lim_{\text{SP} \ni T_{k+i}, \text{ID} \ni k \rightarrow \infty} \left(\Psi(\tau) \left(\prod_{i=k}^{k+\ell-1} [\Psi(T_i)]x^*(t_k) + \sum_{i=k}^{k+\ell-1} \prod_{j=i+1}^{k+\ell-1} [\Psi(T_j)]\Gamma(T_i)r(t_i) \right) + \Gamma(\tau)r(t_{k+\ell}) \right),
 \end{aligned} \tag{4.2}$$

for all $\ell \in \bar{p}$ and some finite $p \in \mathbf{N}_0$, for some $\tau \in \mathbf{R}_{0+}$. Thus,

$$\begin{aligned} x^* \left(\bar{t}_i + \sum_{i=1}^p \bar{T}_i + \tau \right) &= x^* (\bar{t}_i) \\ &= \left(\Psi(\tau) \left(\prod_{j=i}^{i+p-1} [\Psi(\bar{T}_j)] \right) x^* (\bar{t}_i) \right. \\ &\quad \left. + \sum_{j=1}^{i+p-1} \prod_{\ell=j+1}^{i+p-1} [\Psi(\bar{T}_\ell)] \Gamma(\bar{T}_j) r(\bar{t}_j) \right) \Gamma(\tau) r(\bar{t}_{i+p}), \end{aligned} \quad (4.3)$$

where the following limits have to exist for $T_k \in \text{SP}$, $t_k \in \text{SI}$, $k \in \text{ID}$ for all $i \in \bar{p}$, that is, for the sampling periods and sampling instants with a certain repeated string sequence, where \bar{t}_0 in arbitrary starting limit reference sampling instant:

$$\begin{aligned} \exists \lim_{\text{ID} \ni k \rightarrow \infty} T_{k+i} &:= \bar{T}_i = \bar{T}_{p+i}, \\ \exists \lim_{\text{ID} \ni k \rightarrow \infty} t_{k+i} &:= \bar{t}_i = \bar{t}_{p+i}, \\ \exists \lim_{\text{ID} \ni k \rightarrow \infty} r(t_{k+i}) &:= r(\bar{t}_i) = r(\bar{t}_{p+i}). \end{aligned} \quad (4.4)$$

Then, the period of the limit oscillation is $T_{\text{per}} := \sum_{i=1}^p \bar{T}_i + \tau$, some real $\tau \in [0, \bar{T}_1)$. A similar limiting equation using (4.2) into the output equation: $y(t_{k+\ell} + \tau) = c^T x(t_k + \tau)$ describes the limit oscillation in the output as $t_k \rightarrow \infty$. The following four lemmas related to the necessary condition of the existence of a limit cycle independent of a particular SC follow from (4.3)-(4.4) and simple topological considerations about uniqueness of the state- trajectory solution:

Lemma 4.1. *If (4.3), subject to (4.4), holds, then $\exists \lim_{t \rightarrow \infty} x(t + \tau) = x^*(\tau) = x^*(\tau + T_{\text{per}})$; for all $\tau \in [0, T_{\text{per}})$ and then a limit oscillation of the state-trajectory solution exists.*

Proof. Note that $\lim_{k \rightarrow \infty} x(t_{k+i}) = x^*(\bar{t}_i) = x^*(\bar{t}_i + T_{\text{per}})$ and $\lim_{k \rightarrow \infty} x(t_k + \tau) = x^*(\tau) = x^*(\tau + T_{\text{per}})$ for $T_{\text{per}} := \sum_{i=1}^p \bar{T}_i + \tau$, for some parameterizing $\tau \in [0, \bar{T}_1)$, for all $i \in \bar{p}$ that is, at a discrete set of $(p+1)$ limit sampling instants as time tends to infinity, some $p \in \mathbf{N}_0$ and this sequence of identities is repeated with period T_{per} . The state-trajectory inbetween consecutive samples is prescribed according to the values of the limit reference and the state trajectory components cannot intersect at any time so that the periodic limit identity holds in continuous-time as time tends to infinity, and the result is proven. \square

Lemma 4.2. *Assume that distinct double points $x^*(\bar{t}_i)$ ($i \in \bar{p}$) exist satisfying (4.3), subject to (4.4) for some $p \in \mathbf{N}_0$, or equivalently,*

$$\left(I_n - \Psi(\tau) \left(\prod_{j=i}^{i+p-1} [\Psi(\bar{T}_j)] \right) \right) x^*(\bar{t}_i) = \sum_{j=1}^{i+p-1} \prod_{\ell=j+1}^{i+p-1} [\Psi(\bar{T}_\ell)] \Gamma(\bar{T}_j) r(\bar{t}_j), \quad (4.5)$$

for all $i \in \bar{p}$ and some $\tau \in [0, T_1)$. Assume also that if $p = 1$ then $x^*(\bar{t}_i)$ satisfying the above identity is not an equilibrium point. Then, the existing limit oscillation may be tested by any of the double points, in particular, by the limit double point $x^*(\bar{t}_1)$ satisfying:

$$\left(I_n - \Psi(\tau) \left(\prod_{j=1}^p [\Psi(\bar{T}_j)] \right) \right) x^*(\bar{t}_1) = \sum_{j=1}^p \prod_{\ell=j+1}^p [\Psi(\bar{T}_\ell)] \Gamma(\bar{T}_j) r(\bar{t}_j). \quad (4.6)$$

If the reference sequence is identically zero, then a limit oscillation exists verifying double points

$$\begin{aligned} x \in \text{Ker} \frac{(I_n - \Psi(\tau) (\prod_{j=1}^p [\Psi(\bar{T}_j)]))}{\{P_{\text{eq}}\}} \\ \text{if } \text{Ker} \frac{(I_n - \Psi(\tau) (\prod_{j=1}^p [\Psi(\bar{T}_j)]))}{\{P_{\text{eq}}\}} \neq \{0\}. \end{aligned} \quad (4.7)$$

Proof. It follows from (4.3)-(4.4) and Lemma 4.1 since the state-trajectory solution is unique for any initial conditions, sampling periods and reference sequence and a periodic limit oscillation exist. Since the limit double points are distinct, they are not equilibrium points since the state-trajectory solution is unique if $p > 1$. If $p = 1$, the double point is not an equilibrium one as a requirement of the lemma statement. \square

Lemma 4.3. *If Lemmas 4.1-4.2 hold for a given set of $p \in \mathbf{N}_0$ limit sampling periods \bar{T}_i for all $i \in \bar{p}$ and some real $\tau \in [0, T_1)$, then there is no other limit oscillation for the same sets of limit sampling periods and limit reference sequence neighboring the one with oscillation period $T_{\text{per}} := \sum_{i=1}^p \bar{T}_i + \tau$.*

Proof. If $\tau \rightarrow \tau + \Delta\tau$ then $T_{\text{per}} \rightarrow T_{\text{per}} + \Delta\tau$ provided identical limit sampling periods \bar{T}_i for all $i \in \bar{p}$. Since all state-trajectories are distinct, any two closed trajectories cannot be everywhere identical. Thus, two trajectories with identical initial conditions should bifurcate to different subtrajectories to complete both distinct closed paths at points inside the common parts of both trajectories. This contradicts the fact that state-trajectory solutions are unique. \square

Lemma 4.4. *All the closed state trajectory solutions verifying Lemmas 4.1-4.3 are either stable or any unstable one, if any, is surrounded by two stable ones, namely, point-wise strictly bounded from above and below by two distinct stable closed state-trajectory solutions. Furthermore, any two closed stable trajectories cannot be arbitrarily close to each other.*

Proof. The two matrices $\Psi(\tau)(\prod_{j=1}^p [\Psi(\bar{T}_j)])$ have to possess at least two complex conjugate eigenvalues at the unit circumference for both tuples $(p_\ell, \tau_\ell, \bar{T}_{j\ell}; j \in \bar{p}_1); \ell = 1, 2$ associated with the limit closed state-trajectory solutions. Otherwise, the system would be either BIBO stable or unstable from Theorems 2.4-2.5. Then, since all the eigenvalues are within the closed unity circle, the system is BIBO stable from Theorem 2.4 so that any state-trajectory solution can be unbounded. Thus, all existing limit oscillations are bounded for all time and then either stable or surrounded by two stable ones. On the other hand, if any two stable trajectories are arbitrary and close to each other then it would be destroyed by any arbitrarily small disturbance so they would not be stable. \square

Lemma 4.4 dictates that potential limit cycles of the solutions are separated to each other so that there is no accumulation closed attractor of the state-space trajectories. The interpretation of the implications of Lemma 4.4 for a linear dynamics of dimension $n = 2$ is direct. For $n > 2$, it is possible to interpret the lemma consequences in a plane corresponding to a 2nd-dimensional system for two of the state components in the same above way while, for the remaining components, we can consider the surrounding trajectories being equal to that one under consideration. The whole surrounding closed trajectories are still distinct from the study for the second-order subsystem.

5. Limit Oscillations for the Constant and Sampling-Dependent Amplitude Difference Sampling Criteria

5.1. The CADSC

The presence of limit oscillations is now discussed for the CADSC from the study of oscillations for sampling criteria in the above section. The conditions of stable limit oscillation for the CADSC are from (4.4)–(4.6) for a double point $x^*(\bar{t}_1)$ to exist satisfying:

$$\left(I_n - \Psi(\tau) \left(\prod_{j=1}^p [\Psi(\bar{T}_j)] \right) \right) x^*(\bar{t}_1) = \sum_{j=1}^p \prod_{\ell=j+1}^p [\Psi(\bar{T}_\ell)] \Gamma(\bar{T}_j) r(\bar{t}_j), \quad (5.1)$$

($i \in \bar{p}$) for some $p \in \mathbf{N}_0$ and some real $\tau \in [0, \bar{T}_1)$; and, furthermore,

$$\begin{aligned} & [y^*(\bar{t}_1), y^*(\bar{t}_2), \dots, y^*(\bar{t}_p), y^*(\bar{t}_p + \tau)] \\ &= c^T \left[I_n, \Psi(\bar{T}_1), \dots, \prod_{j=1}^p [\Psi(\bar{T}_j)], \Psi(\tau) \left(\prod_{j=1}^p [\Psi(\bar{T}_j)] \right) \right] x^*(\bar{t}_1) \\ &= [\ell\delta, (\ell+1)\delta, \dots, (\ell+p_1-1)\delta, (\ell+p_1-2)\delta, \dots \\ & \quad (\ell+p_1-p_2-1)\delta, (\ell+p_1-p_2)\delta, \dots, (\ell+p-1)\delta, (\ell+p-1)\delta + \tau] \end{aligned} \quad (5.2)$$

for some $p_1 = p_1(\ell)$, $p_2 = p_2(\ell) \leq p \in \mathbf{N}_0$ and some finite $\ell \in \mathbf{N}_0$. Since the limit oscillation may be tested starting at any sampling point, it turns out that any limit oscillation verifying (5.2) for some $\ell \in \mathbf{N}$ is also verified $1 \leq j \leq \ell$ by redefining the integers $p_1(j)$; $p_2(j) \leq p$. A brief intuitive explanation of (5.2) follows. Take any positive value of the limit oscillation for a certain $\ell \in \mathbf{N}$ such that (5.1)–(5.2) hold together for some $p_1(\ell)$, $p_2(\ell) \leq p \in \mathbf{N}_0$ for some $p \in \mathbf{N}_0$ and some real $\tau \in [0, \bar{T}_1)$. Then, a limit oscillation with the system output satisfying (5.2) starting with no loss in generality by a positive value, continuing to increase p_1 times by step-by-step positive increments δ at each sampling instant, then decreasing p_2 times, then increasing again to complete the closed trajectory. Note that for any $\ell \in \mathbf{N}_0$ satisfying the given constraints, $p = (p - p_2(\ell)) + p_2(\ell)$ and $|p - 2p_2(\ell)| \leq 2$ since the number of negative increments is some $p_2(\ell) \leq p \in \mathbf{N}_0$, and the number of positive increments is then $p - 2p_2(\ell)$, while the absolute difference of amplitude increments is of at most two.

5.2. The SDADSC

The study of oscillations can be directly generalized to the SDADSC as follows in the subsequent technical result whose proofs is obvious from (5.1)-(5.2).

Theorem 5.1. (1) Assume that a SDADSC is defined to generate the set of sampling instants SI with a potential set of amplitudes $ST\delta := \{\delta_k \in \mathbf{R}_+ : 0 < \underline{\delta} \leq \delta_k \leq \bar{\delta} < \infty, \text{ for all } k \in ID\}$ so that ID has a finite cardinal card $ID = \text{fl}_0 \geq 1$ (If $\gamma_0 = 1$, one has the particular CADSC);

(2) consider any strictly ordered finite sequence of $\gamma \geq \gamma_0$ amplitudes with possible repetitions $\overline{ST\delta} = \overline{ST\delta}(\gamma) := \{\bar{\delta}_1, \bar{\delta}_2, \dots, \bar{\delta}_\gamma : \bar{\delta}_i = \delta_j \in STD, \forall i \in \bar{\gamma}, \text{ some } j \in \bar{\gamma}_0\}$;

Let a finite real number M be defined as $M = M(j) := \sum_{i=0}^{\ell} \delta(t_i)$, where $\ell \in \mathbf{N}_0$ is some finite positive integer defined according to $\delta(t_i) = \delta_k \in ST\delta$ for all $t_i \in SI$ and some chosen $k = k(t_i) \in \bar{\gamma}_0$. Also, define accordingly a set of real numbers $M_1 = M + \bar{\delta}_1$ and $M_{i+1} = M_i + \bar{\delta}_i$ for all $i \in \bar{\gamma}$ is defined from the given set $ST\delta$ of amplitudes. If $\gamma > \gamma_0$, then $\overline{ST\delta}$ contains $(\gamma - \gamma_0)$ repeated elements.

Thus, if (5.1) is defined with $p = j + \gamma + 1$ and (5.2) holds with its right-hand side being replaced with the tuple $[M, M_1, \dots, M_{\gamma+1} + \tau]$, then a limit oscillation exists which satisfies the extended version of (5.1) under the above replacements.

Note from Theorem 5.1 that by appropriate choice of the limiting sequence $\overline{ST\delta}$ of amplitudes, the amplitudes of sustained oscillations might be reduced compared to the use of a single amplitude.

Example 5.2. First, consider the linear dynamic system of transfer function:

$$G(s) = \frac{Y(s)}{U(s)} = K \frac{s+1}{s^2}, \quad (5.3)$$

where $Y(s)$ and $U(s)$ are the Laplace transforms in the Laplace argument “ s ” of the output $y(t)$ of a linear time-invariant dynamic system and its time-differentiable control input $u(t)$ under zero initial conditions. Under linear unit control feedback $u(t) = -y(t)$, the closed-loop differential equation becomes $y''(t) + K(y'(t) + y(t)) = 0$ subject to any initial conditions $y(0) = y_0$ and $y'(0) = y_{01}$. Note the following features in the context of limit sustained oscillations:

(1) this system is globally asymptotically stable (then, its solution $y(t)$ is oscillation-free) for any $K > 0$ and might describe a wide set of real processes, for instance, a mechanical system subject to damping and stiffness or the control of the angular position of a satellite with respect to its axis under a derivative or tachometric control. It can also describe mathematically a linear electric circuit with two energy storing devices specified by capacitors and/or inductors with at least one dissipative device, that is, a resistor which can be in practice either a separated dissipative device or dissipative effects of the inductors/capacitors;

(2) on the other hand, note that, in order to design an electronic oscillator, that is, an electronic system whose asymptotic solution is periodic irrespective of the initial conditions, a nonlinear effect should be included in the system. In this context, note that the solution of the above damped second-order differential equation converges asymptotically to zero for any initial conditions and then it is not periodic so that it cannot be used in that way for the design of oscillators. Note, furthermore, that a typical and well-used class of electric oscillators in

applications consists of those being typically synthesized with a saturation function $f(u(t)) = \text{sat}_{u_{M_0}, u_M}(u(t))$ of a certain amplifier linear gain $K = u_M/u_{M_0}$ in the linear mode, of saturation threshold u_{M_0} and saturated value u_M , that is,

$$\text{sat}_{u_{M_0}, u_M}(u(t)) = \begin{cases} Ku(t) & \text{if } |u(t)| \leq u_{M_0}, \\ u_M \text{sign}(u(t)) & \text{otherwise,} \end{cases} \quad (5.4)$$

together with a linear electric network being of at least third order. The oscillation condition at a frequency ω_0 is that the first-harmonic of the closed-loop response of the frequency domain satisfies $1 + KG(j\omega_0) = 0$, obtained under the replacement $s \rightarrow j\omega_0$ with $j = \sqrt{-1}$, provided that such an equation has a real solution $\omega_0 > 0$. The amplitude \bar{y} of the such a first harmonic $y(t) = \bar{y} \sin(\omega_0 t + \varphi)$ is approximately calculated from the companion complex identity $G(j\omega_0) = C_{\text{sat}}(u_{M_0}, u_M)$, where $C_{\text{sat}}(u_{M_0}, u_M)$ is the critical locus (i.e., the minus describing function $-(\text{sat } u(t))/y(t)$, an extended concept of the frequency response for certain separable or analytical nonlinearities, [36]) which is real for the case of saturations parameterized by the pair (u_{M_0}, u_M) so that the gain in the linear mode of the saturation is $K = u_M/u_{M_0}$. Note that $1 + C_{\text{sat}}(u_{M_0}, u_M)G(j\omega_0) = 0$ replaces intuitively the condition $1 + KG(j\omega_0) = 0$ of complex conjugate modes for the replacement $(-1/K) \rightarrow C_{\text{sat}}(u_{M_0}, u_M)$. The precision of the computation of the locus $C_{\text{sat}}(u_{M_0}, u_M)$, and then the precision of the calculated amplitude of the oscillation first-harmonic, depends on the type of describing function calculated for the saturation. See [36] and references therein for a number of useful describing functions/critical locus for different nonlinearities through different, but mutually close, useful definitions of describing function. The temporal asymptotic solution $y(t)$ tends to the limit cycle of first-harmonic $y(t) = \bar{y} \sin(\omega_0 t + \varphi)$ for any initial conditions;

(3) it is well known that electronic oscillators with basic saturated amplifiers of gain K (in their linear mode) require also linear network of at least third-order to be synthesized. See, for instance, [36]. This is because the impulse response hodograph $G(j\omega)$ (being the Fourier transform, if it exists, of the impulse response of the dynamic system) of the linear feed-forward part of first- and second-order jointly stable and inversely stable systems (i.e., both poles and zeros are in $\text{Re } s < 0$) are always in the third and fourth quadrants of the complex plane. As a result, they cannot cut the critical locus of a saturation nonlinearity for some frequency since such a critical locus is always allocated in the negative real semi axis;

(4) it is now described, in the context of the current problem at hand, how sustained oscillations can be obtained from the above described CADSC and SDADSC criteria by using just second-order systems of transfer functions $G(s) = K(s+1)/s^2$ in the feedforward loop. This implies that the order of the auxiliary linear network to synthesize the oscillator can be diminished related to the typical design using electronic circuitry whose basic amplifier in saturation mode needs the use of an auxiliary network of at least third-order. Then, consider again the feedback differential equation referred to above but under discrete control at, in general, nonperiodic sampling for CADSC and SDADSC:

$$y''(t) + K(y'(t) + 1) = Ku(t), \quad u(t) = -y(t_i), \quad \forall t \in [t_i, t_{i+1}), \quad (5.5)$$

where $\{t_i\}$ is the real sequence of sampling instants, and $\{T_i = t_{i+1} - t_i\}$ is the real sequence of sampling periods under the CADSC or the SDADSC sampling criteria for all $i \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

The above feedback system is a regulator since the control signal $u(t) = e(t) = -y(t_i)$ for all $t \in [t_i, t_{i+1})$ is generated by a zero-order hold under an identically zero external reference signal. If the control is identically zero then the resulting linear feedback is globally asymptotically Lyapunov stable to the origin which is the sole stable equilibrium point. However, the use of the sampling criteria translates into the presence of limit cycles that is, asymptotic oscillations being the limits of the solution trajectories in the phase plane. The following values are taken: $K = 1$, $A = 0.2154$ for the single threshold case as the sampling amplitude in the CADSC criterion and the set of amplitudes $S_{\text{amp}} \equiv \{A_1, A_2, A_3, A_4\} = \{0.12, 0.1677, 0.2154, 0.2631\}$ for the SDADSC criterion injected in this order to implement the sampling criterion. Since the control is a regulator, $r \equiv 0$ for all time, then the sampling criterion becomes

$$|e(t) - e(t_i)| = |y(t) - y(t_i)| = A, \quad \forall t \in [t_i, t_{i+1}), \quad (5.6)$$

for the CADSC, and

$$|e(t) - e(t_i)| = |y(t) - y(t_i)| = A(t_i) = A_{[i/4]}, \quad \forall t \in [t_i, t_{i+1}), \quad (5.7)$$

where $[i/4] = \text{Integer Part}(i/4)$; for all $i \in \mathbf{N}_0 = \mathbf{N} \cup \{0\}$ for the SDADSC. Note that the sampling criteria (5.6) and (5.7) can be interpreted as a separated nonlinearity of the dynamic systems (5.5) consisting of a multi-relay with hysteresis displayed in Figure 1 where $m(t_i) = -u(t_i)$. In such a way, the sampling and hold device with the sampling criteria is equivalent to such a nonlinearity which could be potentially be generated in a completely different way by using relays with hysteresis. In the case of the SDADSC, the amplitudes are taken to vary consecutively in the defined order in the set of amplitudes of the sampling criterion. The asymptotic phase plot for the CADSC criterion and SDADSC criterion are, a limit cycle of fundamental amplitude and frequency 0.39 and 0.3079 cycles/sec., and another limit cycle of amplitude 0.31 and frequency 0.3095 cycles/sec., respectively. Both limit cycles to which the phase portraits of the trajectory solutions asymptotically converge are shown in Figure 2. A direct interpretation of why the asymptotic solution is a stable limit cyclic, so that the solution is bounded and the whole system is (nonasymptotically) stable relies on the equivalence of the tandem sampling criteria CADSC and SDADSC through a companion zero-order hold to two variants of the multiple hysteretic relay nonlinearity of Figure 1. Note that it is well known that nonlinear systems under certain conditions can generate limit cycles. The sequences of constant asymptotic sampling periods reached in both cases are also listed in the figure. Note that the fundamental amplitude of the second limit cycle corresponding to the SDADSC is reduced more that 20% with respect to the first one while the fundamental frequencies differ only in about 0.50%, the second one being very slightly larger than the first one. It has been also observed under exhaustive inspection of related examples by modifying their parameterizations that the duration of the transient time interval towards the limit cycle solution is slightly shorter under the first criterion compared to the second one.

If the sampling criterion is modified to the constant amplitude-based sampling criterion $|y(t) + y(t_i)| = A$ for all $t \in [t_i, t_{i+1})$ (i.e., modified CADSC) and to the multithreshold sampling criterion obtained via its right-hand-side replacement by the same multithreshold sequence as above, that is, $S_{\text{amp}} \equiv \{A_1, A_2, A_3, A_4\} = \{0.12, 0.1677, 0.2154, 0.2631\}$ (i.e., modified SDADSC) then one gets the results of Figure 3 below. Note the complex geometry of the asymptotic oscillations of the standard criteria displayed in Figure 2 compared to the more smooth shaped ones of the modified ones displayed in Figure 3. It can be pointed

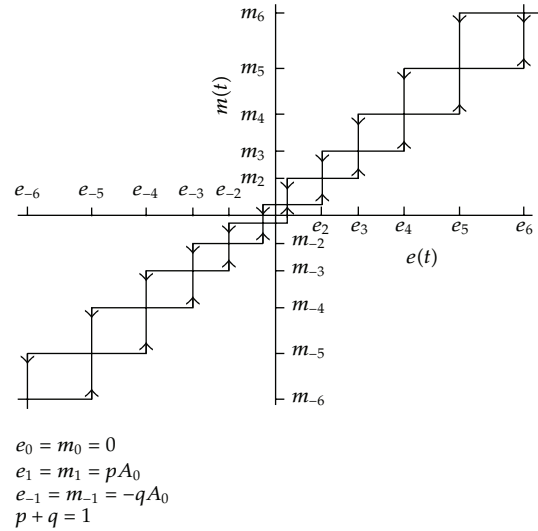


Figure 1: Multirelay with hysteresis nonlinear characteristics of the sampling criterion.

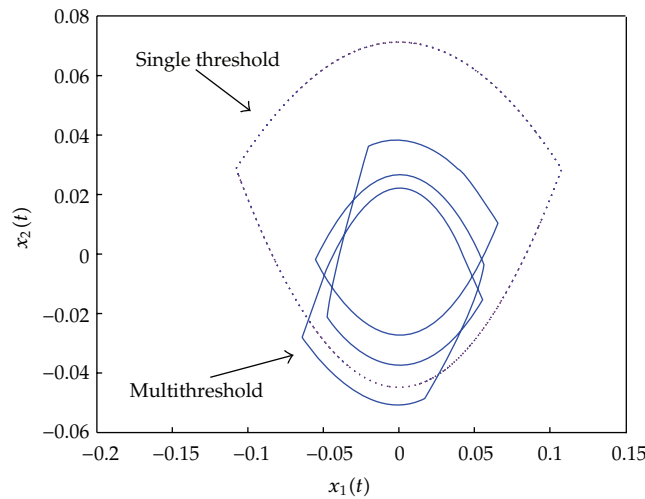


Figure 2: Phase plane plot for the solution versus its first time derivative exhibiting closed limit trajectories.

out that other potential sampling criteria, not been subject to constant or varying (within a prescribed set) differences of amplitude can lead to asymptotically stable solutions provided that the admissibility domain for the sampling intervals defined by such sampling criteria is constrained to the stability domain of a constant sampling provided that the continuous part of the dynamic system is globally asymptotically stable. See, for instance, Theorem 2.3.

Remark 5.3. Note that Example 5.2 is based on a transfer function description of the linear part. Thus, the above mathematical results on the limit asymptotic solutions are applicable to any minimal state-space realization, since in this case the dimension of the linear system coincides with the order of the transfer function (i.e., its number of poles). In the case of nonminimal realizations (then being either noncontrollable or nonobservable or both), the above discussed results still hold if the cancelled modes are strictly stable since their

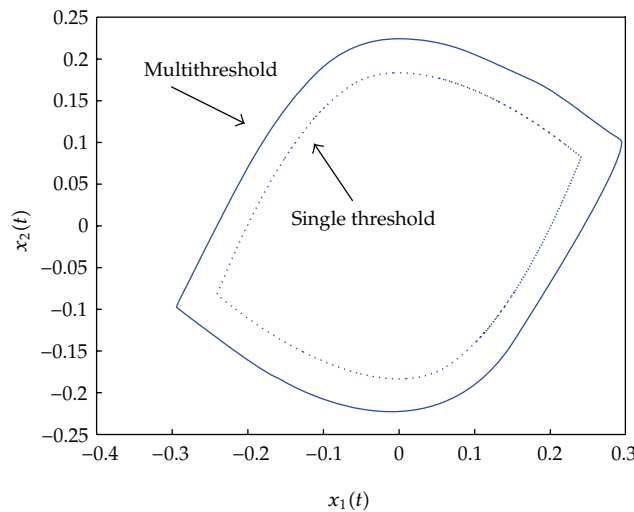


Figure 3: Phase plane plot of the modified sampling criteria for the solution versus its first time derivative exhibiting limit cycles.

contribution to the state-space trajectories and their time derivatives the relevant order vanish asymptotically as time tends to infinity.

6. Conclusions

This paper has been devoted to investigate the solutions and, in particular, their stability and instability properties as well as the possible presence of sustained oscillations in discrete linear dynamic systems under sampling laws which generate time-varying sampling periods in general. Two sampling criteria have been specially emphasized, namely, (a) the so-called *constant amplitude difference sampling criterion (CADSC)*, under which the signal of interest is sampled at each time that it reaches a prescribed threshold variation which is the positive real constant defining the sampling criterion; (b) the more general sampling criterion is referred to as a *sampling-dependent amplitude difference sampling criterion (SDADSC)* which involves a set of at least two distinct of such amplitudes. Both sampling criteria possess the property that, together with their associate sampling and zero-order hold device, are characterized as a relay with multiple hysteresis. Such a nonlinear model is expected to potentially generate potentially sustained limit oscillations of the solution. The analysis has been fully performed in the time domain so that, contrarily to the case of the use of frequency-domain analysis methods, no specific assumption is needed about low-pass filtering constraints of the linear auxiliary network in order to perform the analysis of the first-harmonic of the existing sustained oscillations. It is noticed that, the proposed analysis, no separation of the first-order harmonic of the whole oscillation has to be taken in mind.

Acknowledgments

The authors thank the Spanish Ministry of Education and the Basque Government for their support of this work through Grants DPI2009-07197, IT378-10, and SAIOTEK S-PE09UN12. They also thank the editor of the special issue of "Modelling and Control of Complex

Dynamic Systems: Applied Mathematics Aspects" and the anonymous referees and editor for their useful comments which helped the authors to improve the former versions of the paper.

References

- [1] J. C. Soto and M. De La Sen, "On the derivation and analysis of a non-linear model for describing a class of adaptive sampling laws," *International Journal of Control*, vol. 42, no. 6, pp. 1347–1368, 1985.
- [2] M. de la Sen, "A method for improving the adaptation transient using adaptive sampling," *International Journal of Control*, vol. 40, no. 4, pp. 639–665, 1984.
- [3] M. de la Sen, "Adaptive sampling for improving the adaptation transients in hybrid adaptive control," *International Journal of Control*, vol. 41, no. 5, pp. 1189–1205, 1985.
- [4] M. De La Sen, "A new modelling for aperiodic sampling systems," *International Journal of Systems Science*, vol. 15, no. 3, pp. 315–328, 1984.
- [5] G. Ge and W. Zhiwen, "Stability of control systems with time-varying sampling," in *Proceedings of the 6th World Congress on Intelligent Control and Automation (WCICA '06)*, pp. 651–656, June 2006.
- [6] S. Luo, "Error estimation for non-uniform sampling in shift invariant space," *Applicable Analysis*, vol. 86, no. 4, pp. 483–496, 2007.
- [7] W. R. Dieter, S. Datta, and W. K. Kai, "Power reduction by varying sampling rate," in *Proceedings of the International Symposium on Low Power Electronics and Design*, pp. 227–232, August 2005.
- [8] G. M. Andrew, "Control and guidance systems with automatic aperiodic sampling," *Journal of Spacecraft and Rockets*, vol. 12, no. 1, p. 59, 1975.
- [9] M. de la Sen and A. Gallego, "Adaptive control for nonperiodic sampling using bilinear models," *International Journal of Systems Science*, vol. 22, no. 8, pp. 1403–1418, 1991.
- [10] M. J. Chambers, "Testing for unit roots with flow data and varying sampling frequency," *Journal of Econometrics*, vol. 144, no. 2, pp. 524–525, 2008.
- [11] T. Shibata, T. Bando, and S. Ishii, "Visual tracking achieved by adaptive sampling from hierarchical and parallel predictions," in *Proceedings of the 14th International Conference on Neural Information Processing (ICONIP '07)*, vol. 4984 of *Neural Information Processing, Part I Book Series: Lecture Notes in Computer Science*, pp. 604–613, Kitakyushu, Japan, 2008.
- [12] M. H. F. Zarandi, A. Alaeddini, and I. B. Turksen, "A hybrid fuzzy adaptive sampling—run rules for Shewhart control charts," *Information Sciences*, vol. 178, no. 4, pp. 1152–1170, 2008.
- [13] M. De la Sen, "On the properties of reachability, observability, controllability, and constructibility of discrete-time positive time-invariant linear systems with aperiodic choice of the sampling instants," *Discrete Dynamics in Nature and Society*, vol. 2007, Article ID 84913, 23 pages, 2007.
- [14] M. de la Sen, "Application of the nonperiodic sampling to the identifiability and model matching problems in dynamic systems," *International Journal of Systems Science*, vol. 14, no. 4, pp. 367–383, 1983.
- [15] M. Delasen, J. M. Sandoval, and Yu. Kryukov, "An adaptive sampling approach to the problem of learning and control in manipulator systems," *Revista de Informatica y Automatica*, vol. 22, no. 4, pp. 31–44, 1989.
- [16] M. de la Sen, "Non-periodic and adaptive sampling. A tutorial review," *Lithuanian Academy of Sciences. Informatica*, vol. 7, no. 2, pp. 175–228, 1996.
- [17] M. De la Sen, J. J. Miñambres, A. J. Garrido, A. Almansa, and J. C. Soto, "Basic theoretical results for expert systems. Application to the supervision of adaptation transients in planar robots," *Artificial Intelligence*, vol. 152, no. 2, pp. 173–211, 2004.
- [18] J. C. Soto and M. De La Sen, "Nonlinear oscillations in nonperiodic sampling systems," *Electronics Letters*, vol. 20, no. 20, pp. 816–818, 1984.
- [19] C. Kadilar, Y. Unyazici, and H. Cingi, "Ratio estimator for the population mean using ranked set sampling," *Statistical Papers*, vol. 50, no. 2, pp. 301–309, 2009.
- [20] C. Kadilar and H. Cingi, "Improvement in variance estimation in simple random sampling," *Communications in Statistics. Theory and Methods*, vol. 36, no. 9–12, pp. 2075–2081, 2007.
- [21] C. Kadilar and H. Cingi, "Ratio estimators for the population variance in simple and stratified random sampling," *Applied Mathematics and Computation*, vol. 173, no. 2, pp. 1047–1059, 2006.
- [22] C. Kadilar and H. Cingi, "Improvement in estimating the population mean in simple random sampling," *Applied Mathematics Letters*, vol. 19, no. 1, pp. 75–79, 2006.

- [23] M. Miskowicz, "Asymptotic effectiveness of the event-based sampling according to the integral criterion," *Sensors*, vol. 7, no. 1, pp. 16–37, 2007.
- [24] M. Miskowicz, "Send-on-delta concept: an event-based data reporting strategy," *Sensors*, vol. 6, no. 1, pp. 49–63, 2006.
- [25] E. Oral and C. Kadilar, "Improved ratio estimators via modified maximum likelihood," *Pakistan Journal of Statistics*, vol. 27, no. 3, pp. 269–282, 2011.
- [26] A. J. Garrido, M. De la Sen, and R. Bárcena, "Semi-heuristically obtained discrete models for LTI Systems under real sampling with choice of the hold device," in *Proceedings of the American Control Conference*, vol. 1–6, pp. 71–76, June 2003.
- [27] M. De la Sen, R. Bárcena, and A. J. Garrido, "On the intrinsic limiting zeros as the sampling period tends to zero," *IEEE Transactions on Circuits and Systems. I. Regular Papers*, vol. 48, no. 7, pp. 898–900, 2001.
- [28] M. Mellado, S. Dormido, and J. M. Guillen, "Introduction to control of fixed interval systems," *Anales de Fisica*, vol. 66, no. 1-2, p. 33, 1970.
- [29] J. Sánchez, M. A. Guarnes, and S. Dormido, "On the application of different event-based sampling strategies to the control of a simple industrial process," *Sensors*, vol. 9, no. 9, pp. 6795–6818, 2009.
- [30] S. Dormido, M. Mellado, J. Ruiz, and J. M. Guillen, "Sistemas de muestreo adaptivo mediante un criterio de diferencia de amplitudes constante," *Revista de Informatica y Automatica*, vol. 16, pp. 13–17, 1973.
- [31] S. Dormido and M. Mellado, "Determinacion de ciclos limite en sistemas de muestreo adaptivo," *Revista de Informatica y Automatica*, vol. 26, no. 4, pp. 21–33, 1975.
- [32] M. B. Paz, M. de la Sen, S. Dormido, and M. Mellado, "Compensation of discrete Systems to variations in their parameters by changing sampling period," *Electronics Letters*, vol. 18, no. 10, pp. 404–406, 1982.
- [33] K. J. Åström and B. Bernhardsson, "Systems with Lebesgue sampling," in *Directions in Mathematical Systems Theory and Optimization*, A. Rantzer and C. I. Byrnes, Eds., vol. 286 of *Lecture Notes in Control and Information Sciences*, pp. 1–13, Springer, Berlin, Germany, 2003.
- [34] E. Kofman and J. H. Braslavsky, "Level crossing sampling in feedback stabilization under data-rate constraints," in *Proceedings of the 45th IEEE Conference on Decision and Control (CDC '06)*, vol. 1–14, pp. 4423–4428, December 2006.
- [35] M. De la Sen, "Sufficiency-type stability and stabilization criteria for linear time-invariant systems with constant point delays," *Acta Applicandae Mathematicae*, vol. 83, no. 3, pp. 235–256, 2004.
- [36] D. P. Atherton, *Nonlinear Control Engineering*, Van Nostrand Reinhold, London, UK, 1975.

Research Article

The Analysis and Control for Singular Ecological-Economic Model with Harvesting and Migration

Qingling Zhang,¹ Hong Niu,¹ Lichun Zhao,² and Fenglan Bai³

¹ Institute of Systems Science, Northeastern University, Liaoning, Shenyang 110819, China

² Department of Mathematics, Anshan Normal University, Liaoning, Anshan 114007, China

³ School of Science, Dalian Jiaotong University, Liaoning, Dalian 116028, China

Correspondence should be addressed to Qingling Zhang, qlzhang@mail.neu.edu.cn

Received 12 March 2012; Accepted 11 April 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 Qingling Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To keep the resources renewable, a singular ecological-economic model is proposed for the populations with harvesting and migration. The local stability and the dynamic behavior of the model are studied. Singular induced bifurcation appears when economic interest is zero, which is different from the ordinary differential models. In order to apply variable structure control to eliminate these complex behaviors, the singular model is transformed into a single-input and single-output model with parameter varying within definite intervals. And then, a variable structure controller is designed to make the model stable. Finally, an inshore-offshore fishery model is given to illustrate the proposed method, and some numerical simulations are shown to demonstrate the control results.

1. Introduction

The management of renewable resources is important for the development of human and society. In exploiting the biological resources, both the economic profit and the environmental effects should be taken into account, which initiates a new research area: biomathematics. Interactions of mathematics and biology promote the development of the biosciences greatly in a certain extent. Since most of biological theories evolve rapidly, it is necessary to develop some useful mathematical models to describe the consequences of these biological models.

Singular model as a branch of modern control theory can describe a class of practical models more accurately. Compared with the ordinary differential models, singular models exhibit more complicated dynamics, such as the impulse phenomenon. They have more applications in power systems, aerospace engineering, chemical processes, social economic systems, biological systems, network analysis, and so forth. With the help of the singular

models for the power systems and bifurcation theory, complex dynamical behaviors of the power systems have been extensively studied, which reveal the instability mechanism of power systems [1–3]. Applications of singular models are also found in neural networks [4], fault diagnosis [5, 6], robotics [7, 8] and epidemic [9–11], economics [12, 13], and chemistry [14]. As far as the singular system theory is concerned, there are a few research results in biology. Since a singular biological economics model with stage structure was established to model the biological systems in [15], some singular biological models appeared [16–19]. These ideas are based on the economic theory [20]:

$$\text{Net Economic Revenue} = \text{Total Revenue} - \text{Total Cost}. \quad (1.1)$$

This formula presents some solid preliminary on singular biological systems.

In biology, many mathematicians, ecologists, and economists are concerned with the exploitation of renewable resources in recent years, and some results are achieved [21–24]. Though the harvesting can bring economic profit for people, the overexploitation may cause the extinction of some populations. In order to prevent the population from damages, some methods are introduced, such as, to raise taxes or to make the young population forbidden to be harvested. We propose a singular ecological-economic model to model such a problem. Singular model is often strongly nonlinear and unstable. In this case, one of control methods, which are able to perform high-quality automatic control, is demanded.

Variable structure control is considered to be used in this paper. It is a flexible control method to deal with some models with uncertain parameters and external disturbances. The main advantage of this technique is that once the system state variables reach a sliding surface, the structure of the feedback loop is adaptively altered to slide the state variables along the sliding surface. Thereafter, the system response depends on the gradients of the sliding surface and remains insensitive to parameter variations and external disturbances. Variable structure control with sliding mode was first proposed by Emelyanov [25] and was elaborated in the 1970s [26, 27]. In their pioneer works, variable structure controls are used to handle some linear models, and then expanded to nonlinear models, multi-input and multioutput models, discrete time models, infinite-dimensional models and stochastic models [28–33]. In recent years, variable structure control is applied to a wide variety of engineering fields successfully, such as robot control, flight control, motor control, and power control [34–36].

The main contents of the paper are as follows. In Section 1, in order to prevent the extinction of some populations, a singular ecological-economic model is proposed for the populations with harvesting and migration. In Section 2, when the local stability and the dynamic behavior for the model are discussed, singular induced bifurcation appears, and a control method is demanded to eliminate this bifurcation. In Section 3, in order to apply variable structure control, the singular model is transformed into a single-input and single-output model with parameters varying within definite intervals. In Section 4, an inshore-offshore model is given to illustrate the analysis results, and the simulations illustrate the effectiveness of the proposed method.

2. Modeling

In order to model growth of the populations, numerous models have been introduced. The generalized logistic growth model can provide an adequate approximation for the growth

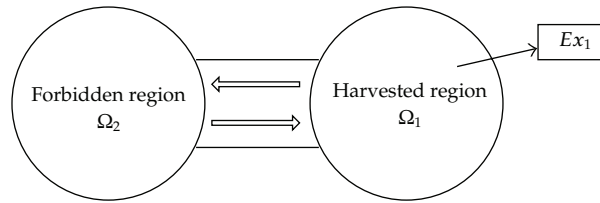


Figure 1: The migration between the harvesting permitted region and the harvesting forbidden region.

of the populations. However, if there is no harvesting, the populations would continue to increase rapidly. Therefore, the harvesting is an effective measure to maintain the diversity of species and protect the renewable resources. The equation of the harvested populations reads

$$\frac{dx}{dt} = \gamma x \left(1 - \frac{x}{K}\right) - qEx, \quad (2.1)$$

where x is the number of population, γ is a positive constant which is called the intrinsic growth rate, E is the harvesting effort, q is catch-ability coefficient, and K is usually the environment carrying capacity or saturation level. Some papers studied the model (2.1) with a constant effort E . But it is only suitable for some special case. In practice, the harvesting effort E is usually time-varying. For convenience in calculation, the condition $q = 1$ is usually assumed. If $E > \gamma$, a rapid collapse of the populations will occur. The extinction of population is inevitable, and the ecological balance will be destroyed.

In order to keep the resources renewable and prevent the extinction of some populations, the populations can be divided into two regions: Ω_1 and Ω_2 . In region Ω_1 , the harvesting is permitted, while the harvesting is forbidden in region Ω_2 . If there is a difference between region Ω_1 and region Ω_2 , the migration can occur between two regions, which is assumed to be proportional to the difference, and the proportional coefficient is positive. To better understand relation of the populations between two regions, a concise schematic is shown in Figure 1.

From Figure 1, it can be seen that the two regions are connected, and the population can migrate freely between two regions. The number of population in region Ω_1 is x_1 , and x_2 is the number of population in region Ω_2 . The ultimate aim in harvesting the biological resources is to get economic profits and practical value. Generally, from the economic aspect, we know that the harvesting behavior changes with many market factors. Therefore, studying the relation between economic profits and the harvesting can help us better protect sustainable resources. If considering the economic profits in the model (2.1), the following mathematical model of the exploited population with protective region, called a singular ecological-economic model, is proposed:

$$\begin{aligned} \frac{dx_1}{dt} &= \gamma x_1 \left(1 - \frac{x_1}{K_1}\right) - \alpha(x_1 - x_2) - Ex_1, \\ \frac{dx_2}{dt} &= \gamma x_2 \left(1 - \frac{x_2}{K_2}\right) + \alpha(x_1 - x_2), \\ 0 &= (px_1 - c)E - m, \end{aligned} \quad (2.2)$$

where p is the unit price, c is the unit cost, m is the economic profit, and $\alpha > 0$ is the migration proportional coefficient between two regions. Considering the practical significance, p and c are positive constants. K_1 and K_2 are the environment carrying capacity of Ω_1 and Ω_2 , respectively. The differential equations are the growth rate of the population in region Ω_1 and Ω_2 . The algebraic equation is an economic model, which represents the relations of the total income, the total cost, and the economic profits.

Remark 2.1. In the management of sustainable resources, the model (2.2) not only considers the ecological balance but also includes the economic profits when the population is harvested, which combines the biological control problems with the economic problems. The model (2.2) provides an effective way for human being to maintain the ecological resources sustainable when we get economic profit.

Considering the biological significance, the model (2.2) is discussed in the following interval:

$$R_+^3 = \{\chi = (x_1, x_2, E) \mid x_1 \geq 0, x_2 \geq 0, E \geq 0\}. \quad (2.3)$$

If $m = 0$, the economic equilibrium occur, that is, the income is equal to the cost. When the economic profit is zero, the population reaches the maximum harvesting effort, and it is called the overfishing. In the exploitation of population resources, a collapse of the population may occur.

Due to the limitation of the environment, the number of populations cannot exceed the environment maximum carrying capacity. Otherwise, due to the crowded environment, a large number of populations will die gradually. So the state variables and the parameters satisfy the following conditions:

$$0 < x_1 < K_{1\max}, \quad 0 < x_2 < K_{2\max}, \quad 0 < K_1 < K_{1\max}, \quad 0 < K_2 < K_{2\max}, \quad 0 < E < \gamma, \quad (2.4)$$

where $K_{1\max}$ and $K_{2\max}$ are the maximum environment carrying capacities of Ω_1 and Ω_2 , respectively.

3. Local Stability Analysis

For convenience, the environment carrying capacity in Ω_1 is assumed to be proportional to that in Ω_2 , and the ratio is η , that is, $\eta K_1 = K_2$ ($\eta > 0$). For the model (2.2), the equilibrium points are the solutions for the equations:

$$\begin{aligned} \gamma x_1 \left(1 - \frac{x_1}{K_1}\right) - \alpha(x_1 - x_2) - Ex_1 &= 0, \\ \gamma x_2 \left(1 - \frac{x_2}{\eta K_1}\right) + \alpha(x_1 - x_2) &= 0, \\ (px_1 - c)E - m &= 0. \end{aligned} \quad (3.1)$$

By solving (3.1), we get two equilibrium points for the model (2.2):

$$p_0 = \left(0, 0, -\frac{m}{c}\right),$$

$$p_1 = (x_{10}, x_{20}, E_0) = \left(x_0, \left(\frac{m}{\alpha(px_0 - c)} + 1 - \frac{\gamma}{\alpha} + \frac{\gamma x_0}{\alpha K_1}\right)x_0, \frac{m}{px_0 - c}\right). \quad (3.2)$$

Here $x_0 \neq 0$ is the root of the equation:

$$C_0 x^3 + C_1 x^2 + C_2 x + C_3 = 0, \quad (3.3)$$

where $C_0 = \gamma p(K_1 \alpha - K_1 \gamma + \gamma)^2$, $C_1 = 2\eta K_1 p m \gamma (K_1 \alpha - K_1 \gamma + \gamma) - 2pc\gamma(K_1 \alpha - K_1 \gamma + \gamma)^2$, $C_2 = \gamma c^2(K_1 \alpha - K_1 \gamma + \gamma)^2 - 2\eta K_1 c m \gamma (K_1 \alpha - K_1 \gamma + \gamma) + \alpha p \eta^3 K_1^3 m(\alpha - \gamma) + K_1^2 m^2 \gamma$, $C_3 = -\alpha \eta^3 K_1^3 m c(\alpha - \gamma)$.

When the coefficients C_i ($i = 0, 1, 2, 3$) satisfy certain conditions, there is a positive solution for (3.3). Here, we suppose that the positive equilibrium point p_1 exists. We are interested in the local stability of the model (2.2) at the equilibrium points p_0 and the positive equilibrium point p_1 . In order to analyze the local stability of the model (2.2), let

$$\mathbf{F}(\mathbf{X}, E) = \begin{pmatrix} \gamma x_1 \left(1 - \frac{x_1}{K_1}\right) - \alpha(x_1 - x_2) - E x_1 \\ \gamma x_2 \left(1 - \frac{x_2}{\eta K_1}\right) + \alpha(x_1 - x_2) \end{pmatrix} \quad (3.4)$$

$$G(\mathbf{X}, E) = (px_1 - c)E - m,$$

where $\mathbf{X} = [x_1 \ x_2]$. The local stability of the model (2.2) at the equilibrium points p_0 is discussed by the following theorem.

Theorem 3.1. *If $0 < \gamma < \alpha + m/2c$ and $\gamma(\gamma - 2\alpha) + (\gamma - \alpha + m/c) > 0$, the model (2.2) is locally stable at p_0 .*

Proof. $p_0 = (0, 0, -m/c)$ is an equilibrium point of the model (2.2). Since $\det D_E G|_{p_0} = -c \neq 0$, Jacobian matrix of the model (2.2) at p_0 is given by

$$J = \left[D_{\mathbf{X}} \mathbf{F} - D_E \mathbf{F} (D_E G)^{-1} D_{\mathbf{X}} G \right] \Big|_{p_0} = \begin{bmatrix} \gamma - \alpha + \frac{m}{c} & \alpha \\ \alpha & \gamma - \alpha \end{bmatrix}, \quad (3.5)$$

where $D_E G$ denote the derivative of the function G on the variable E .

The characteristic equation of Jacobian matrix (3.5) can be obtained:

$$\lambda^2 - \left[2(\gamma - \alpha) + \frac{m}{c} \right] \lambda + \gamma(\gamma - 2\alpha) + \left(\gamma - \alpha + \frac{m}{c} \right) = 0. \quad (3.6)$$

If $0 < \gamma < \alpha + m/2c$ and $\gamma(\gamma - 2\alpha) + (\gamma - \alpha + m/c) > 0$, the roots of the characteristic equation (3.6) all have negative real part. Therefore, the model (2.2) is locally stable at p_0 . \square

In order to analyze the local stability at the positive equilibrium point p_1 , a linear transformation $\chi^T = \mathbf{Q}\mathbf{S}^T$ is used, where

$$\chi = [x_1 \ x_2 \ E], \quad \mathbf{S} = [u \ v \ \bar{E}], \quad \mathbf{Q} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{pE_0}{px_0 - c} & 0 & 1 \end{pmatrix}. \quad (3.7)$$

Thus, $D_\chi G(\chi_0)\mathbf{Q} = \begin{bmatrix} 0 & 0 & px_0 - c \end{bmatrix}$, $u = x_1$, $v = x_2$, $\bar{E} = pE_0/(px_0 - c) + E$. The model (2.2) is changed into the following form:

$$\begin{aligned} \frac{du}{dt} &= \gamma u \left(1 - \frac{u}{K_1}\right) - \alpha(u - v) - \bar{E}u + \frac{pE_0}{px_0 - c}u^2, \\ \frac{dv}{dt} &= \gamma v \left(1 - \frac{v}{\eta K_1}\right) + \alpha(u - v), \\ 0 &= (pu - c) \left(\bar{E} - \frac{pE_0}{px_0 - c}u\right) - m. \end{aligned} \quad (3.8)$$

Now the local stability of the model (3.8) at the positive equilibrium point p_1 will be analyzed. First, the diffeomorphism ψ is defined as follows:

$$\begin{bmatrix} u & v & \bar{E} \end{bmatrix}^T = \psi(\bar{\mathbf{Z}}) = \mathbf{S}_0^T + \mathbf{U}_0\bar{\mathbf{Z}} + \mathbf{V}_0h(\bar{\mathbf{Z}}), \quad (3.9)$$

where $\mathbf{U}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\mathbf{V}_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, $\bar{\mathbf{Z}} = [y_1 \ y_2]^T$, $\mathbf{S}_0 = [u_0 \ v_0 \ \bar{E}_0]$, $h : \mathbb{R}^2 \rightarrow \mathbb{R}^1$ is a smooth mapping. Jacobi matrix $D\psi$ is a 3×1 real matrix.

Second, by differentiating $G(\psi(\bar{\mathbf{Z}})) = 0$, the following equation is obtained:

$$DG(\chi)D\psi(\bar{\mathbf{Z}}) = 0. \quad (3.10)$$

Differentiating (3.9) and multiplying on the left by \mathbf{U}_0^T , it can be obtained that:

$$\mathbf{U}_0^T D\psi(\bar{\mathbf{Z}}) = \mathbf{I}_2, \quad (3.11)$$

where \mathbf{I}_2 is a 2×2 unit matrix. From (3.10) and (3.11), the following formula is gotten:

$$D\psi(\bar{\mathbf{Z}}) = \begin{bmatrix} DG(\mathbf{S}_0) \\ \mathbf{U}_0^T \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \mathbf{I}_2 \end{bmatrix}. \quad (3.12)$$

Furthermore, the following model is further obtained [37]:

$$\frac{d\bar{\mathbf{Z}}}{dt} = \mathbf{U}_0^T f(\psi(\bar{\mathbf{Z}})) = \mathbf{U}_0^T Df(\chi_0) \begin{bmatrix} DG(\mathbf{S}_0) \\ \mathbf{U}_0^T \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \mathbf{I}_2 \end{bmatrix} + Y(\bar{\mathbf{Z}}), \quad (3.13)$$

where $Y(\bar{\mathbf{Z}}) = o(\bar{\mathbf{Z}})$ ($\bar{\mathbf{Z}} \rightarrow 0^+$).

From the transformation above and (3.13), the coefficient matrix of linear model corresponding to the model (3.8) is gotten as follows:

$$\begin{aligned} E(\mathbf{S}_0) &= \begin{pmatrix} D_S f_1(\mathbf{S}_0) \\ D_S f_2(\mathbf{S}_0) \end{pmatrix} \begin{pmatrix} D_S G(\mathbf{S}_0) \\ \mathbf{U}_0^T \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \gamma - \bar{E}_0 - \alpha - \frac{2\gamma u_0}{K_1} + \frac{2pE_0 u_0}{px_0 - c} & \alpha \\ \alpha & \gamma - \alpha - \frac{2\gamma v_0}{\eta K_1} \end{pmatrix}, \end{aligned} \quad (3.14)$$

where $u_0 = x_0$, $v_0 = (m/\alpha(px_0 - c) + 1 - \gamma/\alpha + \gamma x_0/\alpha K_1)x_0$, $\bar{E}_0 = pE_0/(px_0 - c) + E_0$.

Thus, the characteristic equation of the matrix (3.14) is given by

$$\lambda^2 + D_1 \lambda + D_2 = 0, \quad (3.15)$$

where $D_1 = 2(\alpha - \gamma) + (3pE_0 x_0 + E_0(p - c))/(px_0 - c) + 2\gamma x_0/\eta K_1((m/\alpha(px_0 - c)) + 2 - (\gamma/\alpha) + (\gamma/\alpha K_1))$, $D_2 = [\gamma - \alpha + (3pE_0 x_0 + E_0(p - c))/(px_0 - c)][\gamma - \alpha - (2\gamma x_0/\eta K_1)((m/\alpha(px_0 - c)) + 2 - (\gamma/\alpha) + (\gamma/\alpha K_1))]$.

About the local stability of the model (2.2) at the positive equilibrium point p_1 , we have the following theorem.

Theorem 3.2. *For the model (2.2):*

- (a) *if $D_1 > 0$ and $D_2 > 0$, the model (2.2) is locally stable at the positive equilibrium point p_1 ;*
- (b) *if $D_1 < 0$ or $D_2 < 0$, the model (2.2) is unstable at the positive equilibrium point p_1 .*

Proof. The model (3.8) and the model (2.2) are isomorphic. The local stability of them is discussed by the eigenvalues of the coefficient matrix $E(\mathbf{S}_0)$. When $D_1 > 0$ and $D_2 > 0$, two roots of the characteristic equation (3.15) all have negative real part. The model (3.8) and the model (2.2) are all locally stable at the positive equilibrium point p_1 .

However, when $D_1 < 0$ or $D_2 < 0$, at least one of the eigenvalues of $E(\mathbf{S}_0)$ has nonnegative real part. We can conclude that the model (2.2) is unstable at the positive equilibrium point p_1 . Thus, the proof is completed. \square

To further study the dynamic behavior of the model (2.2), x_0 is given a specified value. If $x_0 = c/p$, the positive equilibrium point of the model (2.2) is

$$p_1(x_{10}, x_{20}, E_0) = \left(\frac{c}{p}, \frac{\gamma - \alpha + \theta}{2p\gamma}, (\gamma - \alpha) \left(\frac{\alpha}{2c\gamma} + 1 \right) - \frac{\gamma c}{pK_1} + \frac{\alpha\theta}{2c\gamma} \right), \quad (3.16)$$

where $\theta = \sqrt{(\gamma - \alpha)^2 p^2 \eta^2 K_1^2 + 4p\eta K_1 \gamma \alpha c}$. By analysis, we know that there is a bifurcation at the positive equilibrium point p_1 for the model (2.2), which is shown in the following theorem.

Theorem 3.3. *If $\gamma - (\gamma - \alpha + \theta)/K_1 p - \alpha \neq 0$, there is a singular induced bifurcation for the model (2.2) at the positive equilibrium point p_1 , and $m = 0$ is a bifurcation value.*

Proof. Let m be a bifurcation parameter for the model (2.2). $x_1 = c/p$ makes $\Delta = \det[D_E G] = px_1 - c = 0$. If $\gamma - (\gamma - \alpha + \theta)/K_1 p - \alpha \neq 0$, the following three conditions are satisfied:

$$\begin{aligned} & \text{(i) } \text{trace}[D_E F \text{adj}(D_E G)(D_{x_1} G \ D_{x_2} G)]_{p_1} \\ & \quad = \begin{pmatrix} -pEx_1 & 0 \\ 0 & 0 \end{pmatrix}_{p_1} \\ & \quad = -c(\gamma - \alpha) \left(\frac{\alpha}{2c\gamma} + 1 \right) + \frac{\gamma c^2}{pK_1} - \frac{\alpha\theta}{2\gamma} \neq 0; \\ & \text{(ii) } \begin{vmatrix} D_X F & D_E F \\ D_X G & D_E G \end{vmatrix}_{p_0} \\ & \quad = \begin{vmatrix} \gamma - \frac{2\gamma}{K_1} x_1 - \alpha - E & \alpha & -x_1 \\ \alpha & \gamma - \frac{2\gamma}{\eta K_1} x_2 - \alpha & 0 \\ pE & 0 & px_1 - c \end{vmatrix}_{p_1} \\ & \quad = \left[c(\gamma - \alpha) \left(\frac{\alpha}{2c\gamma} + 1 \right) - \frac{\gamma c^2}{p\eta K_1} + \frac{\alpha\theta}{2\gamma} \right] \left(\gamma - \frac{\gamma - \alpha + \theta}{K_1 p} - \alpha \right) \neq 0; \\ & \text{(iii) } \begin{vmatrix} D_X F & D_E F & D_m F \\ D_X G & D_E G & D_m G \\ D_X \Delta & D_E \Delta & D_m \Delta \end{vmatrix}_{p_0} \\ & \quad = \begin{vmatrix} \gamma - \frac{2\gamma}{K_1} x_1 - \alpha - E & \alpha & -x_1 & 0 \\ \alpha & \gamma - \frac{2\gamma}{\eta K_1} x_2 - \alpha & 0 & 0 \\ pE & 0 & px_1 - c & -1 \\ p & 0 & 0 & 0 \end{vmatrix}_{p_1} \\ & \quad = -\frac{c}{p} \left(\gamma - \frac{\gamma - \alpha + \theta}{K_1 p} - \alpha \right) \neq 0. \end{aligned} \quad (3.17)$$

Thus, we can conclude that there exists a smooth curve in \mathbf{R}^3 which passes through the positive equilibrium point p_1 , and it is transversal to the singular surface at the positive equilibrium point p_1 . And we can get the following equations:

$$\begin{aligned}
 i &= -\text{trace}[D_E \text{Fadj}(D_E G)(D_{x_1} G \ D_{x_2} G)]_{p_1} = c(\gamma - \alpha) \left(\frac{\alpha}{2c\gamma} + 1 \right) - \frac{\gamma c^2}{pK_1} + \frac{\alpha\theta}{2\gamma}; \\
 j &= D_m \Delta - (D_X \Delta \ D_E \Delta) \begin{pmatrix} D_X F & D_E F \\ D_X G & D_E G \end{pmatrix}^{-1} \begin{pmatrix} D_m F \\ D_m G \end{pmatrix} \\
 &= -(p \ 0 \ 0) \begin{pmatrix} \gamma - \frac{2\gamma}{K_1} x_1 - \alpha - E & \alpha & -x_1 \\ \alpha & \gamma - \frac{2\gamma}{\eta K_1} x_2 - \alpha & 0 \\ pE & 0 & px_1 - c \end{pmatrix}_{p_1}^{-1} \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \\
 &= \frac{1}{(\gamma - \alpha)(\alpha/2c\gamma + 1) - \gamma c/pK_1 + \alpha\theta/2c\gamma}.
 \end{aligned} \tag{3.18}$$

From above we can get that $i/j = c[(\gamma - \alpha)(\alpha/2c\gamma + 1) - \gamma c/pK_1 + \alpha\theta/2c\gamma]^2$. Obviously, $i/j > 0$. According to Theorem 3 in [38], when m passes through 0, one eigenvalue of matrix $J = D_X F - D_E F(D_E G)^{-1} D_X G$ moves from C^- to C^+ along the real axis by diverging through ∞ . There is a singular induced bifurcation for the model (2.2), and the model turns to unstable. The proof is completed. \square

Remark 3.4. When the economic profit is zero, it is called the overfishing in economics. One eigenvalue of the model (2.2) is approaching to endless, and the impulse occurs in the model (2.2). This would lead to the collapse of the population and destroy the ecological balance. It is necessary to find an effective method to make that the population develop sustainably.

4. Controller Design

Variable structure control is often used to deal with some models with internal varying parameters and external disturbances since it provides effective means to design robust state feedback controllers. In this section, variable structure control is introduced to eliminate the bifurcation behavior and ensure the system stable. This approach makes direct use of the nonlinear model and the full biological state information. In order to facilitate the controller design, differentiating the second differential equation in the model (2.2) and substituting the other two equations into it, the model (2.2) is transformed into a second-order differential equation [39]:

$$\begin{aligned}
 \frac{d^2 x_2}{dt^2} + \left(\alpha + \frac{2\gamma x_2}{K_2} - \gamma \right) \frac{dx_2}{dt} - \alpha^2 x_2 \\
 = \frac{K_1 p E (\alpha \gamma m - E c \alpha \gamma - \alpha^2 m - \alpha^2 c E - \alpha m E) - \alpha \gamma m (m + c E + c E) - c \alpha^2 \gamma E^2}{K_1 p^2 E^2} + \frac{c}{p} E.
 \end{aligned} \tag{4.1}$$

Equation (4.1) can be rewritten as a single-input and single-output model with the parameters varying within definite intervals:

$$\frac{d^2 y}{dt^2} + a_1 \frac{dy}{dt} + a_0 y = b_0 u + \beta, \quad (4.2)$$

where $y = x_2$, $u = E$, $a_1 = \alpha + 2\gamma x_2 / K_2 - \gamma$, $a_0 = -\alpha^2$, $b_0 = c/p$, $\beta = (K_1 p E (\alpha \gamma m - E c \alpha \gamma - \alpha^2 m - \alpha^2 c E - \alpha m E) - \alpha \gamma m (m + 2cE) - c \alpha^2 \gamma E^2) / K_1 p^2 E^2$.

Obviously, a_0 and b_0 are fixed, while a_1 and β change with the parameters and the variables. From the varying intervals (2.4), we can get the varying intervals of the coefficients a_1 and β :

$$\begin{aligned} \alpha - \gamma < a_1 < \alpha + \gamma, \\ \frac{-c\alpha\gamma - (\alpha^2 + \alpha)m}{p} < \beta < \frac{(\alpha\gamma - \alpha^2)m}{p\gamma} - \frac{(\alpha\gamma + \alpha^2)c + \alpha m}{p} - \frac{\alpha m(m + 2c\gamma)}{K_1 p^2 \gamma} - \frac{c\alpha^2 \gamma}{K_1 p^2}. \end{aligned} \quad (4.3)$$

In order to make the number of the population in protecting region Ω_2 reach the carrying capacity, let

$$e = K_2 - y, \quad (4.4)$$

where e is the error of y and K_2 . Here y is the number of population in Ω_2 , while K_2 is the carrying capacity of region Ω_2 .

Differentiating the formula (4.4) twice and considering the model (4.2), the following equation is obtained:

$$\frac{d^2 e}{dt^2} + a_1 \frac{de}{dt} + a_0 e = -b_0 u + (a_0 K_2 - \beta). \quad (4.5)$$

For the differential equation (4.5), $a_0 K_2 - \beta$ is considered as an external disturbance. According to the transformation, the model (4.1) is considered as a linear uncertain system with the control input. And then the model (4.5) is transformed into

$$\begin{aligned} \frac{de_1}{dt} &= e_2, \\ \frac{de_2}{dt} &= -a_0 e_1 - a_1 e_2 - b_0 u + a_0 K_2 - \beta, \end{aligned} \quad (4.6)$$

where $e_1 = e$.

The model (4.6) can be rewritten as a matrix form:

$$\frac{d\mathbf{w}}{dt} = \mathbf{A}\mathbf{w} + \mathbf{B}u + \mathbf{C}K_2 + \mathbf{D}, \quad (4.7)$$

where $\mathbf{w} = [e_1 \ e_2]^T$, $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -a_0 & -a_1 \end{pmatrix}$, $\mathbf{B} = [0 \ -b_0]^T$, $\mathbf{C} = [0 \ a_0]^T$, and $\mathbf{D} = [0 \ -\beta]^T$.

To stabilize the model (4.7), the variable structure controller is designed as

$$u = -(\lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 K_2) \text{Sgn}(\delta(\mathbf{w})), \quad (4.8)$$

where λ_i ($i = 1, 2, 3$) are switching coefficients and $\text{Sgn}(\delta(\mathbf{w}))$ is a sign function. $\delta(\mathbf{w})$ is called sliding surface, which divides the phase plane into two regions. The function $\delta(\mathbf{w})$ contains only endpoints of the trajectories of the model (4.7) coming from both sides of the surface and is defined as

$$\delta(\mathbf{w}) = f e_1 + e_2, \quad (4.9)$$

where $f > 0$ is a constant. To suppress the effect of the uncertainty and drive the trajectories of the model (4.7) toward the sliding surface until intersection occurs, the following reachable condition is established:

$$\delta(\mathbf{w}) \frac{d\delta(\mathbf{w})}{dt} < 0, \quad \text{for } \delta(\mathbf{w}) \neq 0. \quad (4.10)$$

That is

$$\frac{d\delta(\mathbf{w})}{dt} = f \frac{de_1}{dt} + \frac{de_2}{dt} = \begin{cases} (\lambda_1 - a_0)e_1 + (f + \lambda_2 - a_1)e_2 + (a_0 + \lambda_3)K_2 < 0, & \delta(\mathbf{w}) > 0 \\ (-\lambda_1 - a_0)e_1 + (f + \lambda_2 - a_1)e_2 + (a_0 - \lambda_3)K_2 > 0, & \delta(\mathbf{w}) < 0. \end{cases} \quad (4.11)$$

According to the reachable condition (4.10), we get the variable structure controller for the model (4.7):

$$u = \begin{cases} u^+ = -(\lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 K_2), & \delta(\mathbf{w}) > 0 \\ u^- = \lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 K_2, & \delta(\mathbf{w}) < 0. \end{cases} \quad (4.12)$$

Using the controller $u = u^+$ in the model (4.7), the controlled model is

$$\begin{bmatrix} \frac{de_1}{dt} \\ \frac{de_2}{dt} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \lambda_1 - a_0 & \lambda_2 - a_1 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} + \begin{bmatrix} 0 \\ a_0 + \lambda_3 \end{bmatrix} K_2 + \begin{bmatrix} 0 \\ -\beta \end{bmatrix}. \quad (4.13)$$

Let $e'_1 = e_1 + (\lambda_3 + a_0)K_2/(\lambda_1 - a_0)$ and $e'_2 = e_2$, then

$$\begin{bmatrix} \frac{de'_1}{dt} \\ \frac{de'_2}{dt} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \lambda_1 - a_0 & \lambda_2 - a_1 \end{bmatrix} \begin{bmatrix} e'_1 \\ e'_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -\beta \end{bmatrix}. \quad (4.14)$$

Obviously, the model (4.13) and the model (4.14) have the same state matrix $M = \begin{bmatrix} 0 & 1 \\ \lambda_1 - a_0 & \lambda_2 - a_1 \end{bmatrix}$. β is a bounded constant, and it does not influence the local stability of the controlled model. Thus, we have the following theorem.

Theorem 4.1. *If $a_0 - \lambda_1 > 0$, $a_1 - \lambda_2 > 0$, the model (4.7) can be stabilized by the controller u^+ .*

Proof. When the model (4.7) is controlled by the controller u^+ , it is transformed into the linear model (4.13). The characteristic equation of the state matrix M is

$$|\lambda E - M| = \lambda^2 - (\lambda_2 - a_1)\lambda - (\lambda_1 - a_0) = 0. \quad (4.15)$$

According to the Routh-Hurwitz criterion, if $a_0 - \lambda_1 > 0$, $a_1 - \lambda_2 > 0$, two eigenvalues for the state matrix M have negative real part. Therefore, the model (4.13) is locally stable. That is to say, the model (4.7) can be stabilized by the controller u^+ . \square

If $u = u^-$ in the model (4.7), the controlled model is

$$\begin{bmatrix} \frac{de_1}{dt} \\ \frac{de_2}{dt} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\lambda_1 - a_0 & -\lambda_2 - a_1 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} + \begin{bmatrix} 0 \\ a_0 - \lambda_3 \end{bmatrix} K_2 + \begin{bmatrix} 0 \\ -\beta \end{bmatrix}. \quad (4.16)$$

Let $e'_1 = e_1 + (\lambda_3 - a_0)K_2/(\lambda_1 + a_0)$ and $e'_2 = e_2$, then,

$$\begin{bmatrix} \frac{de'_1}{dt} \\ \frac{de'_2}{dt} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\lambda_1 - a_0 & -\lambda_2 - a_1 \end{bmatrix} \begin{bmatrix} e'_1 \\ e'_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -\beta \end{bmatrix}. \quad (4.17)$$

The model (4.16) and the model (4.17) also have the same state matrix $N = \begin{bmatrix} 0 & 1 \\ -\lambda_1 - a_0 & -\lambda_2 - a_1 \end{bmatrix}$. The transformation does not change the local stability of the model (4.16). Furthermore, we have another theorem.

Theorem 4.2. *If $\lambda_1 + a_0 > 0$, $\lambda_2 + a_1 > 0$, the model (4.7) can be stabilized by the controller u^- .*

Proof. When the model (4.7) is controlled by the controller u^- , it is transformed into the linear model (4.16). The characteristic equation of the state matrix N is

$$|\lambda E - N| = \lambda^2 - (\lambda_2 + a_1)\lambda - (\lambda_1 + a_0) = 0. \quad (4.18)$$

According to the Routh-Hurwitz criterion, if $\lambda_1 + a_0 > 0$, $\lambda_2 + a_1 > 0$, two eigenvalues for the state matrix N have negative real part. The model (4.16) is locally stable, and the model (4.7) can be stabilized by the controller u^- . \square

From the condition (4.11), Theorems 4.1 and 4.2, we get the varying range of the switching coefficients λ_i ($i = 1, 2, 3$):

$$\lambda_1 \geq \max|-a_0|, \quad \lambda_2 \geq \max|f - a_1|, \quad \lambda_3 \geq \max|a_0|. \quad (4.19)$$

According to the condition $\delta(\mathbf{w}) = 0$ and $d\delta(\mathbf{w})/dt = 0$, the equivalent control on the sliding surface $\delta(\mathbf{w}) = 0$ can be obtained. If $\delta(\mathbf{w}) = 0$, there is a state variable represented by the remaining state variables. From the condition $\text{Sgn}(\delta(\mathbf{w})) = 0$, we have

$$\frac{de_1}{dt} = -fe_1. \quad (4.20)$$

Remark 4.3. When applying variable structure control, the singular model is transformed into a linear model with parameters varying within definite intervals. Since the sliding surface can be designed as required and has nothing to do with the parameters and disturbance, it makes the discontinuous control insensitive to internal parameter variations and extraneous disturbance and decreases the chattering phenomenon. Variable structure control can stabilize the nonlinear system effectively.

5. Simulations

Fishery production is an important aspect in human life. In order to guarantee the sustainable development of the fishery, people have taken many necessary measures. Therefore, to study the structure model for the inshore-offshore fishery is necessary. It is a good idea to divide the population into two categories in keeping resources sustainable, a harvesting-permitted category and a harvesting-forbidden category. Some inshore-offshore models in an aquatic environment have ever been studied to keep the fishery sustainable [40–42]. But these papers did not consider the economic profits that the fishery brings for people. In this paper, the sustainable fishery and the economic interest are discussed for the inshore-offshore fishery model.

The sea around Zhoushan is a famous fishing ground in Zhejiang province. The total sea area is about more than 10800 km². The area of the inshore region is about 3700 km², and the offshore region is about 7100 km² [43]. The coiliasspp is a kind of fish, and it is about 1099 million in the whole sea area [44]. To protect the fishery resources, the coiliasspp in the inshore region is permitted to be harvested, while the offshore region is forbidden. In the inshore region, the density of the coiliasspp is greater than that in the offshore region because of the environment effect. So the environment carrying capacity of the inshore region is about 423 million, and the offshore environment carrying capacity is 676 million. The intrinsic growth rate γ is assumed to be 0.2. When the number of the fish in two regions are different, they migrate between two regions at the proportional $\alpha = 0.6$. It is supposed that they are sold at the average unit price $p = 11$, and its unit cost c is 6.

Considering these conditions, the following singular ecological-economic model can be established:

$$\begin{aligned}\frac{dx_1}{dt} &= 0.2x_1\left(1 - \frac{x_1}{423}\right) - 0.6(x_1 - x_2) - Ex_1, \\ \frac{dx_2}{dt} &= 0.2x_2\left(1 - \frac{x_2}{676}\right) + 0.6(x_1 - x_2), \\ 0 &= (11x_1 - 6)E - m.\end{aligned}\tag{5.1}$$

When the economic profit m varies, there are some complex dynamic behaviors for the model (5.1), such as the singular induced bifurcation. When the economic profit $m = 0$, the model (5.1) has a positive equilibrium point $p^*(0.545, 0.818, 0.499)$. When economic profit $m = 0.001$, there are two eigenvalues for the matrix $J = D_x F - D_E F(D_E G)^{-1}D_x G$, -1.2998 and -0.0002 . The eigenvalues became -1.2998 and 0.0017 when the parameter $m = -0.001$. It obvious that one eigenvalue remains constant, and the other eigenvalue moves from C^- to C^+ along the real axis by diverging through ∞ . It is called the overexploitation, and it causes the extinction of the coiliassp. In order to avoid such phenomena, a variable structure controller is designed to make the coiliassp in the offshore region reach the environment carrying capacity 676 million. According to the varying range of the switching coefficients (4.19), the variable structure controller is designed as follows:

$$u = -(10e_1 + 76e_2 + 15800)\text{Sgn}(\delta(\mathbf{w})),\tag{5.2}$$

where the sliding surface is chosen as $\delta(\mathbf{w}) = 9.6e_1 + e_2$. By controlling the harvesting effort E , x_2 reaches the environment carrying capacity 676 million in Ω_2 . Figure 2 shows the control result of x_1 , x_2 , and E with variable structure control.

In Figure 2, when the harvesting effort E is controlled at 0.16 million, the number of fish in region Ω_2 reaches 675.2 million controlled by the controller u . Due to the migration between Ω_1 and Ω_2 , the coiliassp in inshore region reaches 424.3 million accordingly. The state variables stay in a stable situation, and the singular induced bifurcation is eliminated by the controller u . In practical, we can regulate the harvesting behavior by the revenue to keep the harvesting and the reproduction in balance. Therefore, the sustainable development of the fishery can be realized by this controller. Further, we know that the corresponding nonlinear singular ecological-economic model can be stabilized by variable structure control.

6. Conclusions

In this paper, the population is divided into the harvesting region and the protecting region, in which the population can migrate between two regions. In harvesting the population resources, when the economic interest and the environmental effects are taken into account, a singular ecological-economic model is established. The local stability and the dynamic behavior for this model are discussed. As the parameters changing, the singular model undergoes the singular induced bifurcation. In order to apply variable structure control to eliminate this complex behavior, the singular model is transformed into a linear single-input and single-output model with parameters varying within definite intervals. Variable

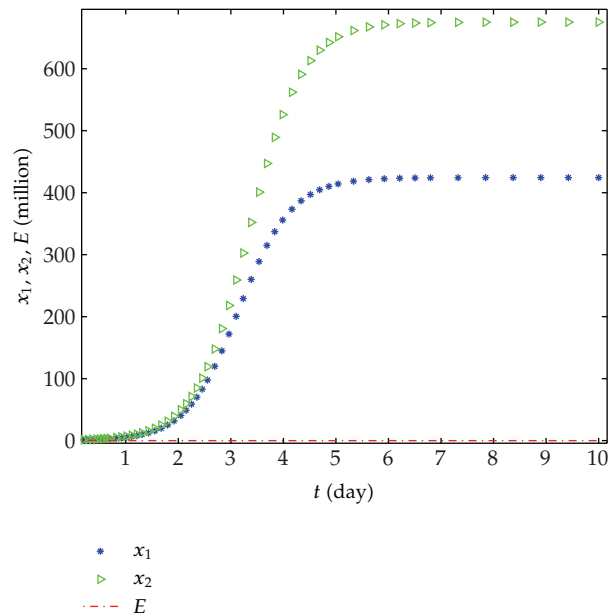


Figure 2: The state response of x_1 , x_2 , and E when $m = 1325$ with the controller u .

structure control with sliding mode is designed to stabilize the model. An inshore-offshore fishery model illustrates the analysis result. Some simulations show the effectiveness of the control method.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (60974004).

References

- [1] W. G. Marszalek and Z. W. Trzaska, "Singularity-induced bifurcations in electrical power systems," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 312–320, 2005.
- [2] M. Yue and R. Schlueter, "Bifurcation subsystem and its application in power system analysis," *IEEE Transactions on Power Systems*, vol. 19, no. 4, pp. 1885–1893, 2004.
- [3] T. Y. Guo and R. A. Schlueter, "Identification of generic bifurcation and stability problems in power system differential-algebraic model," *IEEE Transactions on Power Systems*, vol. 9, no. 2, pp. 1032–1038, 1994.
- [4] R. Riaza and P. J. Zufiria, "Differential-algebraic equations and singular perturbation methods in recurrent neural learning," *Dynamical Systems*, vol. 18, no. 1, pp. 89–105, 2003.
- [5] Z. W. Gao and S. X. Ding, "State and disturbance estimator for time-delay systems with application to fault estimation and signal compensation," *IEEE Transactions on Signal Processing*, vol. 55, no. 12, pp. 5541–5551, 2007.
- [6] Z. Gao and D. W. C. Ho, "State/noise estimator for descriptor systems with application to sensor fault diagnosis," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1316–1326, 2006.
- [7] R. J. Spiteri, U. M. Ascher, and D. K. Pai, "Numerical solution of differential systems with algebraic inequalities arising in robot programming," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 2373–2380, May 1995.

- [8] R. J. Spiteri, D. K. Pai, and U. M. Ascher, "Programming and control of robots by means of differential algebraic inequalities," *IEEE Transactions on Robotics and Automation*, vol. 16, no. 2, pp. 135–145, 2000.
- [9] W. M. Liu, H. W. Hethcote, and S. A. Levin, "Dynamical behavior of epidemiological models with nonlinear incidence rates," *Journal of Mathematical Biology*, vol. 25, no. 4, pp. 359–380, 1987.
- [10] N. Yi, Z. Zhao, and Q. Zhang, "Bifurcations of an SEIQS epidemic model," *International Journal of Information & Systems Sciences*, vol. 5, no. 3-4, pp. 296–310, 2009.
- [11] N. Yi, Q. Zhang, K. Mao, D. Yang, and Q. Li, "Analysis and control of an SEIR epidemic system with nonlinear transmission rate," *Mathematical and Computer Modelling*, vol. 50, no. 9-10, pp. 1498–1513, 2009.
- [12] D. G. Luenberger and A. Arbel, "Singular dynamic Leontief system," *Econometrics*, vol. 45, no. 32, pp. 991–995, 1997.
- [13] J. S. Zhang, *Singular System Economic Cybernetics*, Tsinghua University Press, Beijing, China, 1990.
- [14] A. Kumar and P. Daoutidis, *Control of Nonlinear Differential Algebraic Equation Systems*, vol. 397 of *Chapman & Hall/CRC Research Notes in Mathematics*, Chapman & Hall/CRC, Boca Raton, Fla, USA, 1999.
- [15] Y. Zhang, Q. L. Zhang, and L. C. Zhao, "Bifurcations and control in singular biological economic model with stage structure," *Journal of Systems Engineering*, vol. 22, no. 3, pp. 233–238, 2007.
- [16] C. Liu, Q. L. Zhang, Y. Zhang, and X. D. Duan, "Bifurcation and control in a differential-algebraic harvested prey-predator model with stage structure for predator," *International Journal of Bifurcation and Chaos*, vol. 18, no. 10, pp. 3159–3168, 2008.
- [17] Y. Zhang and Q. Zhang, "Optimal control of logistic bioeconomic model with singularity-induced bifurcation," *International Journal of Information & Systems Sciences*, vol. 5, no. 3-4, pp. 369–379, 2009.
- [18] C. Liu, Q. Zhang, and X. Zhang, "Dynamic analysis in a harvested differential-algebraic prey-predator model," *Journal of Mechanics in Medicine and Biology*, vol. 9, no. 1, pp. 123–140, 2009.
- [19] X. Zhang, Q.-l. Zhang, and Y. Zhang, "Bifurcations of a class of singular biological economic models," *Chaos, Solitons & Fractals*, vol. 40, no. 3, pp. 1309–1318, 2009.
- [20] H. S. Gordon, "The economic theory of a common property resource: the Fishery," *Journal of Political Economy*, vol. 62, pp. 124–142, 1954.
- [21] W. J. Reed, "Recruitment variability and age structure in harvested animal populations," *Mathematical Biosciences*, vol. 65, no. 2, pp. 239–268, 1983.
- [22] T. K. Kar, "Modelling and analysis of a harvested prey-predator system incorporating a prey refuge," *Journal of Computational and Applied Mathematics*, vol. 185, no. 1, pp. 19–33, 2006.
- [23] L. Liu and J. Sun, "Existence of periodic solution for a harvested system with impulses at variable times," *Physics Letters, Section A*, vol. 360, no. 1, pp. 105–108, 2006.
- [24] M. Dieter, "Analysis of trade in illegally harvested timber: accounting for trade via third party countries," *Forest Policy and Economics*, vol. 11, no. 8, pp. 600–607, 2009.
- [25] S. V. Emelyanov, *Variable Structure Control Systems*, Nauka, Moscow, Russia, 1967.
- [26] Y. Itkis, *Control Systems of Variable Structure*, John Wiley & Sons, New York, NY, USA, 1976.
- [27] V. I. Utkin, *Sliding Modes and Their Application in Variable Structure Systems*, Nauka, Moscow, Russia, 1978.
- [28] R. Suarez-Cortez, J. Alvarez-Gallegns, and E. Gonzalez-Mora, "Sliding controller design for a non-linear fermentation system," *Biotechnology and Bioengineering*, vol. 33, no. 4, pp. 377–385, 1989.
- [29] A. S. Zinober, *Deterministic Control of Uncertain Systems*, vol. 40 of *IEE Control Engineering Series*, Peter Peregrinus, London, UK, 1990.
- [30] V. I. Utkin, *Sliding Modes in Control and Optimization*, Communications and Control Engineering Series, Springer, Berlin, Germany, 1992.
- [31] P. Park, D. J. Choi, and S. G. Kong, "Output feedback variable structure control for linear systems with uncertainties and disturbances," *Automatica*, vol. 43, no. 1, pp. 72–79, 2007.
- [32] C.-C. Wang and J.-P. Su, "A new adaptive variable structure control for chaotic synchronization and secure communication," *Chaos, Solitons & Fractals*, vol. 20, no. 5, pp. 967–977, 2004.
- [33] F. Juan and L. Gerard, "Variable structure control for power systems stabilization," *International Journal of Electrical Power and Energy Systems*, vol. 32, no. 2, pp. 101–107, 2010.
- [34] C.-C. Wang and J.-P. Su, "A new adaptive variable structure control for chaotic synchronization and secure communication," *Chaos, Solitons & Fractals*, vol. 20, no. 5, pp. 967–977, 2004.
- [35] S. H. Ling, F. H. F. Leung, H. K. Lam, and H. H. C. Iu, "Variable-structure neural network with real-coded genetic algorithm and its application on short-term load forecasting," *International Journal of Information and Systems Sciences*, vol. 5, no. 1, pp. 23–40, 2009.

- [36] J. Fernández-Vargas and G. Ledwich, "Variable structure control for power systems stabilization," *International Journal of Electrical Power and Energy Systems*, vol. 32, no. 2, pp. 101–107, 2010.
- [37] B. S. Chen, X. X. Liao, and Y. Q. Liu, "Normal forms and bifurcations for differential-algebraic systems," *Acta Mathematicae Applicatae Sinica*, vol. 23, no. 3, pp. 429–443, 2000 (Chinese).
- [38] V. Venkatasubramanian, H. Schättler, and J. Zaborszky, "Local bifurcations and feasibility regions in differential-algebraic systems," *IEEE Transactions on Automatic Control*, vol. 40, no. 12, pp. 1992–2013, 1995.
- [39] P. Zlateva, "Variable-structure control of nonlinear systems," *Control Engineering Practice*, vol. 4, no. 7, pp. 1023–1028, 1996.
- [40] T. Pradhan and K. S. Chaudhuri, "Bioeconomic modelling of selective harvesting in an inshore-offshore fishery," *Differential Equations and Dynamical Systems*, vol. 7, no. 3, pp. 305–320, 1999.
- [41] B. Dubey, P. Chandra, and P. Sinha, "A model for fishery resource with reserve area," *Nonlinear Analysis: Real World Applications*, vol. 4, no. 4, pp. 625–637, 2003.
- [42] M. Jerry, N. Raissi, and A. Rapaport, "A viability analysis for an explicit inshore-offshore model," *Journal of Applied Mathematics*, vol. 1, no. 1, pp. 41–60, 2011.
- [43] <http://wiki.cnki.com.cn/HotWord/7044734.htm>.
- [44] <http://info.china.alibaba.com/news/detail/v0-d1023779518.html>.

Research Article

A Matrix Method for Determining Eigenvalues and Stability of Singular Neutral Delay-Differential Systems

Jian Ma,^{1,2} Baodong Zheng,¹ and Chunrui Zhang²

¹ Department of Mathematics, Harbin Institute of Technology, Harbin 150001, China

² Department of Mathematics, Northeast Forestry University, Harbin 150040, China

Correspondence should be addressed to Baodong Zheng, zbd@hit.edu.cn

Received 29 November 2011; Revised 6 April 2012; Accepted 8 April 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 Jian Ma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The eigenvalues and the stability of a singular neutral differential system with single delay are considered. Firstly, by applying the matrix pencil and the linear operator methods, new algebraic criteria for the imaginary axis eigenvalue are derived. Second, practical checkable criteria for the asymptotic stability are introduced.

1. Introduction

Nowadays, the time-delay systems have become an important natural models in physics, engineering, multibody mechanics, computer-aided design, and economic systems. The theory on ordinary differential equations with delays have been discussed for decades in a wide range, so there are very many results for them. Especially, the eigenvalues and the stability analysis of time-delay systems have received much attention of researchers and many excellent results have been obtained, see [1–6]. Certainly most of them had been focused on the analytical methods or numerical methods, such as V-functional methods, Laplace transformation, Runge-Kutta methods, and linear multistep methods. In [7–9], the numerical techniques for the computation of the eigenvalues were discussed. In [10], Zhu and Petzold researched the asymptotic stability of delay-differential-algebraic equations by applying the θ -methods, Runge-Kutta methods, and linear multistep methods. These methods play the key roles at last. But in recent years, algebraic methods are developing fast, especially for the research on the more complex systems, such as the n -dimensional systems. Though the algebraic methods as a new and effective tool is also applied to analyze the time-delay systems [2, 3], the results are very few.

In this paper, we will discuss the differential-algebraic equations by the algebraic methods. Their dynamics have not been well understood yet.

Example 1.1. Consider the simple differential-algebraic system:

$$\begin{aligned}\dot{x}_1(t) &= f_1(t), \\ x_1(t) - x_2(t - \tau) &= f_2(t),\end{aligned}\tag{1.1}$$

where $t \geq 0$, $\tau \geq 0$, and x_1 and x_2 are given by continuous functions on the initial interval $(-\tau, 0]$. So we have the solution:

$$\begin{aligned}x_1(t) &= \int_0^t f_1(s)ds + c, \\ x_2(t) &= -f_2(t + \tau) + \int_0^{t+\tau} f_1(s)ds + c,\end{aligned}\tag{1.2}$$

where c is a constant. From the solution, we find that the solution depends on future integrals of the input $f(t)$. This interesting phenomenon arrested many scholars to research. For a general n -dimensional differential equations with delay, we can note by

$$B\dot{X}(t) = A_0X(t) + A_1X(t - \tau),\tag{1.3}$$

where $B, A_0, A_1 \in R^{n \times n}$, $\text{Rank } B \leq n$, $t \geq 0$, $\tau \geq 0$, and $X(t) \in R^n$ is given by continuous functions on the initial interval $(-\tau, 0]$. When $\text{Rank } B = n$, we called it the retarded differential equations. It can be improved as

$$\dot{X}(t) = A_0X(t) + A_1X(t - \tau).\tag{1.4}$$

Many scholars have widely researched the delay-independent or delay-dependent stability and asymptotic stability by analytic methods or numerical methods. When $\text{Rank } B < n$, it is called a singular (or degenerated) delay-differential equations. The imaginary axis eigenvalues are discussed by using matrix pencil, see [2]. But because of the complex nature of the singular differential systems with delay, the research is very difficult by using the analytical treatment. So few studies on the stability and the bifurcations have been conducted so far. Particular, for the singular neutral differential systems with delays, there are hardly flexible and efficient verdicts.

In this paper, we will apply the algebraic methods to discuss the stability of a singular neutral differential system with a single delay, as follows:

$$B_0\dot{x}(t) + B_1\dot{x}(t - \tau) = A_0x(t) + A_1x(t - \tau),\tag{1.5}$$

where $B_0, B_1, A_0, A_1 \in R^{n \times n}$, $x \in R^n$, $\tau \geq 0$. For the system (1.5), if $\det B_0 \neq 0$, we can improve it as the form

$$\dot{x}(t) + B_1\dot{x}(t - \tau) = A_0x(t) + A_1x(t - \tau).\tag{1.6}$$

It is a neutral differential equation with a single delay. The problem of computing imaginary axis eigenvalues on the system (1.6) has been previously studied in [11]. Here, we consider the state rank $B_0 \leq n$. The solvability of the system (1.5), which is essentially the existence and the uniqueness of the solution, is determined by the regularity. The matrix pencil (B_0, A_0) is said to be regular if $sB_0 + A_0$ is not identically singular for any complex s . If (B_0, A_0) is regular, the zero s of $\det(sB_0 + A_0)$ is called the eigenvalue of the matrix pencil (B_0, A_0) . From [12], we know that the system (1.5) is solvable if and only if (B_0, A_0) is regular. So, in this paper, we suppose that (B_0, A_0) is regular. In the following, we will analyze the eigenvalues and the stability of the system (1.5).

2. The Algebraic Criteria for Determining Imaginary Axis Eigenvalues

Firstly, we research an ordinary differential equation, which will motivate our analysis. Consider

$$\begin{aligned} B_0 \dot{X}(t) + B_1 \dot{Y}(t) &= A_0 X(t) + A_1 Y(t), \\ \dot{X}(t) B_1^T + \dot{Y}(t) B_0^T &= -X(t) A_1^T - Y(t) A_0^T, \end{aligned} \quad (2.1)$$

where $B_0, B_1, A_0, A_1 \in \mathbb{R}^{n \times n}$, $X, Y \in \mathbb{C}^{n \times n}$. Let V denote the vector space $V = \mathbb{C}^{n \times n} \times \mathbb{C}^{n \times n}$ and E, F denote the operators on V , given by

$$E \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} B_0 X + B_1 Y \\ X B_1^T + Y B_0^T \end{pmatrix}, \quad F \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} A_0 X + A_1 Y \\ -X A_1^T - Y A_0^T \end{pmatrix}, \quad \forall X, Y \in \mathbb{C}^{n \times n}. \quad (2.2)$$

With $Z(t) = \begin{pmatrix} X(t) \\ Y(t) \end{pmatrix}$, the system (2.1) can be written as

$$E \dot{Z}(t) = F Z(t). \quad (2.3)$$

Supposing $\tilde{Z}(t) = \begin{pmatrix} \tilde{X} \\ \tilde{Y} \end{pmatrix} = \begin{pmatrix} X_0 e^{st} \\ Y_0 e^{st} \end{pmatrix}$ is a matrix solution of the system (2.1), we have

$$\begin{aligned} (sB_0 - A_0) \tilde{X} + (sB_1 - A_1) \tilde{Y} &= 0, \\ \tilde{X} (sB_1^T + A_1^T) + \tilde{Y} (sB_0^T + A_0^T) &= 0. \end{aligned} \quad (2.4)$$

For any complex s , let $T = T(s)$ be the operator $T = sE - F$; then

$$T \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} (sB_0 - A_0)X + (sB_1 - A_1)Y \\ X(sB_1^T + A_1^T) + Y(sB_0^T + A_0^T) \end{pmatrix}, \quad \forall X, Y \in \mathbb{C}^{n \times n}. \quad (2.5)$$

For any complex s , let $T^+ = T^+(s) : V \rightarrow V$, satisfying

$$T^+ \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} X(sB_0^T + A_0^T) - (sB_1 - A_1)Y \\ -X(sB_1^T + A_1^T) + (sB_0 - A_0)Y \end{pmatrix}, \quad \forall X, Y \in \mathbb{C}^{n \times n}. \quad (2.6)$$

For any complex s , let $\Lambda = \Lambda(s)$, satisfying

$$\Lambda X = (sB_0 - A_0)X(sB_0^T + A_0^T) - (sB_1 - A_1)X(sB_1^T + A_1^T), \quad \forall X \in \mathbb{C}^{n \times n}. \quad (2.7)$$

By simple computations, we can get

$$\begin{aligned} (sB_0 - A_0)X(sB_0^T + A_0^T) - (sB_1 - A_1)X(sB_1^T + A_1^T) &= X_0(sB_0^T + A_0^T) - (sB_1 - A_1)Y_0, \\ (sB_0 - A_0)Y(sB_0^T + A_0^T) - (sB_1 - A_1)Y(sB_1^T + A_1^T) &= -X_0(sB_1^T + A_1^T) + (sB_0 - A_0)Y_0. \end{aligned} \quad (2.8)$$

Expressing by the operator language, that is,

$$\mathbf{T}^+ \mathbf{T} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \Lambda X \\ \Lambda Y \end{pmatrix}. \quad (2.9)$$

In the following, we will convert matrix ordinary differential equation (2.3) to vector form. Let ξ be the elementary transform, $\xi : \mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{n^2}$, that is,

$$\xi A = \begin{pmatrix} a_1^T \\ \vdots \\ a_n^T \end{pmatrix}, \quad \forall A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad a_i^T \in \mathbb{C}^n, \quad i = 1, 2, \dots, n. \quad (2.10)$$

Let $x = \xi X, y = \xi Y, z = \begin{pmatrix} x \\ y \end{pmatrix}$, and

$$E_0 = \begin{pmatrix} B_0 \otimes I & B_1 \otimes I \\ I \otimes B_1 & I \otimes B_0 \end{pmatrix}, \quad F_0 = \begin{pmatrix} A_0 \otimes I & A_1 \otimes I \\ -I \otimes A_1 & -I \otimes A_0 \end{pmatrix}. \quad (2.11)$$

Using the property of the Kronecker product, we have

$$\xi(AXB) = A \otimes B^T \xi X, \quad (2.12)$$

where $A, B, X \in \mathbb{C}^{n \times n}$. So (2.3) can be written as

$$E_0 \dot{z}(t) = F_0 z(t). \quad (2.13)$$

Similarly, by denoting $T_0 = T_0(s)$, $T_0^+ = T_0^+(s)$, $\Lambda_0 = \Lambda_0(s)$ as follows

$$\begin{aligned} T_0 &= T_0(s) = \begin{pmatrix} (sB_0 - A_0) \otimes I & (sB_1 - A_1) \otimes I \\ I \otimes (sB_1 + A_1) & I \otimes (sB_0 + A_0) \end{pmatrix} = sE_0 - F_0, \\ T_0^+ &= T_0^+(s) = \begin{pmatrix} I \otimes (sB_0 + A_0) & -(sB_1 - A_1) \otimes I \\ -I \otimes (sB_1 + A_1) & (sB_0 - A_0) \otimes I \end{pmatrix}, \\ \Lambda_0 &= \Lambda_0(s) = (sB_0 - A_0) \otimes (sB_0 + A_0) - (sB_1 - A_1) \otimes (sB_1 + A_1), \end{aligned} \quad (2.14)$$

we have

$$T_0^+ T_0 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \Lambda_0 x \\ \Lambda_0 y \end{pmatrix}, \text{ that is, } T_0^+ T_0 = \begin{pmatrix} \Lambda_0 & 0 \\ 0 & \Lambda_0 \end{pmatrix}. \quad (2.15)$$

Lemma 2.1. For all complex s , $\det T_0^+(s) = \det T_0(s)$, and so $\det T_0(s) = \pm \det \Lambda_0(s)$.

Proof. Let $T_0 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where $a = (sB_0 - A_0) \otimes I$, $b = (sB_1 - A_1) \otimes I$, $c = I \otimes (sB_1 + A_1)$, and $d = I \otimes (sB_0 + A_0)$. Noting the Kronecker product identities $(I \otimes M)(N \otimes I) = (N \otimes I)(I \otimes M) = N \otimes M$, we have $db = bd$, $bc = cb$. By regularity of (B_0, A_0) , we know $sB_0 + A_0$ is nonsingular for enough complex s . So by the property of the polynomial and easy computations, we can get that $\det T_0^+(s) = \det T_0(s)$ for any complex s . Elsewhere,

$$T_0^+ T_0 = \begin{pmatrix} \Lambda_0 & 0 \\ 0 & \Lambda_0 \end{pmatrix}, \quad (2.16)$$

so

$$(\det T_0)^2 = \det T_0^+ \det T_0 = (\det \Lambda_0)^2. \quad (2.17)$$

Thus

$$\det T_0(s) = \pm \det \Lambda_0(s). \quad (2.18)$$

□

Theorem 2.2. Any imaginary axis eigenvalue of the systems (1.5) is a zero point of $\det \Lambda_0(s)$ and thus also one of the eigenvalues of the matrix pencil (E_0, F_0) .

Proof. The matrix polynomial for the system (1.5) is

$$p(s, e^{-s\tau}) = sB_0 - A_0 + (sB_1 - A_1)e^{-s\tau}. \quad (2.19)$$

Let $s = i\omega$ be an imaginary axis eigenvalue of the system (1.5) and v is associated eigenvector, $\|v\| = 1$. We have $p(s, e^{-s\tau})v = 0$. By conjugating and transforming, we can get

$$(sB_0 - A_0)vv^* (sB_0^T + A_0^T) - (sB_1 - A_1)vv^* (sB_1^T + A_1^T) = 0. \quad (2.20)$$

Via the elementary transform ξ , we get

$$[(sB_0 - A_0) \otimes (sB_0 + A_0) - (sB_1 - A_1) \otimes (sB_1 + A_1)]\xi(vv^*) = 0, \quad (2.21)$$

that is, $\Lambda_0(s)u = 0$, $u = \xi(vv^*)$. We know that $\det \Lambda_0(s) = 0$, and so

$$\det(sE_0 - F_0) = \pm \det \Lambda_0(s) = 0. \quad (2.22)$$

From Theorem 2.2, we know that all of the imaginary axis eigenvalues of the system (1.5) are zero points of the algebraic equation

$$\det[(sB_0 - A_0) \otimes (sB_0 + A_0) - (sB_1 - A_1) \otimes (sB_1 + A_1)] = 0. \quad (2.23)$$

□

Corollary 2.3. *If $\det(B_0 \otimes B_0 - B_1 \otimes B_1) \neq 0$, then E_0 is invertible, and any imaginary axis eigenvalue of the systems (1.5) is the eigenvalue of $F_0 E_0^{-1}$.*

Proof. By proof of Lemma 2.1, we have

$$\det E_0 = \begin{vmatrix} B_0 \otimes I & B_1 \otimes I \\ I \otimes B_1 & I \otimes B_0 \end{vmatrix} = \det(B_0 \otimes B_0 - B_1 \otimes B_1). \quad (2.24)$$

The corollary follows immediately from Theorem 2.2. □

Corollary 2.4. *Any imaginary axis eigenvalue of the system with single delay*

$$\dot{x}(t) = A_0 x(t) + A_1 x(t - \tau) \quad (2.25)$$

is an eigenvalue of $F_0 = \begin{pmatrix} A_0 \otimes I & A_1 \otimes I \\ -I \otimes A_1 & -I \otimes A_0 \end{pmatrix}$.

Proof. It follows immediately from Theorem 2.2. □

In fact, the above result contains the system (1.6). For the system (1.6), we also have the following corollary.

Corollary 2.5 (see Jarlebring and Hochstenbach [2, Theorem 1]). *For the system (1.6), the imaginary axis eigenvalues are the roots of the equation*

$$\det[(sI - A_0) \otimes (sI + A_0) - (sB_1 - A_1) \otimes (sB_1 + A_1)] = 0. \quad (2.26)$$

Remark 2.6. In fact, (2.23) or (2.26) is usually called a polynomial eigenvalue problem. The classical and most widely used approach to research the polynomial eigenvalue problems is linearization, where the polynomial is converted into a larger matrix pencil with the same eigenvalues. There are many forms for the linearization: the companion form is most typically commission. The linearization method is also an important tool to research the characteristic equations in algebraic methods, see [2, 13].

From the above results, we find that the imaginary axis eigenvalues of the system (1.5) or (1.6) can be computed via the algebraic equation (2.23) and (2.26). The imaginary eigenvalues play an important role in the stability. Next, we will use the results to discuss the stability of the systems. At first, we will give the condition of the delay-independent stability on the system (1.5). Secondly, we will address the problem of finding the critical delays of the system (1.6), that is, the delay such that the system (1.6) has purely imaginary eigenvalues.

3. The Algebraic Criteria of the Asymptotic Stability

The stability of the delay ordinary differential equations has been widely discussed [14, 15]. It is well known that the Lyapunov-Krasovskii functional approach is the important analytic method to find the delay-independent stability criteria, which do not include any information on the size of delay. The main ideas for developing algebraic criteria of the stability analysis on the systems can be found in many works, such as “The Degenerate Differential Systems with Delay” (W. Jiang, 1998, [12]). The results for singular neutral differential equations are still very few, especially by algebraic methods.

Next we first research the delay-independent stability of the system (1.5). The characteristic equation for the system(1.5) is denoted again by

$$P(s, z) = \det[s(B_0 + B_1 z) - (A_0 + A_1 z)], \quad \forall s \in \mathbb{C}, z = e^{-s\tau}. \quad (3.1)$$

From [12], we known that the solution of the neutral time-delay systems is asymptotically stable if all roots of (3.1) have negative real part bounded away from 0, that is, there exists a number $\delta > 0$, such that $\operatorname{Re}(s) \leq -\delta < 0$ for any root s of (3.1). Especially for the system (1.5), Zhu and Petzold had found that there must exist the δ , if the condition $|u^T B_0 u| \geq |u^T B_1 u|$ for all $u \in \mathbb{R}^n$ holds, see [10]. So from [3, 10], we can get the following theorem.

Theorem 3.1. *Let s be the zeros of (2.23). If the coefficient matrices of the system (1.5) satisfies the following conditions:*

- (i) $\operatorname{Re} \lambda < 0, \lambda \in \sigma(B_0, -A_0),$
- (ii) $\max_{\operatorname{Re} s=0} \rho((sB_0 - A_0)^{-1}(sB_1 - A_1)) < 1,$

then the system (1.5) is asymptotically stable for all $\tau \geq 0$, that is, the stability of the systems (1.5) is delay independent.

Proof. It follows immediately from [10].

In the following, we consider the neutral system (1.6), whose characteristic equation is denoted by

$$P(s, z) = \det[s(I + B_1 z) - (A_0 + A_1 z)], \quad \forall s \in \mathbb{C}, z = e^{-s\tau}. \quad (3.2)$$

It is well known that the spectrum of the neutral delay systems exhibits some discontinuity properties, that is to say, an infinitesimal change of the delay parameter may cause the stability of the system to shift. These discontinuity properties are closely related to the essential spectrum of the system. The critical condition for a stability switch of a neutral delay system is that the rightmost eigenvalue goes from the left complex half-plane into the right complex half-plane by passing the imaginary axis. So the appearance of the imaginary axis eigenvalue is the critical condition. In Section 2, we find all of the imaginary axis eigenvalues. Let the delay be a parameter. In the following, we will find the critical value of the delay parameter such that the stability switch occurs. It is known that, if a neutral delay system is

stable, it is necessary that its neutral part must be stable. For the system (1.6), this requirement concerns the stability of the difference equation:

$$x(t) + B_1 x(t - \tau) = 0. \quad (3.3)$$

The eigenvalues of (3.3) are called the essential spectrum of the system (1.6). We know that (3.3) is stable if $\rho(B_1) < 1$. It is important to point out that, under this assumption, the condition $\operatorname{Re}(s) \leq -\delta < 0$ can be improved to $\operatorname{Re}(s) < 0$. Our task is to find the critical delay where the system (1.6) becomes unstable. So we have the following theorem. \square

Theorem 3.2. *Supposing all of the eigenvalues of matrix A_0 have negative real part and $\rho(B_1) < 1$, one has the following.*

- (i) *If for any root s of (3.2), $\operatorname{Re}(s) < 0$, else if for any root $s = i\omega$, $\omega > 0$, $z = e^{-s\tau}$, $|z| < 1$, then system (1.6) is asymptotic stability for any $\tau \geq 0$, that is, stability is delay independent.*
- (ii) *Otherwise, for each root $s = i\omega$, $\omega > 0$, $|z| = 1$, one can get the minimal critical value of delay parameter τ^* , such that if $\tau \in [0, \tau^*)$, then system (1.6) is asymptotic stability, and when $\tau \geq \tau^*$, the stability of system (1.6) changes, that is, the system (1.6) is delay-dependent stable and generates bifurcation in $\tau = \tau^*$.*

Example 3.3. Consider a neutral neural networks with a single delay, see [16]:

$$\begin{aligned} \dot{x}_1(t) &= -x_1(t) + af(x_1(t - \tau)) + bg(x_2(t - \tau)) + bg(x_3(t - \tau)) \\ &\quad + a_2 f_2(\dot{x}_1(t - \tau)) + b_2 g_2(\dot{x}_2(t - \tau)) + b_2 g_2(\dot{x}_3(t - \tau)), \\ \dot{x}_2(t) &= -x_2(t) + bg(x_1(t - \tau)) + af(x_2(t - \tau)) + bg(x_3(t - \tau)) \\ &\quad + b_2 g_2(\dot{x}_1(t - \tau)) + a_2 f_2(\dot{x}_2(t - \tau)) + b_2 g_2(\dot{x}_3(t - \tau)), \\ \dot{x}_3(t) &= -x_3(t) + bg(x_2(t - \tau)) + af(x_3(t - \tau)) \\ &\quad + b_2 g_2(\dot{x}_2(t - \tau)) + a_2 f_2(\dot{x}_3(t - \tau)). \end{aligned} \quad (3.4)$$

Now we rewrite system (3.4) as the matrix equations:

$$\dot{X}(t) + B\dot{X}(t - \tau) = AX(t) + C_1 f(X(t - \tau)) + C_2 g(X(t - \tau)). \quad (3.5)$$

The linearization of the system (3.4) around the origin is given by

$$\dot{X}(t) + B_1 \dot{X}(t - \tau) = A_0 X(t) + A_1 X(t - \tau), \quad (3.6)$$

where

$$A_0 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad A_1 = \begin{pmatrix} a & b & b \\ b & a & b \\ 0 & b & a \end{pmatrix}, \quad B_1 = \begin{pmatrix} a_2 & b_2 & b_2 \\ b_2 & a_2 & b_2 \\ 0 & b_2 & a_2 \end{pmatrix}. \quad (3.7)$$

Assume that $a = -2$, $b = -1.5$, $a_2 = 0.3$, $b_2 = -0.3$, and $f(x) = g(x) = \tanh(x)$. We carry out the numerical simulations for system (3.4). From (2.26), by MATLAB computation, we can get

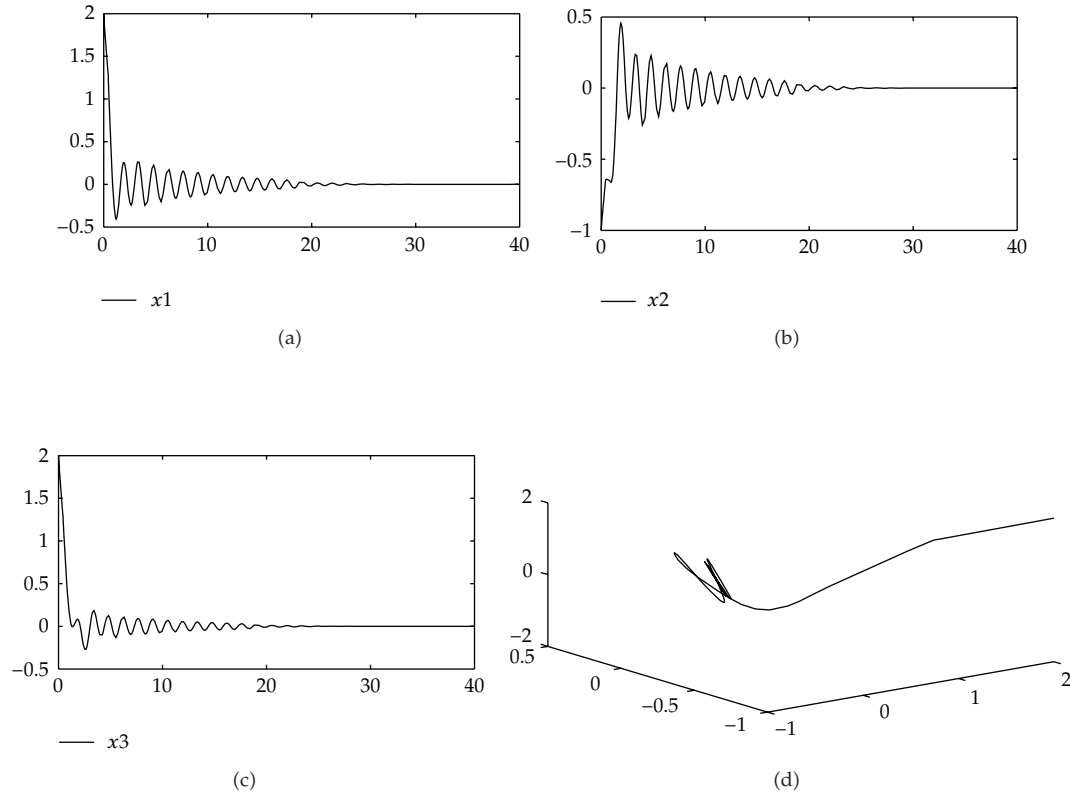


Figure 1: For system (3.4), when $\tau = 0.44 < \tau^*$, the equilibrium is asymptotically stable.

the imaginary eigenvalue $s \doteq \pm 0.445i$, $\tau^* \doteq 0.45$. From Theorem (3.2), we know that the zero solution of the system (3.4) is delay-dependent stable. The direction of the Hopf bifurcation at $\tau = \tau^*$ is supercritical and the bifurcating periodic solutions are asymptotically stable. The simulation results are shown in Figures 1 and 2.

4. Conclusion

In this paper, we consider a singular neutral differential system with a single delay. Via applying the algebraic method, that is, the matrix pencil and the linear operators, we discussed the eigenvalues and the stability of the time-delay systems (1.5) and (1.6). By using MATLAB, we could easily compute imaginary eigenvalues from the algebraic equation (2.23) or (2.26). In fact, we only find the imaginary axis eigenvalues, which are the small part of the infinite eigenvalues. So compared with the analytic methods and the numerical methods, the algebraic methods are more simple and more explicit for some time-delay system. Certainly, applying the algebraic methods to analyze the dynamical properties of the singular neutral differential systems with delays is a new and immature field. So we believe that the algebraic methods used to research the stability of the dynamical systems would be more interesting in the future.

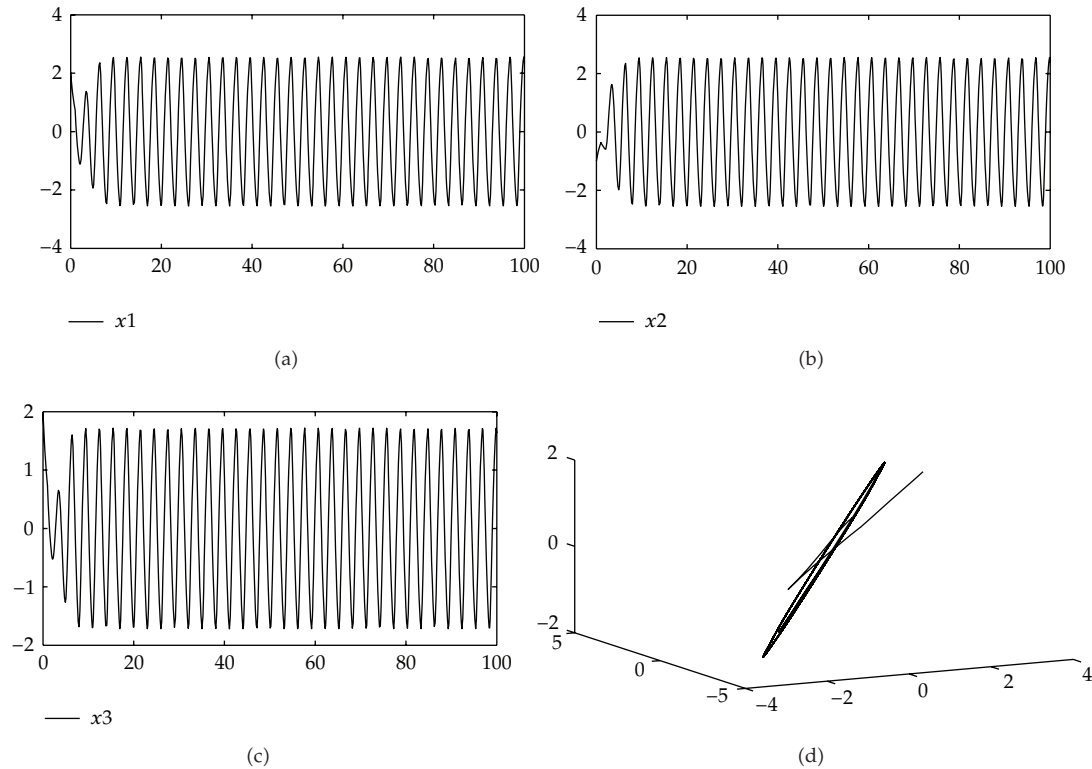


Figure 2: For system (3.4), when $\tau = 0.5 > \tau^*$, the periodic solution bifurcates from the equilibrium.

Acknowledgments

The National Natural Science Foundation of China (10871056) and the Fundamental Research Funds for the Central Universities (DL12BB24) are greatly acknowledged.

References

- [1] J. K. Hale, E. F. Infante, and F. S. P. Tsen, "Stability in linear delay equations," *Journal of Mathematical Analysis and Applications*, vol. 105, no. 2, pp. 533–555, 1985.
- [2] E. Jarlebring and M. E. Hochstenbach, "Polynomial two-parameter eigenvalue problems and matrix pencil methods for stability of delay-differential equations," *Linear Algebra and its Applications*, vol. 431, no. 3–4, pp. 369–380, 2009.
- [3] S. L. Campbell and V. H. Linh, "Stability criteria for differential-algebraic equations with multiple delays and their numerical solutions," *Applied Mathematics and Computation*, vol. 208, no. 2, pp. 397–415, 2009.
- [4] Z. Gao, "PD observer parametrization design for descriptor systems," *Journal of the Franklin Institute*, vol. 342, no. 5, pp. 551–564, 2005.
- [5] Z. Gao and S. X. Ding, "Actuator fault robust estimation and fault-tolerant control for a class of nonlinear descriptor systems," *Automatica*, vol. 43, no. 5, pp. 912–920, 2007.
- [6] Z. Gao, T. Breikin, and H. Wang, "Reliable observer-based control against sensor failures for systems with time delays in both state and input," *IEEE Transactions on Systems, Man, and Cybernetics A*, vol. 38, no. 5, pp. 1018–1029, 2008.
- [7] T. E. Simos, "Closed Newton-Cotes trigonometrically-fitted formulae of high order for long-time integration of orbital problems," *Applied Mathematics Letters*, vol. 22, no. 10, pp. 1616–1621, 2009.

- [8] S. Stavroyiannis and T. E. Simos, "Optimization as a function of the phase-lag order of nonlinear explicit two-step P -stable method for linear periodic IVPs," *Applied Numerical Mathematics*, vol. 59, no. 10, pp. 2467–2474, 2009.
- [9] T. E. Simos, "Exponentially and trigonometrically fitted methods for the solution of the Schrödinger equation," *Acta Applicandae Mathematicae*, vol. 110, no. 3, pp. 1331–1352, 2010.
- [10] W. Zhu and L. R. Petzold, "Asymptotic stability of linear delay differential-algebraic equations and numerical methods," *Applied Numerical Mathematics*, vol. 24, no. 2-3, pp. 247–264, 1997.
- [11] J. Louisell, "A matrix method for determining the imaginary axis eigenvalues of a delay system," *Institute of Electrical and Electronics Engineers*, vol. 46, no. 12, pp. 2008–2012, 2001.
- [12] W. Jiang, *The Degenerate Differential Systems with Delay*, Anhui University, Hefei, China, 1998.
- [13] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, "Vector spaces of linearizations for matrix polynomials," *SIAM Journal on Matrix Analysis and Applications*, vol. 28, no. 4, pp. 971–1004, 2006.
- [14] W. Michiels and T. Vyhídal, "An eigenvalue based approach for the stabilization of linear time-delay systems of neutral type," *Automatica*, vol. 41, no. 6, pp. 991–998, 2005.
- [15] E. Jarlebring, "On critical delays for linear neutral delay systems," in *Proceedings of the European Control Conference (ECC '07)*, Kos, Greece, 2007.
- [16] L. Li and Y. Yuan, "Dynamics in three cells with multiple time delays," *Nonlinear Analysis. Real World Applications*, vol. 9, no. 3, pp. 725–746, 2008.

Research Article

Type-K Exponential Ordering with Application to Delayed Hopfield-Type Neural Networks

Bin-Guo Wang

School of Mathematics and Statistics, Lanzhou University, Lanzhou, Gansu 730000, China

Correspondence should be addressed to Bin-Guo Wang, wangbinguo@lzu.edu.cn

Received 10 December 2011; Accepted 9 April 2012

Academic Editor: Chuanhou Gao

Copyright © 2012 Bin-Guo Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Order-preserving and convergent results of delay functional differential equations without quasimonotone condition are established under type-K exponential ordering. As an application, the model of delayed Hopfield-type neural networks with a type-K monotone interconnection matrix is considered, and the attractor result is obtained.

1. Introduction

Since monotone methods have been initiated by Kamke [1] and Müller [2], and developed further by Krasnoselskii [3, 4], Matano [5], and Smith [6], the theory and application of monotone dynamics have become increasingly important (see [7–18]).

It is well known that the quasimonotone condition is very important in studying the asymptotic behaviors of dynamical systems. If this condition is satisfied, the solution semiflows will admit order-preserving property. There are many interesting results, for example, [6, 8–12, 14–17] for competitive (cooperative) or type-K competitive (cooperative) systems and [6, 7, 13] for delayed systems. In particular, for the scalar delay differential equations of the form

$$x'(t) = g(x(t), x(t-r)), \quad (1.1)$$

if the quasimonotone condition $(\partial g(x, y))/\partial y > 0$ holds, then (1.1) generates an eventually strongly monotone semiflow on the space $C([-r, 0], \mathbb{R})$, which is one of sufficient conditions for obtaining convergent results. In other words, the right hand side of (1.1) must be strictly increasing in the delayed argument. This is a severe restriction, and so the quasimonotone conditions are not always satisfied in applications. Recently, many researchers have tried

to relax the quasimonotone condition by introducing a new cone or partial ordering, for example, the exponential ordering [6, 18, 19]. In particular, Smith [6] and Wu and Zhao [18] considered a new cone parameterized by a nonnegative constant, which is applicable to a single equation. Replacing the previous constant by a quasipositive matrix, the exponential ordering is generalized to some delay differential systems by Smith [6] and Y. Wang and Y. Wang [19]. However, the above results are not suitable to the type-K systems (see [6] for its definition). A typical example is a Hopfield-type neural network model with a type-K monotone interconnection matrix, which implies that the interaction among neurons is not only excitatory but also inhibitory. For this purpose, we introduce a type-K exponential ordering and establish order-preserving and convergent results under the weak quasimonotone condition (WQM) (see Section 2) and then apply the result to a network model with a type-K monotone interconnection matrix.

This paper is arranged as follows. In next section, the type-K exponential ordering parameterized by a type-K monotone matrix is introduced, and convergent result is established. In Section 3, we apply our results to a delayed Hopfield-type neural network.

2. Type-K Exponential Ordering

In this section, we establish a new cone and introduce some order-preserving and convergent results.

Let (X_i, X_i^+) , $i \in N = \{1, 2, \dots, n\}$, be ordered Banach spaces with $\text{Int}X_i^+ \neq \emptyset$. For $x_i, y_i \in X_i$, we write $x_i \leq_{X_i} y_i$ if $y_i - x_i \in X_i^+$; $x_i <_{X_i} y_i$ if $y_i - x_i \in X_i^+ \setminus \{0\}$; $x_i \ll_{X_i} y_i$ if $y_i - x_i \in \text{Int}X_i^+$. For $k \in N$, we denote $I = \{1, 2, \dots, \kappa\}$ and $J = N \setminus I = \{\kappa + 1, \dots, n\}$. Thus, we can define the product space $X = \prod_{i=1}^n X_i$ which generates two cones $X^+ = \prod_{i=1}^n X_i^+$ and $K = \prod_{i=1}^{\kappa} X_i^+ \times \prod_{i=\kappa+1}^n (-X_i^+)$ with nonempty interiors $\text{Int}X^+ = \prod_{i=1}^n \text{Int}X_i^+$ and $\text{Int}K = \prod_{i=1}^{\kappa} \text{Int}X_i^+ \times \prod_{i=\kappa+1}^n (-\text{Int}X_i^+)$. The ordering relation on X^+ and K is defined in the following way:

$$\begin{aligned}
 x \leq_X y &\iff x_i \leq_{X_i} y_i, \quad \forall i \in N, \\
 x <_X y &\iff x \leq y, \quad x_i <_{X_i} y_i, \quad \text{for some } i \in N, \text{ that is, } x \leq_X y, \quad x \neq y, \\
 x \ll_X y &\iff x_i \ll_{X_i} y_i, \quad \forall i \in N, \\
 x \leq_K y &\iff x_i \leq_{X_i} y_i, \quad \forall i \in I, \quad x_i \geq_{X_i} y_i, \quad \forall i \in J, \\
 x <_K y &\iff x \leq_K y, \quad x_i <_{X_i} y_i, \quad \text{for some } i \in I \quad \text{or} \quad x_i >_{X_i} y_i, \quad \text{for some } i \in J, \\
 x \ll_K y &\iff x_i \ll_{X_i} y_i, \quad \forall i \in I, \quad x_i \gg_{X_i} y_i, \quad \forall i \in J.
 \end{aligned} \tag{2.1}$$

A semiflow on X is a continuous mapping $\Phi: X \times \mathbb{R}_+ \rightarrow X$, $(x, t) \rightarrow \Phi(x, t)$, which satisfies (i) $\Phi_0 = id$ and (ii) $\Phi_t \cdot \Phi_s = \Phi_{t+s}$ for $t, s \in \mathbb{R}_+$. Here, $\Phi_t(x) \equiv \Phi(x, t)$ for $x \in X$ and $t \geq 0$. The orbit of x is denoted by $O(x)$:

$$O(x) = \{\Phi_t(x) : t \geq 0\}. \tag{2.2}$$

An equilibrium point is a point x for which $\Phi_t(x) = x$ for all $t \geq 0$. Let E be the set of all equilibrium points for Φ . The omega limit set $\omega(x)$ of x is defined in the usual way. A point $x \in X$ is called a quasiconvergent point if $\omega(x) \subset E$. The set of all such points is denoted by \mathcal{Q} .

A point $x \in X$ is called a *convergent point* if $\omega(x)$ consists of a single point of E . The set of all convergent points is denoted by C .

The semiflow Φ is said to be *type-K monotone* provided

$$\Phi_t(x) \leq_K \Phi_t(y) \quad \text{whenever } x \leq_K y \quad \forall t \geq 0. \quad (2.3)$$

Φ is called *type-K strongly order preserving* (for short type-K SOP), if it is type-K monotone, and whenever $x <_K y$, there exist open subsets U, V of X with $x \in U, y \in V$ and $t_0 > 0$, such that

$$\Phi_t(U) \leq_K \Phi_t(V) \quad \forall t \geq t_0. \quad (2.4)$$

The semiflow Φ is said to be *strongly type-K monotone* on X if Φ is type-K monotone, and whenever $x <_K y$ and $t > 0$, then $\Phi_t(x) \ll_K \Phi_t(y)$. We say that Φ is *eventually strongly type-K monotone* if it is type-K monotone, and whenever $x <_K y$, there exists $t_0 > 0$ such that $\Phi_{t_0}(x) \ll_K \Phi_{t_0}(y)$. Clearly, strongly type-K monotonicity implies eventually strongly type-K monotonicity.

An $n \times n$ matrix M is said to be *type-K monotone* if it has the following manner:

$$M = \begin{pmatrix} \bar{A} & -\bar{B} \\ -\bar{C} & \bar{D} \end{pmatrix}, \quad (2.5)$$

where $\bar{A} = (a_{ij})_{k \times k}$ satisfies $(a_{ij}) \geq 0$ if $i \neq j$, similarly for the $(n-k) \times (n-k)$ matrix \bar{D} and $\bar{B} \geq 0, \bar{C} \geq 0$.

In this paper, the following lemma is necessary.

Lemma 2.1. *If M is a type-K monotone matrix, then e^{Mt} remains type-K monotone with diagonal entries being strictly positive for all $t > 0$.*

Proof. The product of two type-K monotone matrices remains type-K monotone; the rest is obvious and we omit it here. \square

Let $r > 0$ be fixed and let $C := C([-r, 0], X)$. The ordering relations on C are understood to hold pointwise. Consider the family of sets parameterized by type-K monotone matrix M given by

$$\tilde{K}_M = \left\{ \phi = (\phi_1, \phi_2, \dots, \phi_n) \in C : \phi(s) \geq_K 0, s \in [-r, 0], \phi(t) \geq_K e^{M(t-s)} \phi(s), 0 \geq t \geq s \geq -r \right\}. \quad (2.6)$$

It is easy to see that \tilde{K}_M is a closed cone in C and generates a partial ordering on C which is written by \geq_M . Assume that $\phi \in C$ is differentiable on $(-r, 0)$, a similar argument to [18, lemma 2.1] implies that $\phi \geq_M 0$ if and only if $\phi(-r) \geq_K 0$ and $d\phi(s)/ds - M\phi(s) \geq_K 0$ for all $s \in (-r, 0)$.

Consider the abstract functional differential equation

$$x'(t) = f(x_t), \quad (2.7)$$

where $f : D \rightarrow X$ is continuous and satisfies a local Lipschitz condition on each compact subset of D and D is an open subset of C . By the standard equation theory, the solution $x(t, \phi)$ of (2.7) can be continued to the maximal interval of existence $[0, \sigma_\phi)$. Moreover, if $\sigma_\phi > r$, then $x(t, \phi)$ is a classical solution of (2.7) for $t \in (r, \sigma_\phi)$. In this section, for simplicity, we assume that, for each $\phi \in D$, (2.7) admits a solution $x(t, \phi)$ defined on $[0, \infty)$. Therefore, (2.7) generates a semiflow on C by $\Phi_t(\phi) \equiv x_t(\phi)$, where $x_t(\phi)(s) = x(t+s, \phi)$ for $t \geq 0$ and $-r \leq s \leq 0$.

In the following, we will seek a sufficient condition for the solution of (2.7) to preserve the ordering \geq_M .

(WQM) Whenever $\phi, \psi \in D$, $\psi \geq_M \phi$, then

$$f(\psi) - f(\phi) \geq_K M(\psi(0) - \phi(0)). \quad (2.8)$$

Theorem 2.2. Suppose that (WQM) holds. If $\psi \geq_M \phi$, then $x_t(\psi) \geq_M x_t(\phi)$ for all $t \geq 0$.

Proof. Let $\eta \in \text{Int}K$. For any $\varepsilon > 0$, define $f_\varepsilon(\phi) = f(\phi) + \varepsilon\eta$ for $\phi \in D$, and let $x_t^\varepsilon(\psi)$ be a unique solution of the following equation:

$$\begin{aligned} x'(t) &= f_\varepsilon(x_t), \quad t \geq 0, \\ x(s) &= \psi(s), \quad -r \leq s \leq 0. \end{aligned} \quad (2.9)$$

Let $y^\varepsilon(t) = x^\varepsilon(t, \psi) - x(t, \phi)$ and define

$$S = \{t \in [0, \infty) : y_t^\varepsilon \geq_M 0\}. \quad (2.10)$$

Since $\psi \geq_M \phi$, S is closed and nonempty. We first prove the following two claims.

Claim 1. If $t_0 \in S$, there exists $\delta_0 > 0$ such that $[t_0, t_0 + \delta_0] \subset S$.

According to the integral expression of (2.9) we have

$$y^\varepsilon(t) = e^{M(t-s)} y^\varepsilon(s) + \int_s^t e^{M(\tau-s)} [f(x_\tau^\varepsilon(\psi)) - f(x_\tau(\phi)) - M(x^\varepsilon(\tau, \psi) - x(\tau, \phi)) + \varepsilon\eta] d\tau. \quad (2.11)$$

Since $t_0 \in S$ and (WQM) hold, we have

$$f(x_t^\varepsilon(\psi)) - f(x_t(\phi)) - M(x^\varepsilon(t, \psi) - x(t, \phi)) + \varepsilon\eta|_{t=t_0} \geq_K \varepsilon\eta \gg_K 0. \quad (2.12)$$

By the characteristic of a cone, there is $\delta_0 > 0$ such that

$$f(x_t^\varepsilon(\psi)) - f(x_t(\phi)) - M(x^\varepsilon(t, \psi) - x(t, \phi)) + \varepsilon\eta \geq_K 0, \quad \forall t \in [t_0, t_0 + \delta_0]. \quad (2.13)$$

By Lemma 2.1, we have

$$y^\epsilon(t) \geq_K e^{M(t-s)} y^\epsilon(s), \quad \forall t_0 \leq s \leq t \leq t_0 + \delta_0, \quad (2.14)$$

which, together with the definition of \tilde{K}_M , implies that

$$x_t^\epsilon(\psi) \geq_M x_t(\phi), \quad \forall t \in [t_0, t_0 + \delta_0]. \quad (2.15)$$

Claim 2. Let $S_1 = \{t : [0, t] \subset S\}$. Then $\sup S_1 = \infty$.

If $t^* = \sup S_1 < \infty$, then there is a sequence $\{t_n\} \subset S_1 \subset S$ such that $t_n \rightarrow t^*$ as $n \rightarrow \infty$. From the closeness of S we have $t^* \in S$. By Claim 1, $[t^*, t^* + \delta^*] \subset S$ for some $\delta^* > 0$, which contradicts the definition of t^* . Therefore, $\sup S_1 = \infty$, which implies $S = [0, \infty)$.

Since $f_\epsilon \rightarrow f$ uniformly on bounded subset of D as $\epsilon \rightarrow 0^+$, then

$$\lim_{\epsilon \rightarrow 0^+} x_t^\epsilon(\psi) = x_t(\psi), \quad \forall t \geq 0. \quad (2.16)$$

Letting $\epsilon \rightarrow 0^+$ in $y_t^\epsilon = x_t^\epsilon(\psi) - x_t(\phi) \geq_M 0$, we have $x_t(\psi) - x_t(\phi) \geq_M 0$, which implies that $x_t(\psi) \geq_M x_t(\phi)$. \square

By the definition of the semiflow Φ_t , it is easy to see from (WQM) that Φ_t is monotone with respect to \geq_M in the sense that $\Phi_t(\psi) \geq_M \Phi_t(\phi)$ whenever $\psi \geq_M \phi$ for all $t \geq 0$.

As we all know the strongly order-preserving property is necessary for obtaining some convergent results. However, it is easy to check that the cone \tilde{K}_M has empty interior on C ; we cannot, therefore, expect to show that the semiflow generated by (2.7) is eventually strongly type-K monotone in C . Let $\varphi(\cdot) \in \text{Int}K$ and define

$$\begin{aligned} C_\varphi &= \{\phi \in C : \text{there exist } \gamma \geq 0 \text{ such that } -\gamma\varphi \leq_M \phi \leq_M \gamma\varphi\}, \\ \|\phi\|_\varphi &= \inf\{\gamma \geq 0 : -\gamma\varphi \leq_M \phi \leq_M \gamma\varphi\}. \end{aligned} \quad (2.17)$$

It is easy to check that $(C_\varphi, \|\phi\|_\varphi)$ is a Banach space, $K_M = C_\varphi \cap \tilde{K}_M$ is a cone with nonempty interior $\text{Int}K_M$ (see [20]), and $i : C_\varphi \rightarrow C$ is continuous. Using the smoothing property of the semiflow Φ on C^+ and fundamental theory of abstract functional differential equations, we deduce that for all $t > r$, $\Phi_t C \subset C \cap C_\varphi$, $\Phi_t : C \rightarrow C \cap C_\varphi$ is continuous, and $\Phi_t(\psi - \phi) \in \text{Int}K_M$ for any $\psi, \phi \in C$ with $\psi >_M \phi$. Thus, from Theorem 2.2, type-K strongly order-preserving property can be obtained.

Theorem 2.3. Assume that (WQM) holds. If $\psi >_M \phi$, then $x_t(\psi) \gg_M x_t(\phi)$ in K_M for all $t \geq r$.

In order to obtain the main result of this paper, which says that the generic solution converges to equilibrium, the corresponding compactness assumption will be required.

- (A1) f maps bounded subset of D to bounded subset of \mathbb{R}^n . Moreover, for each compact subset A of D , there exists a closed and bounded subset $B = B(A)$ of D such that $x_t(\phi) \in B$ for each $\phi \in A$ and all large t .

Theorem 2.4. Assume that (WQM) and (A1) hold. Then the set of convergent points in D contains an open and dense subset. If \mathbf{E} consists of a single point, it attracts all solutions of (2.7). If the initial value $x_0 \geq_K 0$ ($x_0 \leq_K 0$) and \mathbf{E} consists of two points or more, we conclude that all solutions converge to one of these.

Proof. By Theorem 2.3, the semiflow is eventually strongly monotone in K_M . Let $\hat{e} = (\hat{1}, \dots, \hat{1}, -\hat{1}, \dots, -\hat{1}) \in K$, where $\hat{1}$ denotes a constant mapping defined on C ; that is, $\hat{1}(s) = 1$ for all $s \in [-r, 0]$. Obviously, $\hat{e} \geq_M \hat{0}$. For any $\varphi \in D$, either the sequence of points $\varphi + (1/n)\hat{e}$ or $\varphi - (1/n)\hat{e}$ is eventually contained in D and approaches φ as $n \rightarrow \infty$, and, hence, each point of D can be approximated either from above or from below in D with respect to \geq_M . The assumption (A1) implies the compactness; that is, $O(x)$ has compact closure in X for each $x \in X$ (see [6]). Therefore, from [6, Theorem 1.4.3], we deduce that the set of quasiconvergent points contains an open and dense subset of D . From the proof of [6, Theorem 6.3.1], we know that the set \mathbf{E} is totally ordered by \geq_M . Reference [6, Remark 1.4.2] implies that the set of convergent points contains an open and dense subset of D . The last two assertions can be obtained from [6, Theorems 2.3.1 and 2.3.2]. \square

Remark 2.5. The above theorem implies that there exists an equilibrium attracting all solutions with initial values in the cone K . If \mathbf{E} consists of a single element, the equilibrium attracts all solutions with initial values in D .

3. Delayed Hopfield-Type Neural Networks

In this section, we will apply our main result to the following system of delayed differential equations:

$$x'_i(t) = -a_i x_i(t) + \sum_{j=1}^n a_{ij} f_j(x_j(t - r_j)) + I_i, \quad i = 1, 2, \dots, n, \quad (3.1)$$

where $a_i > 0$ and $r_j \geq 0$ are constant, $i, j = 1, \dots, n$. The interconnection matrix $(a_{ij})_{n \times n}$ is type-K monotone with the elements in the diagonal being nonnegative. In this situation, the interaction among neurons is not only excitatory but also inhibitory. The external input functions I_i are constants or periodic. The activation functions $f = (f_1, \dots, f_n) : D \rightarrow \mathbb{R}$, where D is an open subset of $X = C([-r, 0], \mathbb{R}^n)$ with $r = \max\{r_j | j \in N\}$, satisfy (A1) and following property.

(A2) There exist constants L_j such that $|f_j(x) - f_j(y)| \leq L_j |x - y|$ for $j = 1, \dots, n$.

First, we consider the case that the external input functions I_i are constants.

Theorem 3.1. Equation (3.1) has an equilibrium which attracts all its solutions coming from the initial value $\phi \geq_K 0$ with $\phi(0)$ being bounded.

Proof. From [21, Theorem 1], we deduce that (3.1) admits at least an equilibrium; that is, the equilibrium points set \mathbf{E} is nonempty.

For $\phi \in X$, we define

$$F_i(\phi) = -a_i\phi_i(0) + \sum_{j=1}^n a_{ij}f_j(\phi_j(-r_j)) + I_i. \quad (3.2)$$

Choosing $M = \text{diag}\{-\mu, \dots, -\mu\}$ with $\mu > 0$, and denoting $L = \max_{1 \leq j \leq n} L_j$, $\alpha = \max_{1 \leq i, j \leq n} |a_{ij}|$ and $\beta = \max_{1 \leq j \leq n} a_j$. Since $\phi(0)$ is bounded, for $\varphi, \phi \in D$ with $\varphi \geq_M \phi$, there exist $\overline{m} \geq 0$ and $\underline{m} \geq 0$ with $\overline{m} \geq \underline{m}$ such that

$$\begin{aligned} \underline{m} &\leq \varphi_j(0) - \phi_j(0) \leq \overline{m}, \quad \forall i \in I, \\ -\overline{m} &\leq \varphi_j(0) - \phi_j(0) \leq -\underline{m}, \quad \forall i \in J. \end{aligned} \quad (3.3)$$

From (A2) and the definition of \tilde{K}_M , if $\varphi \geq_M \phi$, then

$$\begin{aligned} &F_i(\varphi) - F_i(\phi) + \mu(\varphi_i(0) - \phi_i(0)) \\ &= (\mu - a_i)(\varphi_i(0) - \phi_i(0)) + \sum_{j=1}^n a_{ij}(f_j(\varphi_j(-r_j)) - f_j(\phi_j(-r_j))) \\ &\geq (\mu - a_i)(\varphi_i(0) - \phi_i(0)) - \sum_{j=1}^k a_{ij}L_j(\varphi_j(-r_j) - \phi_j(-r_j)) \\ &\quad - \sum_{j=k+1}^n a_{ij}L_j(\varphi_j(-r_j) - \phi_j(-r_j)) \\ &\geq (\mu - a_i)(\varphi_i(0) - \phi_i(0)) - \sum_{j=1}^k a_{ij}L_j e^{\mu r_j}(\varphi_j(0) - \phi_j(0)) \\ &\quad - \sum_{j=k+1}^n a_{ij}L_j e^{\mu r_j}(\varphi_j(0) - \phi_j(0)) \\ &\geq \left(\mu - \beta \frac{\overline{m}}{\underline{m}} - n\alpha L e^{\mu r} \frac{\overline{m}}{\underline{m}} \right) \underline{m}, \end{aligned} \quad (3.4)$$

for all $i \in I$. By a similar argument we have

$$F_i(\varphi) - F_i(\phi) + \mu(\varphi_i(0) - \phi_i(0)) \leq \left(\mu - \beta \frac{\overline{m}}{\underline{m}} - n\alpha L e^{\mu r} \frac{\overline{m}}{\underline{m}} \right) (-\underline{m}) \quad (3.5)$$

for all $i \in J$. Let $H = \beta \overline{m} / \underline{m}$ and let $G = n\alpha L \overline{m} / \underline{m}$, and define $g(\mu) = \mu - H - Ge^{\mu r}$. If $r = 0$, we have $g(\mu) \geq 0$ for $\mu \geq H + G$. If $r > 0$ and $Ge^{Hr}r < 1/e$, we deduce that $g(\mu)$ reaches its positive maximum value at $\mu = H + (1/r) \ln(1/Ge^{Hr}r) > 0$. Thus, there exists a positive constant μ such that (WQM) holds; the conclusion can be obtained by Remark 2.5. \square

For the case of the external input functions I_i being periodic functions, we have following result.

Theorem 3.2. *For any periodic external input function $I(t) = (I_1(t), \dots, I_n(t))$, $I_i(t + \omega) = I_i(t)$, $i = 1, \dots, n$, (3.1) admits a unique periodic solution $x^*(t)$ and all other solutions which come from the initial value $\phi \geq_K 0$ with $\phi(0)$ being bounded converge to it as $t \rightarrow \infty$.*

Proof. Let $x(t) = x(t, \phi)$ be the solution of (3.1) for $t \geq 0$ with $x(s) = \phi(s)$ for $s \in [-r, 0]$. From the properties of the solution semiflow we have

$$x(t + \omega) = x(t + \omega, \phi) = x(t, x(\omega, \phi)). \quad (3.6)$$

From the proof of Theorem 3.1, we know that there exists a type-K monotone matrix such that (WQM) holds; Theorem 2.4 tells us that every orbit of (3.1) is convergent to a same equilibrium, denoted by ϕ^* , and then,

$$\lim_{n \rightarrow \infty} x(n\omega, \phi) = \phi^*. \quad (3.7)$$

We have, therefore,

$$x(\omega, \phi^*) = x\left(\omega, \lim_{n \rightarrow \infty} x(n\omega, \phi)\right) = \lim_{n \rightarrow \infty} x(\omega, x(n\omega, \phi)) = \lim_{n \rightarrow \infty} x((n+1)\omega, \phi) = \phi^*. \quad (3.8)$$

From (3.6) and (3.8) we deduce that

$$x(t + \omega, \phi^*) = x(t, x(\omega, \phi^*)) = x(t, \phi^*). \quad (3.9)$$

Therefore, $x(t, \phi^*) =: x^*(t)$ is a unique periodic solution of (3.1). Using the conclusion of Theorem 2.4 again, we have

$$\lim_{t \rightarrow \infty} x(t, \phi) = \lim_{t \rightarrow \infty} x(t, x(t, \phi)) = \lim_{t \rightarrow \infty} x(t, \phi^*). \quad (3.10)$$

Since $x^*(t)$ is a periodic solution, the proof is complete. \square

Remark 3.3. Neural networks have important applications, such as to content-addressable memory [22], shortest path problem [23], and sorting problem [24]. Generally, the monotonicity is always assumed. Here, we relax the monotone condition, and hence neural networks have more extensive applications.

Acknowledgments

This paper is supported by NSF of China under Grant 10926091 and the Fundamental Research Funds for the Central Universities.

References

- [1] E. Kamke, "Zur Theorie der Systeme gewöhnlicher Differentialgleichungen. II," *Acta Mathematica*, vol. 58, no. 1, pp. 57–85, 1932.
- [2] M. Müller, "Über das fundamentaltheorem in der theorie der gewöhnlichen differentialgleichungen," *Mathematische Zeitschrift*, vol. 26, pp. 619–645, 1926.
- [3] M. A. Krasnoselskii, *Positive Solutions of Operator Equations*, Noordhoff Groningen, Groningen, The Netherlands, 1964.
- [4] M. A. Krasnoselskii, *The Operator of Translation Along Trajectories of Differential Equations*, vol. 19 of *Translations of Mathematical Monographs*, American Mathematical Society, Providence, RI, USA, 1968.
- [5] H. Matano, "Existence of nontrivial unstable sets for equilibriums of strongly order-preserving systems," *Journal of the Faculty of Science IA*, vol. 30, no. 3, pp. 645–673, 1984.
- [6] H. L. Smith, *Monotone Dynamics Systems: An Introduction to the Theory of Competitive and Cooperative Systems*, American Mathematical Society, Providence, RI, USA, 1995.
- [7] H. I. Freedman and X.-Q. Zhao, "Global asymptotics in some quasimonotone reaction-diffusion systems with delays," *Journal of Differential Equations*, vol. 137, no. 2, pp. 340–362, 1997.
- [8] M. Gyllenberg and Y. Wang, "Dynamics of the periodic type-K competitive Kolmogorov systems," *Journal of Differential Equations*, vol. 205, no. 1, pp. 50–76, 2004.
- [9] X. Liang and J. Jiang, "On the finite-dimensional dynamical systems with limited competition," *Transactions of the American Mathematical Society*, vol. 354, no. 9, pp. 3535–3554, 2002.
- [10] X. Liang and J. Jiang, "The classification of the dynamical behavior of 3-dimensional type-K monotone Lotka-Volterra systems," *Nonlinear Analysis. Theory, Methods & Applications A*, vol. 51, no. 5, pp. 749–763, 2002.
- [11] X. Liang and J. Jiang, "The dynamical behaviour of type-K competitive Kolmogorov systems and its application to three-dimensional type-K competitive Lotka-Volterra systems," *Nonlinearity*, vol. 16, no. 3, pp. 785–801, 2003.
- [12] X. Liang and J. Jiang, "Discrete infinite-dimensional type-K monotone dynamical systems and time-periodic reaction-diffusion systems," *Journal of Differential Equations*, vol. 189, no. 1, pp. 318–354, 2003.
- [13] R. H. Martin Jr. and H. L. Smith, "Reaction-diffusion systems with time delays: monotonicity, invariance, comparison and convergence," *Journal für die Reine und Angewandte Mathematik*, vol. 413, pp. 1–35, 1991.
- [14] H. L. Smith, "Competing subcommunities of mutualists and a generalized Kamke theorem," *SIAM Journal on Applied Mathematics*, vol. 46, no. 5, pp. 856–874, 1986.
- [15] C. Tu and J. Jiang, "The coexistence of a community of species with limited competition," *Journal of Mathematical Analysis and Applications*, vol. 217, no. 1, pp. 233–245, 1998.
- [16] C. Tu and J. Jiang, "Global stability and permanence for a class of type-K monotone systems," *SIAM Journal on Mathematical Analysis*, vol. 30, no. 2, pp. 360–378, 1999.
- [17] C. Tu and J. Jiang, "The necessary and sufficient conditions for the global stability of type-K Lotka-Volterra system," *Proceedings of the American Mathematical Society*, vol. 127, no. 11, pp. 3181–3186, 1999.
- [18] J. Wu and X.-Q. Zhao, "Diffusive monotonicity and threshold dynamics of delayed reaction diffusion equations," *Journal of Differential Equations*, vol. 186, no. 2, pp. 470–484, 2002.
- [19] Y. Wang and Y. Wang, "Global dynamics of reaction-diffusion systems with delays," *Applied Mathematics Letters*, vol. 18, no. 9, pp. 1027–1033, 2005.
- [20] H. Amann, "Fixed point equations and nonlinear eigenvalue problems in ordered Banach spaces," *SIAM Review*, vol. 18, no. 4, pp. 620–709, 1976.
- [21] H. Huang, J. Cao, and J. Wang, "Global exponential stability and periodic solutions of recurrent neural networks with delays," *Physics Letters A*, vol. 298, no. 5-6, pp. 393–404, 2002.
- [22] S. Grossberg, "Nonlinear neural networks: principles, mechanisms, and architectures," *Neural Networks*, vol. 1, no. 1, pp. 17–61, 1988.
- [23] J. Wang, "A recurrent neural network for solving the shortest path problem," *IEEE Transactions on Circuits and Systems I*, vol. 43, no. 6, pp. 482–486, 1996.
- [24] J. Wang, "Analysis and design of an analog sorting network," *IEEE Transactions on Neural Networks*, vol. 6, no. 4, pp. 962–971, 1995.

Research Article

Mean Square Almost Periodic Solutions for Impulsive Stochastic Differential Equations with Delays

Ruojun Zhang,¹ Nan Ding,^{2,3} and Linshan Wang¹

¹ School of Mathematical Sciences, Ocean University of China, Qingdao 266100, China

² College of Information Engineering, Ocean University of China, Qingdao 266100, China

³ Department of Mathematics, Chongqing Three Gorges University, Chongqing 404100, China

Correspondence should be addressed to Ruojun Zhang, zhangru1626@sina.com

Received 28 December 2011; Revised 15 March 2012; Accepted 15 March 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 Ruojun Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We establish a result on existence and uniqueness on mean square almost periodic solutions for a class of impulsive stochastic differential equations with delays, which extends some earlier works reported in the literature.

1. Introduction

Impulsive effects widely exist in many evolution processes of real-life phenomena in which states are changed abruptly at certain moments of time, involving such areas as population dynamics and automatic control [1–3]. Because delay is ubiquitous in the dynamical system, impulsive differential equations with delays have received much interesting in recent years, intensively researched, some important results are obtained [4–9]. And almost periodic solutions for abstract impulsive differential equations and for impulsive neural networks with delay have been discussed by G. T. Stamov and I. M. Stamova [10], and Stamov and Alzabut [11].

However, besides delay and impulsive effects, stochastic effects likewise exist in real system. A lot of dynamic systems have variable structures subject to stochastic abrupt changes, which may result from abrupt phenomena such as stochastic failures and repairs of components, changes in the interconnections of subsystems, sudden environment changes, and so on [12–14]. Moreover, differential descriptor systems also have abrupt changes [15, 16]. Recently, a large number of stability criteria of stochastic system with delays have

been reported [17–19]. Almost periodic solutions to some functional integro-differential stochastic evolution equations and to some stochastic differential equations have been studied by Bezandry and Diagana [20], and Bezandry [21]. Huang and Yang investigated almost periodic solution for stochastic cellular neural networks with delays [22]. Because it is not easy to deal with the case of coexistence of impulsive, delay and stochastic effects in a dynamical system, there are few results about this problems [23–25]. To the best of our knowledge, there exists no result on the existence and uniqueness of mean square almost periodic solutions for impulsive stochastic differential equations with delays.

Motivated by the above discussions, the main aim of this paper is to study the mean square almost periodic solutions for impulsive stochastic differential equations with delays. By employing stochastic analysis, delay differential inequality technique and fixed points theorem, we obtain some criteria to ensure the existence and uniqueness of mean square almost periodic solutions.

The rest of this paper is organized as follows: in Section 2, we introduce a class of impulsive stochastic differential equations with delays, and the relating notations, definitions and lemmas which would be used later; in Section 3, a new sufficient condition is proposed to ensure the existence and uniqueness of mean square almost periodic solutions; in Section 4, an example is constructed to show the effectiveness of our results. Finally, a conclusion is given in Section 5.

2. Preliminaries

Let $\mathbb{R} = (-\infty, +\infty)$, $\mathbb{N} = \{1, 2, 3, \dots\}$, and $\mathcal{B} = \{\{t_k\} : t_0 = 0 < t_1 < t_2 < \dots < t_k < t_{k+1} < \dots, \lim_{k \rightarrow +\infty} t_k = +\infty\}$ be the set of all sequence unbounded and strictly increasing. For $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times n}$, let $\|x\|$ be any vector norm, and denote the induced matrix norm and the matrix measure, respectively, by

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}, \quad \mu(A) = \lim_{h \rightarrow 0^+} \frac{\|I + hA\| - 1}{h}. \quad (2.1)$$

The norm and measure of vector and matrix are $\|x\| = \max_i |x_i|$, $\|A\| = \max_i \sum_{j=1}^n |a_{ij}|$, $\mu(A) = \max_i \{a_{ii} + \sum_{j \neq i}^n |a_{ij}|\}$.

Consider the following a class of Itô impulsive stochastic differential equations with delay

$$\begin{aligned} dx(t) &= [Ax(t) + Bf(t, x(t)) + Cg(t, x(t-h)) + I(t)]dt + \sigma(t, x(t))d\omega(t), \quad t \geq 0, \quad t \neq t_k, \\ \Delta x(t) &= x(t_k) - x(t_k^-) = D_k x(t_k^-) + V_k(x(t_k^-)) + \beta_k, \quad t = t_k, \quad k \in \mathbb{N}, \\ x(t) &= \phi(t), \quad -h \leq t \leq 0, \end{aligned} \quad (2.2)$$

where $x(t) = (x_1(t), \dots, x_n(t))^T$ is the solution process, $A, B, C, D_k \in \mathbb{R}^{n \times n}$ are constant matrices, $f(t, x) = (f_1(t, x), \dots, f_n(t, x))^T$, $g(t, x) = (g_1(t, x), \dots, g_n(t, x))^T$, $I(t) = (I_1(t), \dots, I_n(t))^T$, $\sigma(t, x) = (\sigma_{ij}(t, x))_{n \times n}$ is the diffusion coefficient matrix, $V_k(x) = (V_{1k}(x), \dots, V_{nk}(x))^T$ is impulsive function, $h > 0$ is delay; $t_k \in \mathcal{B}$ is impulsive time, $\beta_k = (\beta_{1k}, \dots, \beta_{nk})^T$ is a constant vector, $\omega(t) = (\omega_1(t), \dots, \omega_n(t))^T$ is an n -dimensional Brown motion defined on a complete

probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a natural filtration $\{\mathcal{F}_t\}_{t \geq 0}$ generated by $\omega(t)$, and denote by \mathcal{F} the associated σ -algebra generated by $\omega(t)$ with the probability measure \mathbb{P} . Moreover, the initial conditions $\phi(t) = (\phi_1(t), \dots, \phi_n(t))^T \in PCB_{\mathcal{F}_0}^b([-h, 0], \mathbb{R}^n) \triangleq PCB_{\mathcal{F}_0}^b$. Denote by $PCB_{\mathcal{F}_0}^b$ the family of all bounded \mathcal{F}_0 -measurable, $PC([-h, 0], \mathbb{R}^n)$ -valued random variable ζ , satisfying $E\|\zeta\|^2 = E(\sup_{-h \leq \theta \leq 0} \|\zeta(\theta)\|^2) < +\infty$, where $PC([-h, 0], \mathbb{R}^n) = \{\zeta : [-h, 0] \rightarrow \mathbb{R}^n \text{ is continuous}\}$. E denotes the expectation of stochastic process.

Let $(\mathbb{H}, \|\cdot\|)$ be a Hilbert space and $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. Define $L^2(\mathbb{P}, \mathbb{H})$ to be the space of all \mathbb{H} -value random variable Y such that

$$E\|Y\|^2 = \int_{\Omega} \|Y\|^2 d\mathbb{P} < \infty. \quad (2.3)$$

It is then routine to check that $L^2(\mathbb{P}, \mathbb{H})$ is a Hilbert space when it is equipped with its natural norm $\|\cdot\|_2$ defined by

$$\|Y\|_2 = \left(\int_{\Omega} \|Y\|^2 d\mathbb{P} \right)^{1/2} < \infty, \quad (2.4)$$

for each $Y \in L^2(\mathbb{P}, \mathbb{H})$.

Definition 2.1 (see [25]). For any $\phi \in PCB_{\mathcal{F}_0}^b$, a function $x(t) : [-h, +\infty) \rightarrow L^2(\mathbb{P}, \mathbb{H})$ is said to be solution of system (2.2) on $[-h, +\infty)$ satisfying initial value condition, if the following conditions hold:

- (i) $x(t)$ is absolutely continuous on each interval $(t_k, t_{k+1}) \in [0, +\infty)$, $k \in \mathbb{N}$;
- (ii) for any $t_k \in [0, +\infty)$, $k \in \mathbb{N}$, $x(t_k^+)$ and $x(t_k^-)$ exist and $x(t_k^+) = x(t_k^-)$;
- (iii) $x(t)$ satisfies (2.2) for almost everywhere in $[-h, +\infty)$ and at impulsive points $t = t_k$ situated in $[0, +\infty)$, $k \in \mathbb{N}$, may have discontinuity points of the first kind.

Obviously, the solution defined by definition 1 is piecewise continuous.

Definition 2.2 (see [26]). The set of sequences $\{t_k^j\}$, $t_k^j = t_{k+j} - t_k$, $k \in \mathbb{N}$, $j \in \mathbb{N}$, $\{t_k\} \in \mathcal{B}$ is said to be uniformly almost periodic if for any $\varepsilon > 0$, there exists relatively dense set of ε -almost periods common for any sequences.

Definition 2.3. A piecewise continuous function $x(t) : [-h, +\infty) \rightarrow L^2(\mathbb{P}, \mathbb{H})$ with discontinuity points of first kind at $t = t_k$ is said to be mean square almost periodic, if

- (i) the set of sequence $\{t_k^j\}$ is uniformly almost periodic;
- (ii) for any $\varepsilon > 0$, there exists $\delta > 0$, such that if the points t' and t'' belong to one and the same interval of continuity of $x(t)$ and satisfy the inequality $|t' - t''| < \delta$, then $E\|x(t') - x(t'')\|^2 < \varepsilon$;
- (iii) for any $\varepsilon > 0$, there exists a relatively dense set T such that if $\tau \in T$, then $E\|x(t + \tau) - x(t)\|^2 < \varepsilon$ for all $t \in [-h, +\infty)$ satisfying the condition $|t - t_k| > \varepsilon$, $k \in \mathbb{N}$.

The collection of all functions $x(t) : [-h, +\infty) \rightarrow L^2(\mathbb{P}, \mathbb{H})$ with discontinuity points of the first kind at $t = t_k$ which are mean square almost periodic is denoted by

$AP([-h, +\infty); L^2(\mathbb{P}, \mathbb{H}))$, one can check that $AP([-h, +\infty); L^2(\mathbb{P}, \mathbb{H}))$ is a Banach space when it is equipped with the norm:

$$\|x\|_\infty = \sup_{t \in \mathbb{R}} \left(E\|x(t)\|^2 \right)^{1/2}. \quad (2.5)$$

Let $(B_1, \|\cdot\|_1)$ and $(B_2, \|\cdot\|_2)$ be Banach space and $L^2(\mathbb{P}, B_1)$ and $L^2(\mathbb{P}, B_2)$ be their corresponding L^2 -space, respectively.

Lemma 2.4 (see [20]). *Let $f : \mathbb{R} \times L^2(\mathbb{P}, B_1) \rightarrow L^2(\mathbb{P}, B_2), (t, x) \mapsto f(t, x)$ be mean square almost periodic in $t \in \mathbb{R}$ uniformly in $x \in K$, where $K \subset L^2(\mathbb{P}, B_1)$ is compact. Suppose that there exists $L_f > 0$ such that*

$$E\|f(t, x) - f(t, y)\|_2^2 \leq L_f E\|x - y\|_1^2 \quad (2.6)$$

for all $x, y \in L^2(\mathbb{P}, B_1)$ and for each $t \in \mathbb{R}$. Then for any mean square almost periodic function $\varphi(t) : \mathbb{R} \rightarrow L^2(\mathbb{P}, B_1)$, $f(t, \varphi(t))$ is mean square almost periodic.

In this paper, we always assume that:

- (A1) $\det(I + D_k) \neq 0$ and the sequence $\{D_k\}, k \in \mathbb{N}$, is almost periodic, where $I \in R^{n \times n}$ is the identity matrix;
- (A2) the set of $\{t_k^j\}$ is uniformly almost periodic and $\theta = \inf_k \{t_k^1\} > 0$.

Recall [2], consider the following linear system of system(2.2)

$$\begin{aligned} \dot{x}(t) &= Ax(t), \quad t \neq t_k, \\ \Delta x(t_k) &= D_k x(t_k^-), \quad k \in \mathbb{N}, \end{aligned} \quad (2.7)$$

that if $U_k(t, s)$ is the Cauchy matrix for the system

$$\dot{x}(t) = Ax(t), \quad t_{k-1} \leq t < t_k, \quad (2.8)$$

then the Cauchy matrix for the system (2.7) is in the form

$$W(t, s) = \begin{cases} U_k(t, s), & t_{k-1} \leq s \leq t < t_k, \\ U_{k+1}(t, t_k)(I + D_k)U_k(t_k, s), & t_{k-1} \leq s < t_k \leq t < t_{k+1}, \\ U_{k+1}(t, t_k) \prod_{j=k}^{i+1} (I + D_k)U_j(t_j, t_{j+1})(I + D_i)U_i(t_i, s), & t_{i-1} \leq s < t_i < t_k \leq t < t_{k+1}. \end{cases} \quad (2.9)$$

As the special case of Lemma 1 in [10], we have the following lemma.

Lemma 2.5. Assume that (A1), (A2) and the following condition hold. For the Cauchy matrix $W(t, s)$ of system (2.7), there exist positive constants M and λ such that

$$\|W(t, s)\| \leq Me^{-\lambda(t-s)}, \quad t \geq s, \quad t, s \in \mathbb{R}. \quad (2.10)$$

Then for any $\varepsilon > 0$, $t \geq s$, $t, s \in \mathbb{R}$, $|t - t_k| > \varepsilon$, $|s - t_k| > \varepsilon$, $k \in \mathbb{N}$, there must be exist a relatively dense set T of ε -almost periodic of the matrix A and a positive constant Γ such that for $\tau \in T$, it follows:

$$\|W(t + \tau, s + \tau) - W(t, s)\| \leq \varepsilon \Gamma e^{(-\lambda/2)(t-s)}. \quad (2.11)$$

Lemma 2.6 (see [6]). Let $W(t, s)$ be the Cauchy matrix of the linear system (2.7). Given a constant $\eta \geq \|I + D_k\|$ for all $k \in \mathbb{N}$, if $\eta \geq 1$ and $\theta = \inf_k \{t_k^1\} > 0$, then

$$\|W(t, s)\| \leq \eta e^{(\mu(A) + (\ln \eta / \theta))(t-s)}, \quad t \geq s. \quad (2.12)$$

Introduce the following conditions:

(A3) The functions $f, g : \mathbb{R} \times L^2(\mathbb{P}, \mathbb{H}) \rightarrow L^2(\mathbb{P}, \mathbb{H})$ are mean square almost periodic in $t \in \mathbb{R}$ uniformly in $x \in \Theta$, where $\Theta \subset L^2(\mathbb{P}, \mathbb{H})$ is compact, and $f(0, 0) = g(0, 0) = 0$. Moreover, there exist $L_f, L_g > 0$ such that

$$\begin{aligned} E\|f(t, x) - f(t, y)\|^2 &\leq L_f E\|x - y\|^2, \\ E\|g(t, x) - g(t, y)\|^2 &\leq L_g E\|x - y\|^2, \end{aligned} \quad (2.13)$$

for all stochastic processes $x, y \in L^2(\mathbb{P}, \mathbb{H})$ and $t \in \mathbb{R}$.

(A4) The function $\sigma : \mathbb{R} \times L^2(\mathbb{P}, \mathbb{H}) \rightarrow L^2(\mathbb{P}, \mathbb{H})$ is mean square almost periodic in $t \in \mathbb{R}$ uniformly in $x \in \Theta'$, where $\Theta' \subset L^2(\mathbb{P}, \mathbb{H})$ is compact, and $\sigma(0, 0) = 0$. Moreover, there exists $L_\sigma > 0$ such that

$$E\|\sigma(t, x) - \sigma(t, y)\|^2 \leq L_\sigma E\|x - y\|^2, \quad (2.14)$$

for all stochastic processes $x, y \in L^2(\mathbb{P}, \mathbb{H})$ and $t \in \mathbb{R}$.

(A5) The function $I_i(t) : \mathbb{R} \rightarrow \mathbb{R}$ is almost periodic in the sense of Bohr, $\{\beta_k\}_{k \in \mathbb{N}}$ is almost periodic sequence and there exists a constant $\gamma_0 > 0$, such that

$$\max \left\{ \max_k |\beta_k|, \sup_t \|I(t)\| \right\} \leq \gamma_0. \quad (2.15)$$

(A6) The sequence of functions $V_k(x) : L^2(\mathbb{P}, \mathbb{H}) \rightarrow L^2(\mathbb{P}, \mathbb{H})$ is mean square almost periodic uniformly with respect to $x \in \Theta''$, where $\Theta'' \subset L^2(\mathbb{P}, \mathbb{H})$ is compact. Moreover, there exists $L_V > 0$ such that

$$E\|V_k(x) - V_k(y)\|^2 \leq L_V E\|x - y\|^2 \quad (2.16)$$

for all stochastic processes $x, y \in L^2(\mathbb{P}, \mathbb{H})$.

Lemma 2.7 (see [26]). *If conditions (A1)–(A6) are satisfied, then for each $\varepsilon > 0$, there exists ε_1 , $0 < \varepsilon_1 < \varepsilon$ and relatively dense sets T of real numbers and Q of integral numbers, such that*

- (i) $E\|f(t+\tau, y) - f(t, y)\|^2 < \varepsilon$, $E\|g(t+\tau, y) - g(t, y)\|^2 < \varepsilon$, $t \in \mathbb{R}$, $\tau \in T$, $|t - t_k| > \varepsilon$, $k \in \mathbb{N}$, $y \in L^2(\mathbb{P}, \mathbb{H})$;
- (ii) $E\|\sigma(t+\tau, y) - \sigma(t, y)\|^2 < \varepsilon$, $t \in \mathbb{R}$, $\tau \in T$, $|t - t_k| > \varepsilon$, $k \in \mathbb{N}$, $y \in L^2(\mathbb{P}, \mathbb{H})$;
- (iii) $\|I(t+\tau) - I(t)\|^2 < \varepsilon$, $t \in \mathbb{R}$, $\tau \in T$, $|t - t_k| > \varepsilon$;
- (iv) $E\|V_{k+q}(y) - V_k(y)\|^2 < \varepsilon$, $q \in Q$, $k \in \mathbb{N}$;
- (v) $\|\beta_{k+q} - \beta_k\|^2 < \varepsilon$, $q \in Q$, $k \in \mathbb{N}$;
- (vi) $\|t_{k+q} - \tau\|^2 < \varepsilon_1$, $q \in Q$, $\tau \in T$, $k \in \mathbb{N}$.

Lemma 2.8 (see [26]). *Let condition (A2) holds. Then for each $p > 0$, there exists a positive integer N such that on each interval of length p , there are no more than N elements of the sequence $\{t_k\}$, that is,*

$$i(s, t) \leq N(t - s) + N, \quad (2.17)$$

where $i(s, t)$ is the number of points t_k in the interval (s, t) .

3. Main Results

Theorem 3.1. *Assume that (A1)–(A6) hold, then there exists a unique mean square almost periodic solution of system (2.2) if the following conditions are satisfied: There exists a constant $\eta \geq 1$, such that $\|I + D_k\| \leq \eta$, $k \in \mathbb{N}$ and*

$$\mu(A) + \frac{\ln \eta}{\theta} \triangleq -\lambda < 0. \quad (3.1)$$

Furthermore,

$$\rho = 6\eta^2 \left[\frac{2}{\lambda^2} \left(\|B\|^2 L_f^2 + \|C\|^2 L_g^2 \right) + \frac{N^2}{(1 - e^{-\lambda})^2} L_V^2 + \frac{L_\sigma^2}{2\lambda} \right] < 1. \quad (3.2)$$

Proof. Let $D = \{\varphi(t) \in L^2(\mathbb{P}, \mathbb{H}) : \varphi(t) = (\varphi_1(t), \dots, \varphi_n(t))^T\} \subset AP([-h, +\infty); L^2(\mathbb{P}, \mathbb{H}))$ satisfying the equality $E\|\varphi\|^2 < \bar{K}$, where $\bar{K} = 2\eta^2 \gamma_0^2 ((1/\lambda) + (N/(1 - e^{-\lambda})))^2 > 0$.

Set

$$\begin{aligned} x(t) = & W(t, 0)\phi_0 + \int_0^t W(t, s)[Bf(s, x(s)) + Cg(s, x(s-h)) + I(s)]ds \\ & + \sum_{0 \leq t_k < t} W(t, t_k)[V_k(x(t_k)) + \beta_k] + \int_0^t W(t, s)\sigma(s, x(s))d\omega(s), \quad t \geq 0. \end{aligned} \quad (3.3)$$

where $\phi_0 = x(0)$, it is easy to see that $x(t)$ given by (3.3) is the solution of system (2.2) according to [2] and Lemma 2.2 in [27].

By Lemma 2.6 and the conditions of Theorem, we have

$$\|W(t, s)\| \leq \eta e^{-\lambda(t-s)}, \quad t \geq s, \quad t, s \in \mathbb{R}. \quad (3.4)$$

For $z(t) \in D$, we define the operator L in the following way

$$\begin{aligned} (Lz)(t) = & \int_0^t W(t, s)[Bf(s, z(s)) + Cg(s, z(s-h)) + I(s)]ds \\ & + \sum_{0 \leq t_k < t} W(t, t_k)[V_k(z(t_k)) + \beta_k] + \int_0^t W(t, s)\sigma(s, z(s))d\omega(s). \end{aligned} \quad (3.5)$$

Define subset $D^* \subset D$, $D^* = \{z \in D : E\|z - z_0\|^2 \leq \rho \bar{K}/(1 - \rho)\}$, and $z_0 = \int_0^t W(t, s)I(s)ds + \sum_{0 \leq t_k < t} W(t, t_k)\beta_k$.

We have

$$\begin{aligned} E\|z_0\|^2 & \leq 2E\left\|\int_0^t W(t, s)I(s)ds\right\|^2 + 2E\left\|\sum_{0 \leq t_k < t} W(t, t_k)\beta_k\right\|^2 \\ & \leq 2\left[\int_0^t \eta e^{-\lambda(t-s)} \sup_s \|I(s)\| ds\right]^2 + 2\left[\sum_{0 \leq t_k < t} \eta e^{-\lambda(t-t_k)} \max_k |\beta_k|\right]^2 \\ & \leq 2\eta^2 \gamma_0^2 \left(\frac{1}{\lambda} + \frac{N}{1 - e^{-\lambda}}\right)^2 = \bar{K}. \end{aligned} \quad (3.6)$$

Then for $\forall z \in D^*$, from the definition of D^* and (3.6), since $(a + b)^2 \leq 2a^2 + 2b^2$, we have

$$\begin{aligned} E\|z\|^2 & = E\|(z - z_0) + z_0\|^2 \leq 2E(\|z - z_0\|^2 + \|z_0\|^2) \\ & \leq 2\left(\frac{\rho \bar{K}}{1 - \rho} + \bar{K}\right) = \frac{2\bar{K}}{1 - \rho}. \end{aligned} \quad (3.7)$$

For $\forall z \in D^*$, we have

$$\begin{aligned} \|Lz - z_0\| = & \left\| \int_0^t W(t, s) [Bf(s, z(s)) + Cg(s, z(s-h))] ds \right. \\ & \left. + \sum_{0 \leq t_k < t} W(t, t_k) V_k(z(t_k)) + \int_0^t W(t, s) \sigma(s, z(s)) d\omega(s) \right\|. \end{aligned} \quad (3.8)$$

Since $(a + b + c)^2 \leq 3a^2 + 3b^2 + 3c^2$, it follows

$$\begin{aligned} E\|Lz - z_0\|^2 \leq & 3E \left(\int_0^t \|W(t, s)\| \|Bf(s, z(s)) + Cg(s, z(s-h))\| ds \right)^2 \\ & + 3E \left(\left\| \sum_{0 \leq t_k < t} W(t, t_k) V_k(z(t_k)) \right\| \right)^2 + 3E \left(\int_0^t \|W(t, s) \sigma(s, z(s))\| d\omega(s) \right)^2. \end{aligned} \quad (3.9)$$

For first term of the right-hand side, using (3.7), (A3) and Cauchy-Schwarz inequality, we have

$$\begin{aligned} & E \left(\int_0^t \|W(t, s)\| \|Bf(s, z(s)) + Cg(s, z(s-h))\| ds \right)^2 \\ & \leq \eta^2 \left(\int_0^t e^{-\lambda(t-s)} ds \right) \cdot \left(\int_0^t e^{-\lambda(t-s)} \cdot E \|Bf(s, z(s)) + Cg(s, z(s-h))\|^2 ds \right) \\ & \leq \eta^2 \left(\int_0^t e^{-\lambda(t-s)} ds \right) \cdot \left[\int_0^t e^{-\lambda(t-s)} \cdot \left(2\|B\|^2 L_f^2 E \|z(s)\|^2 + 2\|C\|^2 L_g^2 E \|z(s-h)\|^2 \right) ds \right] \\ & \leq \eta^2 \cdot \frac{2\bar{K}}{1-\rho} \left[\frac{2}{\lambda^2} (\|B\|^2 L_f^2 + \|C\|^2 L_g^2) \right]. \end{aligned} \quad (3.10)$$

As to the second term, using (3.7), (A6) and Cauchy-Schwarz inequality, we can write

$$\begin{aligned} & E \left(\left\| \sum_{0 \leq t_k < t} W(t, t_k) V_k(z(t_k)) \right\| \right)^2 \\ & \leq \eta^2 \left(\sum_{0 \leq t_k < t} e^{-\lambda(t-t_k)} \right) \cdot \left(\sum_{0 \leq t_k < t} e^{-\lambda(t-t_k)} E \|V_k(z(t_k))\|^2 \right) \\ & \leq \eta^2 \left(\sum_{0 \leq t_k < t} e^{-\lambda(t-t_k)} \right) \cdot \left(\sum_{0 \leq t_k < t} e^{-\lambda(t-t_k)} L_V^2 E \|z(t_k)\|^2 \right) \\ & \leq \eta^2 \cdot \frac{2\bar{K}}{1-\rho} \left[L_V^2 \cdot \frac{N^2}{(1-e^{-\lambda})^2} \right]. \end{aligned} \quad (3.11)$$

As far as last term is concerned, using (3.7), (A4), and the Itô isometry theorem, we obtain

$$\begin{aligned} E \left(\int_0^t \|W(t, s) \sigma(s, z(s)) d\omega(s)\| \right)^2 &\leq \int_0^t \|W(t, s)\|^2 E \|\sigma(s, z(s))\|^2 ds \\ &\leq \eta^2 \int_0^t e^{-2\lambda(t-s)} L_\sigma^2 E \|z(s)\|^2 ds \leq \eta^2 \cdot \frac{2\bar{K}}{1-\rho} \cdot \frac{L_\sigma^2}{2\lambda}. \end{aligned} \quad (3.12)$$

Thus, by combining (3.9)–(3.12), it follows that

$$E \|Lz - z_0\|^2 \leq 3\eta^2 \cdot \frac{2\bar{K}}{1-\rho} \left[\frac{2}{\lambda^2} (\|B\|^2 L_f^2 + \|C\|^2 L_g^2) + \frac{N^2}{(1-e^{-\lambda})^2} L_V^2 + \frac{L_\sigma^2}{2\lambda} \right] = \frac{\rho \bar{K}}{1-\rho}. \quad (3.13)$$

By Lemmas 2.5 and 2.6, one can obtain

$$\|W(t + \tau, s + \tau) - W(t, s)\| \leq \varepsilon \Gamma e^{-(\lambda/2)(t-s)}. \quad (3.14)$$

Let $\tau \in T, q \in Q$, where the sets T and Q are determined in Lemma 2.7, and we assume that $0 < \varepsilon < 1$, then

$$\begin{aligned} &\|Lz(t + \tau) - Lz(t)\| \\ &= \left\| \int_0^t [W(t + \tau, s + \tau) - W(t, s)] [Bf(s + \tau, z(s + \tau)) + Cg(s + \tau, z(s + \tau - h)) + I(s + \tau)] ds \right. \\ &\quad + \int_0^t W(t, s) \{ [Bf(s + \tau, z(s + \tau)) + Cg(s + \tau, z(s + \tau - h)) + I(s + \tau)] \\ &\quad \quad \quad \left. - [Bf(s, z(s)) + Cg(s, z(s - h)) + I(s)] \} ds \\ &\quad + \sum_{0 \leq t_k < t} [W(t + \tau, t_{k+q}) - W(t, t_k)] [V_{k+q}(z(t_{k+q})) + \beta_{k+q}] \\ &\quad + \sum_{0 \leq t_k < t} W(t, t_k) [V_{k+q}(z(t_{k+q})) - V_k(z(t_k)) + \beta_{k+q} - \beta_k] \\ &\quad + \int_0^t [W(t + \tau, s + \tau) - W(t, s)] [\sigma(s + \tau, z(s + \tau))] d\omega(s) \\ &\quad \left. + \int_0^t W(t, s) [\sigma(s + \tau, z(s + \tau)) - \sigma(s, z(s))] d\omega(s) \right\}. \end{aligned} \quad (3.15)$$

Therefore, we have

$$\begin{aligned}
& E \|Lz(t+\tau) - Lz(t)\|^2 \\
& \leq 3E \left\| \int_0^t [W(t+\tau, s+\tau) - W(t, s)] [Bf(s+\tau, z(s+\tau)) + Cg(s+\tau, z(s+\tau-h)) + I(s+\tau)] ds \right. \\
& \quad + \int_0^t W(t, s) \{ [Bf(s+\tau, z(s+\tau)) + Cg(s+\tau, z(s+\tau-h)) + I(s+\tau)] \\
& \quad \quad \quad \left. - [Bf(s, z(s)) + Cg(s, z(s-h)) + I(s)] ds \} \right\|^2 \\
& + 3E \left\| \sum_{0 \leq t_k < t} [W(t+\tau, t_{k+q}) - W(t, t_k)] [V_{k+q}(z(t_{k+q})) + \beta_{k+q}] \right. \\
& \quad \left. + \sum_{0 \leq t_k < t} W(t, t_k) [V_{k+q}(z(t_{k+q})) - V_k(z(t_k)) + \beta_{k+q} - \beta_k] \right\|^2 \\
& + 3E \left\| \int_0^t [W(t+\tau, s+\tau) - W(t, s)] [\sigma(s+\tau, z(s+\tau))] d\omega(s) \right. \\
& \quad \left. + \int_0^t W(t, s) [\sigma(s+\tau, z(s+\tau)) - \sigma(s, z(s))] d\omega(s) \right\|^2.
\end{aligned} \tag{3.16}$$

We first evaluate the first term of the right hand side

$$\begin{aligned}
& E \left\| \int_0^t [W(t+\tau, s+\tau) - W(t, s)] [Bf(s+\tau, z(s+\tau)) + Cg(s+\tau, z(s+\tau-h)) + I(s+\tau)] ds \right. \\
& \quad + \int_0^t W(t, s) \{ [Bf(s+\tau, z(s+\tau)) + Cg(s+\tau, z(s+\tau-h)) + I(s+\tau)] \\
& \quad \quad \quad \left. - [Bf(s, z(s)) + Cg(s, z(s-h)) + I(s)] ds \} \right\|^2 \\
& \leq 2E \left[\int_0^t \|W(t+\tau, s+\tau) - W(t, s)\| \right. \\
& \quad \times \|Bf(s+\tau, z(s+\tau)) + Cg(s+\tau, z(s+\tau-h)) + I(s+\tau)\| \left. \right]^2 ds \\
& + 2E \left[\int_0^t \|W(t, s)\| \|B(f(s+\tau, z(s+\tau)) - f(s, z(s))) \right. \\
& \quad \left. + C(g(s+\tau, z(s+\tau-h)) - g(s, z(s-h))) (I(s+\tau) - I(s)) \right] \right\|^2 ds \\
& \leq c_1 \varepsilon,
\end{aligned} \tag{3.17}$$

where $c_1 = (96\eta^2/\lambda^2) [\|B\|^2 L_f^2 \cdot ((\bar{K}/(1-\rho)) + 1) + \|C\|^2 L_g^2 \cdot ((\bar{K}/(1-\rho)) + 1) + \gamma_0^2 + 1]$.

For the second term, we can estimate that

$$\begin{aligned}
& E \left\| \sum_{0 \leq t_k < t} [W(t + \tau, t_{k+q}) - W(t, t_k)] [V_{k+q}(z(t_{k+q})) + \beta_{k+q}] \right. \\
& \quad \left. + \sum_{0 \leq t_k < t} W(t, t_k) [V_{k+q}(z(t_{k+q})) - V_k(z(t_k)) + \beta_{k+q} - \beta_k] \right\|^2 \\
& \leq 2E \left\| \sum_{0 \leq t_k < t} [W(t + \tau, t_{k+q}) - W(t, t_k)] [V_{k+q}(z(t_{k+q})) + \beta_{k+q}] \right\|^2 \\
& \quad + 2E \left\| \sum_{0 \leq t_k < t} W(t, t_k) [V_{k+q}(z(t_{k+q})) - V_k(z(t_k)) + \beta_{k+q} - \beta_k] \right\|^2 \\
& \leq c_2 \varepsilon,
\end{aligned} \tag{3.18}$$

where $c_2 = (8\eta^2 N^2 / (1 - e^{-\lambda})) [L_V^2 \cdot ((\bar{K} / (1 - \rho)) + 1) + \gamma_0^2 + 1]$.

For the last term, using (A4) and Itô isometry identity, we have

$$\begin{aligned}
& E \left\| \int_0^t [W(t + \tau, s + \tau) - W(t, s)] [\sigma(s + \tau, z(s + \tau))] d\omega(s) \right. \\
& \quad \left. + \int_0^t W(t, s) [\sigma(s + \tau, z(s + \tau)) - \sigma(s, z(s))] d\omega(s) \right\|^2 \\
& \leq 2E \left\| \int_0^t [W(t + \tau, s + \tau) - W(t, s), \sigma(s + \tau, z(s + \tau))] d\omega(s) \right\|^2 \\
& \quad + 2E \left\| \int_0^t W(t, s) [\sigma(s + \tau, z(s + \tau)) - \sigma(s, z(s))] d\omega(s) \right\|^2 \\
& \leq 2E \int_0^t \|W(t + \tau, s + \tau) - W(t, s)\|^2 \|\sigma(s + \tau, z(s + \tau))\|^2 ds \\
& \quad + 2E \int_0^t \|W(t, s)\|^2 \|\sigma(s + \tau, z(s + \tau)) - \sigma(s, z(s))\|^2 ds \\
& \leq c_3 \varepsilon,
\end{aligned} \tag{3.19}$$

where $c_3 = (2/\lambda) [\Gamma^2 L_\sigma^2 (\bar{K} / (1 - \rho)) + 1]$.

Combining (3.17), (3.18) and (3.19), it follows that

$$E \|Lz(t + \tau) - Lz(t)\|^2 \leq c_0 \varepsilon, \tag{3.20}$$

where $c_0 = 3(c_1 + c_2 + c_3)$.

So, $Lz \in D^*$, that is L is self-mapping from D^* to D^* by (3.13) and (3.20).

Secondly, we will show L is contracting operator in D^* .

For $\forall x, y \in D^*$,

$$\begin{aligned} \|Lx - Ly\| = & \left\| \int_0^t W(t, s) B[f(s, x(s)) - f(s, y(s))] \right. \\ & + C[g(s, x(s-h)) - g(s, y(s-h))] ds \\ & + \sum_{0 \leq t_k < t} W(t, t_k) [V_k(x(t_k)) - V_k(y(t_k))] \\ & \left. + \int_0^t W(t, s) [\sigma(s, x(s)) - \sigma(s, y(s))] d\omega(s) \right\|. \end{aligned} \quad (3.21)$$

By a minor modification of the proof of (3.13), we can obtain

$$\begin{aligned} E\|Lx - Ly\|^2 & \leq 6\eta^2 \left[\frac{2}{\lambda^2} (\|B\|^2 L_f^2 + \|C\|^2 L_g^2) + \frac{N^2}{(1 - e^{-\lambda})^2} L_V^2 + \frac{L_\sigma^2}{2\lambda} \right] \sup_t E\|x(t) - y(t)\|^2 \\ & = \rho \|x - y\|_\infty^2, \end{aligned} \quad (3.22)$$

and therefore, $\|Lx - Ly\|_\infty \leq \rho \|x - y\|_\infty$, it follows that L is contracting operator in D^* , so there exists a unique mean square almost periodic solution of (2.2) by the fixed points theorem. \square

4. Example

Consider the following impulsive stochastic differential equation with delay

$$\begin{aligned} dx_i(t) = & \left[a_i x_i(t) + \sum_{j=1}^2 b_{ij} f_j(x_j(t)) + \sum_{j=1}^2 c_{ij} g_j(x_j(t-0.1)) + I_i(t) \right] dt \\ & + 0.5 x_i(t) d\omega_i(t), \quad t \geq 0, \quad t \neq t_k, \\ \Delta x(t) = & x(t_k) - x(t_k^-) = D_k x(t_k^-) + V_k(x(t_k^-)) + \beta_k, \quad t = t_k, \quad k \in \mathbb{N}, \\ x(t) = & \phi(t), \quad -h \leq t \leq 0, \end{aligned} \quad (4.1)$$

where $t_k = k$, $k \in \mathbb{N}$, $f(x(t)) = [\sin x_1(t), \sin x_2(t)]^T$, $g(x(t-0.1)) = [\cos x_1(t-0.1), \cos x_2(t-0.1)]^T$, $V_{ik} = [0.01 \sin x_1(t), 0.01 \cos x_2(t)]^T$, $\beta_k = 0.1$, $I(t) = [0.1, 0.1]^T$, $\gamma_0 = 0.1$, for convenience, we can choose

$$A = \begin{bmatrix} -2 & 0 \\ 0 & -3 \end{bmatrix}, \quad B = \begin{bmatrix} -0.1 & 0 \\ 0 & -0.1 \end{bmatrix}, \quad C = \begin{bmatrix} 0.2 & 0 \\ 0 & -0.2 \end{bmatrix}, \quad D_k = \begin{bmatrix} -0.5 & 0 \\ 0 & -0.5 \end{bmatrix}. \quad (4.2)$$

Then $\mu(A) = -2$, $\|I + D_k\| = 1/2$, $\|B\| = 0.1$, $\|C\| = 0.2$, $L_f = L_g = 1$, $L_V = 0.01$, $L_\sigma = 0.5$. Choose $\theta = \inf_k \{t_k^1\} = 0.01$, $\eta = 1$, $N = 6$. By simple calculation, we have $\lambda = -(\mu(A) + (\ln \eta / \theta)) = 2$, $\rho \doteq 0.8139 < 1$, $\bar{K} \doteq 1.107$, $(\rho \bar{K} / (1 - \rho)) \doteq 4.841$.

Let $D^* = \{z \in D : E\|z - z_0\|^2 \leq 4.841\}$, so, by Theorem 3.1, system (4.1) has a unique mean square almost periodic solution in D^* .

Remark 4.1. Since there exist no results for almost periodic solutions for impulsive stochastic differential equations with delays, one can easily see that all the results in [10, 11, 20–22, 28] and the references therein cannot be applicable to system (4.1). This implies that the results of this paper are essentially new.

5. Conclusion

In this paper, a class of Itô impulsive stochastic differential equations with delays has been investigated. We conquer the difficulty of coexistence of impulsive, delay and stochastic factors in a dynamic system, and give a result for the existence and uniqueness of mean square almost periodic solutions. The results in this paper extend some earlier works reported in the literature. Moreover, our results have important applications in almost periodic oscillatory stochastic delayed neural networks with impulsive control.

Acknowledgment

This work is supported by the National Science Foundation of China (no. 10771199).

References

- [1] A. M. Samoilenko and N. A. Perestyuk, *Impulsive Differential Equations*, World Scientific, Singapore, 1995.
- [2] D. D. Bainov and P. S. Simeonov, *Theory of Impulsive Differential Equations: Periodic Solutions and Applications*, Longman, Harlow, UK, 1993.
- [3] V. Lakshmikantham, D. D. Bainov, and P. S. Simeonov, *Theory of Impulsive Differential Equations*, World Scientific, Singapore, 1989.
- [4] X. Liu, "Stability of impulsive control systems with time delay," *Mathematical and Computer Modelling*, vol. 39, no. 4-5, pp. 511–519, 2004.
- [5] Z. Yang and D. Xu, "Existence and exponential stability of periodic solution for impulsive delay differential equations and applications," *Nonlinear Analysis*, vol. 64, no. 1, pp. 130–145, 2006.
- [6] Z. Yang and D. Xu, "Robust stability of uncertain impulsive control systems with time-varying delay," *Computers and Mathematics with Applications*, vol. 53, no. 5, pp. 760–769, 2007.
- [7] X. Fu and X. Li, "Global exponential stability and global attractivity of impulsive Hopfield neural networks with time delays," *Journal of Computational and Applied Mathematics*, vol. 231, no. 1, pp. 187–199, 2009.
- [8] C. Li, J. Sun, and R. Sun, "Stability analysis of a class of stochastic differential delay equations with nonlinear impulsive effects," *Journal of the Franklin Institute*, vol. 347, no. 7, pp. 1186–1198, 2010.
- [9] R. P. Agarwal and F. Karako, "A survey on oscillation of impulsive delay differential equations," *Computers and Mathematics with Applications*, vol. 60, no. 6, pp. 1648–1685, 2010.
- [10] G. T. Stamov and I. M. Stamova, "Almost periodic solutions for impulsive neural networks with delay," *Applied Mathematical Modelling*, vol. 31, no. 7, pp. 1263–1270, 2007.
- [11] G. T. Stamov and J. O. Alzabut, "Almost periodic solutions for abstract impulsive differential equations," *Nonlinear Analysis*, vol. 72, no. 5, pp. 2457–2464, 2010.
- [12] X. Mao, *Exponential Stability of Stochastic Differential Equations*, vol. 182, Marcel Dekker, New York, NY, USA, 1994.
- [13] L. Arnold, *Stochastic Differential Equations: Theory and Applications*, Wiley, New York, NY, USA, 1972.
- [14] A. Friedman, *Stochastic Differential Equations and Applications*, Academic Press, New York, NY, USA, 1976.

- [15] Z. Gao and X. Shi, "Stochastic state estimation and control for stochastic descriptor systems," in *Proceedings of the IEEE International Conference on Robotics, Automation and Mechatronics (RAM '08)*, pp. 331–336, 2008.
- [16] Z. Gao and S. X. Ding, "Actuator fault robust estimation and fault-tolerant control for a class of nonlinear descriptor systems," *Automatica*, vol. 43, no. 5, pp. 912–920, 2007.
- [17] L. Ma and F. Da, "Mean-square exponential stability of stochastic Hopfield neural networks with time-varying discrete and distributed delays," *Physics Letters A*, vol. 373, no. 25, pp. 2154–2161, 2009.
- [18] L. Hu and A. Yang, "Fuzzy model-based control of nonlinear stochastic systems with time-delay," *Nonlinear Analysis*, vol. 71, no. 12, pp. e2855–e2865, 2009.
- [19] H. Zhang, H. Yan, and Q. Chen, "Stability and dissipative analysis for a class of stochastic system with time-delay," *Journal of the Franklin Institute*, vol. 347, no. 5, pp. 882–893, 2010.
- [20] P. H. Bezandry and T. Diagana, "Existence of almost periodic solutions to some stochastic differential equations," *Applicable Analysis*, vol. 86, no. 7, pp. 819–827, 2007.
- [21] P. H. Bezandry, "Existence of almost periodic solutions to some functional integro-differential stochastic evolution equations," *Statistics and Probability Letters*, vol. 78, no. 17, pp. 2844–2849, 2008.
- [22] Z. Huang and Q. G. Yang, "Existence and exponential stability of almost periodic solution for stochastic cellular neural networks with delay," *Chaos, Solitons and Fractals*, vol. 42, no. 2, pp. 773–780, 2009.
- [23] Z. Yang, D. Xu, and L. Xiang, "Exponential -stability of impulsive stochastic differential equations with delays," *Physics Letters A*, vol. 359, no. 2, pp. 129–137, 2006.
- [24] L. Xu and D. Xu, "Exponential -stability of impulsive stochastic neural networks with mixed delays," *Chaos, Solitons and Fractals*, vol. 41, no. 1, pp. 263–272, 2009.
- [25] C. Li and J. Sun, "Stability analysis of nonlinear stochastic differential delay systems under impulsive control," *Physics Letters A*, vol. 374, no. 9, pp. 1154–1158, 2010.
- [26] A. M. Samoilenko and N. A. Perestyuk, *Differential Equations with Impulse Effect*, Vyshcha Shkola, Kiev, Russia, 1987.
- [27] L. Xu and D. Xu, "Mean square exponential stability of impulsive control stochastic systems with time-varying delay," *Physics Letters A*, vol. 373, no. 3, pp. 328–333, 2009.
- [28] X. Li, "Existence and global exponential stability of periodic solution for delayed neural networks with impulsive and stochastic effects," *Neurocomputing*, vol. 73, no. 4–6, pp. 749–758, 2010.

Research Article

Bank Liquidity and the Global Financial Crisis

**Frednard Gideon,¹ Mark A. Petersen,²
Janine Mukuddem-Petersen,³ and Bernadine De Waal²**

¹ Department of Mathematics, Faculty of Science, University of Namibia, Private Bag 13301, Windhoek 9000, Namibia

² Research Division, Faculty of Commerce and Administration, North-West University, Private Bag x2046, Mmabatho 2735, South Africa

³ Economics Division, Faculty of Commerce and Administration, North-West University, Private Bag x2046, Mmabatho 2735, South Africa

Correspondence should be addressed to Frednard Gideon, tewaadha@yahoo.com

Received 2 November 2011; Revised 22 January 2012; Accepted 5 February 2012

Academic Editor: Chuanhou Gao

Copyright © 2012 Frednard Gideon et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We investigate the stochastic dynamics of bank liquidity parameters such as liquid assets and nett cash outflow in relation to the global financial crisis. These parameters enable us to determine the liquidity coverage ratio that is one of the metrics used in ratio analysis to measure bank liquidity. In this regard, numerical results show that bank behavior related to liquidity was highly procyclical during the financial crisis. We also consider a theoretical-quantitative approach to bank liquidity provisioning. In this case, we provide an explicit expression for the aggregate liquidity risk when a locally risk-minimizing strategy is utilized.

1. Introduction

During the global financial crisis (GFC), banks were under severe pressure to maintain adequate liquidity. In general, empirical evidence shows that banks with sufficient liquidity can meet their payment obligations while banks with low liquidity cannot. The GFC highlighted the fact that liquidity risk can proliferate quickly with funding sources dissipating and concerns about asset valuation and capital adequacy realizing. This situation underscores the important relationship between funding risk (involving raising funds to bankroll asset holdings) and market liquidity (involving the efficient conversion of assets into liquid funds at a given price). In response to this, the Basel Committee on Banking Supervision (BCBS) is attempting to develop an international framework for liquidity risk measurement, standards, and monitoring (see, e.g., [1]). Although pre-Basel III regulation

establishes procedures for assessing credit, market, and operational risk, it does not provide effective protocols for managing liquidity and systemic risks. The drafting of Basel III represents an effort to address the latter (see, e.g., [2–4]).

Current liquidity risk management procedures can be classified as micro- or macroprudential. In the case of the former, simple liquidity ratios such as credit-to-deposit ratios (nett stable funding ratios), liquidity coverage ratios and the assessment of the gap between short-term liabilities and assets are appropriate to cover the objectives of bank balance sheet analysis. The ratio approach for liquidity risk management is a quantitative international accepted standard for alerting banks about any possible adverse economic downturns. For instance, the credit-to-deposit ratio assesses the relationships between sources and uses of funds held in the bank's portfolio but has limitations which ultimately do not reflect information on market financing with short-term maturity. By contrast, the liquidity coverage ratio (LCR) performs better by ensuring the coverage of some of the immediate liabilities. Since the LCR depends only on bank balance sheet data, it does not take into account the residual maturities on various uses and sources of funds. Also, in a global context, a quantitative approach may not take financial market conditions into account. In this case, a more comprehensive characterization of the bank system's liquidity risk through designed stress testing and constructed contingency plans is considered. The Basel Committee on Banking Supervision suggested best practices related to international liquidity standards. In this case, a well-designed policy monitoring instrument to measure and regulate the dynamics of foreign currency is considered to best take financial market conditions into account. Also, central banks (CBs) have a pivotal role to play in managing liquidity inflows via macroeconomic management of exchange rate and interest rate responses. The modeling of capital markets as well as stock and bond behavior also contribute to the liquidity response for possible stress conditions observed. The above approaches for liquidity analysis take into account the macroprudential liquidity management of banks.

In this paper, in Section 2, we discuss balance sheet items related to liquid assets and nett cash outflow in order to build a stochastic LCR model. Before the GFC, banks were prosperous with high LCRs, high cash inflows, low interest rates, and low nett cash outflows. This was followed by the collapse of liquidity, exploding default rates, and the effects thereof during the GFC. Next, in Section 3, we apply a dynamic provisioning strategy to liquidity risk management. In this case, we address the problem of dynamic liquidity provisioning for a mortgage, Λ , which is an underlying (illiquid) nonmarketable asset, by substituting (liquid) marketable securities, S . In the light of the above, banks prefer to trade in a Treasury bond market because of liquidity reasons. Since the loan process $(\Lambda_t)_{0 \leq t \leq T}$ is not completely correlated with the substitute, it creates the market incompleteness. In other words, we will employ non-self-financing strategy to replicate the trading process. Therefore the banks would require that the uncertainty involved over the remaining of the trading period be minimized. In this case, we specifically minimize at each date, the uncertainty over the next infinitesimal period. In the dynamics trading there is always a residual risk emanating from the imperfection of the correlation between the Brownian motions. Due to the no-arbitrage opportunities there are infinitely many equivalent martingale measures so that pricing is directly linked to risk. Therefore, we choose a pricing candidate (equivalent martingale measure) under which the discounted stock price follows a martingale. This equivalent measure is chosen according to a provisioning strategy which ensures that the value of Λ is the value of the replicating portfolio. We also provide a framework for assessing residual aggregate liquidity risk stemming from the application of the above strategy.

1.1. Literature Review

The documents formulated in response to the proposed Basel III regulatory framework are among the most topical literature on bank liquidity (see, e.g., [1]). During the GFC, unprecedented levels of liquidity support were required from CBs in order to sustain the financial system and even with such extensive support a number of banks failed, were forced into mergers, or required resolution. The crisis illustrated how quickly and severely liquidity risks can crystallize and certain sources of funding can evaporate, compounding concerns related to the valuation of assets and capital adequacy (see, e.g., [2–4]). A key characteristic of the GFC was the inaccurate and ineffective management of liquidity risk. In recognition of the need for banks to improve their liquidity risk management and control their exposures to such risk, the BCBS issued Principles for Sound Liquidity Risk Management and Supervision in September 2008 (see, e.g., [1]). Supervisors are expected to assess both the adequacy of a bank's liquidity risk management framework and its liquidity risk exposure. In addition, they are required to take prompt action to address the banks risk management deficiencies or excess exposure in order to protect depositors and enhance the overall stability of the financial system. To reinforce these supervisory objectives and efforts, the BCBS has recently focussed on further elevating the resilience of internationally active banks to liquidity stresses across the globe, as well as increasing international harmonization of liquidity risk supervision (see, e.g., [1]). The BCBS hopes to develop internationally consistent regulatory standards for liquidity risk supervision as a cornerstone of a global framework to strengthen liquidity risk management and supervision (see, e.g., [2–4]).

In [5] it is asserted that bank liquidity behavior can be described by straightforward indicators constructed from firm-specific balance sheet data (see, also, [6, 7]). Also, their analysis underscores the relevance of using several indicators of liquidity risk at the same time, given the different leads and lags of the measures with systemic risk. Our study is related to theirs in that we make use of balance sheet items to determine bank behavior. Another similarity is that we make use of data from [6] to formulate conclusions in a numerical quantitative framework (compare with the analysis in Section 3 below).

The contribution [8] studies the role of securitization in bank management. A new index of “bank loan portfolio liquidity” which can be thought of as a weighted average of the potential to securitize loans of a given type, where the weights reflect the composition of a bank loan portfolio. The paper uses this new index to show that by allowing banks to convert illiquid loans into liquid funds, securitization reduces banks holdings of liquid securities and increases their lending ability. Furthermore, securitization provides banks with an additional source of funding and makes bank lending less sensitive to cost of funds shocks. By extension, the securitization weakens the ability of regulators to affect banks lending activity but makes banks more susceptible to liquidity and funding crisis when the securitization market is shutdown. We conduct a similar analysis in Section 4 of this paper where illiquid underlying loans are substituted by liquid marketable securities.

In [9], we use actuarial methods to solve a nonlinear stochastic optimal liquidity risk management problem for subprime originators with deposit inflow rates and marketable securities allocation as controls (see [10]). The main objective is to minimize liquidity risk in the form of funding and credit crunch risk in an incomplete market. In order to accomplish this, we construct a stochastic model that incorporates originator mortgage and deposit reference processes. Finally, numerical examples that illustrate the main modeling and optimization features of the paper are provided. Our work in this paper also has a connection with [9] in that the nexus between funding risk and market liquidity is explored.

However, this paper is an improvement on the aforementioned in that bank balance sheet features play a more prominent role (see, Sections 2, 3, and 4).

1.2. Main Questions and Article Outline

In this subsection, we pose the main questions and provide an outline of the paper.

1.2.1. Main Questions

In this paper on bank liquidity, we answer the following questions.

Question 1 (banking model). Can we model banks' liquid assets and nett cash outflows as well as LCRs in a stochastic framework? (compare with Section 2).

Question 2 (bank liquidity in a numerical quantitative framework). Can we explain and provide numerical examples of bank liquidity dynamics? (refer to Section 3).

Question 3 (bank liquidity in a theoretical quantitative framework). Can we devise a liquidity provisioning strategy in a theoretical quantitative framework? (compare with Section 4).

1.2.2. Paper Outline

The rest of the paper is organized as follows. Section 1 introduces the concept of liquidity risk while providing an appropriate literature review. A stochastic LCR model for bank liquidity is constructed in Section 2. Issues pertaining to bank liquidity in a numerical quantitative framework are discussed in Section 3. Section 4 treats liquidity in a theoretical quantitative manner. Finally, we provide concluding remarks in Section 5.

2. Bank Liquidity Model

In the sequel, we use the notational convention “subscript t or s ” to represent (possibly) random processes, while “bracket t or s ” is used to denote deterministic processes. The assessment of a bank's relative composition of the stock of high-quality liquid assets (liquid assets) and nett cash outflows, is one of the primary ways of analyzing its liquidity position. In this regard, we consider a measure of liquidity offered by the LCR. Before the GFC, banks were prosperous with high LCRs, high cash inflows, low interest rates, and low nett cash outflows. This was followed by the collapse of liquidity, exploding default rates, and the effects thereof. We make the following assumption to set the space and time index that we consider in our LCR model.

Assumption 2.1 (filtered probability space and time index). Throughout, we assume that we are working with a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $\{\mathcal{F}_t\}_{t \geq 0}$ on a time index set $[0, T]$. We assume that the aforementioned space satisfies the usual conditions. Under \mathbb{P} , $\{W_t; 0 \leq t \leq T, W_0 = 0\}$ is an \mathcal{F}_t -Brownian motion.

Furthermore, we are able to produce a system of stochastic differential equations that provide information about the stock of high-quality liquid assets (liquid assets) at time t with

$x^1 : \Omega \times [0, T] \rightarrow \mathbb{R}^+$ denoted by x_t^1 and nett cash outflows at time t with $x^2 : \Omega \times [0, T] \rightarrow \mathbb{R}^+$ denoted by x_t^2 and their relationship. The dynamics of liquid assets, x_t^1 , is stochastic in nature because it depends in part on the stochastic rates of return on assets and cash inflow and outflow (see [9] for more details) and the securitization market. Also, the dynamics of the nett cash outflow, x_t^2 , is stochastic because its value has a reliance on cash inflows as well as liquidity and market risk that have randomness associated with them. Furthermore, for $x : \Omega \times [0, T] \rightarrow \mathbb{R}^2$ we use the notation x_t to denote

$$x_t = \begin{bmatrix} x_t^1 \\ x_t^2 \end{bmatrix}, \quad (2.1)$$

and represent the LCR with $l : \Omega \times [0, T] \rightarrow \mathbb{R}^+$ by

$$l_t = \frac{x_t^1}{x_t^2}. \quad (2.2)$$

It is important for banks that l_t in (2.2) has to be sufficiently high to ensure high bank liquidity.

2.1. Liquid Assets

In this section, we discuss the stock of high-quality liquid assets constituted by cash, CB reserves, marketable securities, and government/CB bank debt issued.

2.1.1. Description of Liquid Assets

The first component of stock of high-quality liquid assets is *cash* that is made up of banknotes and coins. According to [1], a *CB reserve* should be able to be drawn down in times of stress. In this regard, local supervisors should discuss and agree with the relevant CB the extent to which CB reserves should count toward the stock of liquid assets.

Marketable securities represent claims on or claims guaranteed by sovereigns, CBs, noncentral government public sector entities (PSEs), the Bank for International Settlements (BIS), the International Monetary Fund (IMF), the European Commission (EC), or multilateral development banks. This is conditional on all the following criteria being met. These claims are assigned a 0% risk weight under the Basel II standardized approach. Also, deep repo-markets should exist for these securities and that they are not issued by banks or other financial service entities.

Another category of stock of high-quality liquid assets refers to *government/CB bank debt issued in domestic currencies* by the country in which the liquidity risk is being taken by the bank's home country (see, e.g., [1, 4]).

2.1.2. Dynamics of Liquid Assets

In this section, we consider

$$dh_t = r_t^h dt + \sigma_t^h dW_t^h, \quad h(t_0) = h_0, \quad (2.3)$$

where the stochastic processes $h : \Omega \times [0, T] \rightarrow \mathbb{R}^+$ are the *return per unit of liquid assets*, $r^h \rightarrow \mathbb{R}^+$ is the rate of return per liquid asset unit, the scalar $\sigma^h : T \rightarrow \mathbb{R}$ is the volatility in the rate of returns, and $W^h : \Omega \times [0, T] \rightarrow \mathbb{R}$ is standard Brownian motion. Before the GFC, risky asset returns were much higher than those of riskless assets, making the former a more attractive but much riskier investment. It is inefficient for banks to invest all in risky or riskless securities with asset allocation being important. In this regard, it is necessary to make the following assumption to distinguish between risky (e.g., marketable securities and government/CB bank debt) and riskless assets (cash) for future computations.

Assumption 2.2 (liquid assets). Suppose from the outset that liquid assets are held in the financial market with $n + 1$ asset classes. One of these assets is riskless (cash) while the assets $1, 2, \dots, n$ are risky.

The risky liquid assets evolve continuously in time and are modelled using an n -dimensional Brownian motion. In this multidimensional context, the *asset returns in the k th liquid asset class per unit of the k th class* is denoted by $y_t^k, k \in \mathbb{N}_n = \{0, 1, 2, \dots, n\}$ where $y : \Omega \times [0, T] \rightarrow \mathbb{R}^{n+1}$. Thus, the return per liquid asset unit is

$$y = (C(t), y_t^1, \dots, y_t^n), \quad (2.4)$$

where $C(t)$ represents the return on cash and y_t^1, \dots, y_t^n represents the risky return. Furthermore, we can model y as

$$dy_t = r_t^y dt + \Sigma_t^y dW_t^y, \quad y(t_0) = y_0, \quad (2.5)$$

where $r^y : T \rightarrow \mathbb{R}^{n+1}$ denotes the rate of liquid asset returns, $\Sigma_t^y \in \mathbb{R}^{(n+1) \times n}$ is a matrix of liquid asset returns, and $W^y : \Omega \times [0, T] \rightarrow \mathbb{R}^n$ is standard Brownian motion. Notice that there are only n scalar Brownian motions due to one of the liquid assets being riskless.

We assume that the investment strategy $\pi : T \rightarrow \mathbb{R}^{n+1}$ is outside the simplex

$$S = \left\{ \pi \in \mathbb{R}^{n+1} : \pi = (\pi^0, \dots, \pi^n)^T, \pi^0 + \dots + \pi^n = 1, \pi^0 \geq 0, \dots, \pi^n \geq 0 \right\}. \quad (2.6)$$

In this case, short selling is possible. The *liquid asset return* is then $h : \Omega \times R \rightarrow \mathbb{R}^+$, where the dynamics of h can be written as

$$dh_t = \pi_t^T dy_t = \pi_t^T r_t^y dt + \pi_t^T \Sigma_t^y dW_t^y. \quad (2.7)$$

This notation can be simplified as follows. We denote

$$\begin{aligned}
r^C(t) &= r^{y^0}(t), r^C : T \longrightarrow \mathbb{R}^+, \text{ the rate of return on cash,} \\
r_t^y &= \left(r^C(t), \widetilde{r}_t^{yT} + r^C(t)1_n \right)^T, \quad \widetilde{r}^y : T \longrightarrow \mathbb{R}^n, \\
\pi_t &= \left(\pi_t^0, \widetilde{\pi}_t^T \right)^T = \left(\pi_t^0, \pi_t^1, \dots, \pi_t^k \right)^T, \quad \widetilde{\pi} : T \longrightarrow \mathbb{R}^k, \\
\Sigma_t^y &= \begin{pmatrix} 0 & \cdots & 0 \\ & \widetilde{\Sigma}_t^y & \end{pmatrix}, \quad \widetilde{\Sigma}_t^y \in \mathbb{R}^{n \times n}, \\
\widetilde{C}_t &= \widetilde{\Sigma}_t^y \widetilde{\Sigma}_t^{yT}. \text{ Then, we have that} \\
\pi_t^T r_t^y &= \pi_t^0 r^C(t) + \widetilde{\pi}_t^{jT} \widetilde{r}_t^y + \widetilde{\pi}_t^{jT} r^C(t)1_n = r^C(t) + \widetilde{\pi}_t^T \widetilde{r}_t^y, \\
\pi_t^T \Sigma_t^y dW_t^y &= \widetilde{\pi}_t^T \widetilde{\Sigma}_t^y dW_t^y, \\
dh_t &= \left[r^C(t) + \widetilde{\pi}_t^T \widetilde{r}_t^y \right] dt + \widetilde{\pi}_t^T \widetilde{\Sigma}_t^y dW_t^y, \quad h(t_0) = h_0.
\end{aligned} \tag{2.8}$$

2.2. Nett Cash Outflows

In this section, we discuss nett cash outflows arising from cash outflows and inflows.

2.2.1. Description of Nett Cash Outflows

Cash outflows are constituted by retail deposits, unsecured wholesale funding secured funding and additional liabilities (see, e.g., [1]). The latter category includes requirements about liabilities involving derivative collateral calls related to a downgrade of up to 3 notches, market valuation changes on derivatives transactions, valuation changes on posted noncash or non-high-quality sovereign debt collateral securing derivative transactions, asset backed commercial paper (ABCP), special investment vehicles (SIVs), conduits, special purpose vehicles (SPVs), and the currently undrawn portion of committed credit and liquidity facilities.

Cash inflows are made up of amounts receivable from retail counterparties, amounts receivable from wholesale counterparties, receivables in respect of repo and reverse repo transactions backed by illiquid assets, and securities lending/borrowing transactions where illiquid assets are borrowed as well as other cash inflows.

According to [1], *nett cash inflows* is defined as cumulative expected cash outflows minus cumulative expected cash inflows arising in the specified stress scenario in the time period under consideration. This is the nett cumulative liquidity mismatch position under the stress scenario measured at the test horizon. *Cumulative expected cash outflows* are calculated by multiplying outstanding balances of various categories or types of liabilities by assumed percentages that are expected to roll off and by multiplying specified draw-down amounts to various off-balance sheet commitments. *Cumulative expected cash inflows* are calculated by multiplying amounts receivable by a percentage that reflects expected inflow under the stress scenario.

2.2.2. Dynamics of Nett Cash Outflows

Essentially, mortgagors are free to vary their cash inflow rates. Roughly speaking, this rate should be reduced for high LCRs and increased beyond the normal rate when LCRs are too low. In the sequel, the stochastic process $u^1 : \Omega \times [0, T] \rightarrow \mathbb{R}^+$ is the *normal cash inflow rate per nett cash inflow unit* whose value at time t is denoted by u_t^1 . In this case, $u_t^1 dt$ turns out to be the cash inflow rate per unit of the nett cash inflow over the time period $(t, t + dt)$. A notion related to this is the *adjustment to the cash inflow rate per unit of the nett cash inflow rate for a higher or lower LCR*, $u^2 : \Omega \times [0, T] \rightarrow \mathbb{R}^+$, that will in closed loop be made dependent on the LCR. We denote the sum of u^1 and u^2 by the *cash inflow rate* $u^3 : \Omega \times [0, T] \rightarrow \mathbb{R}^+$, that is,

$$u_t^3 = u_t^1 + u_t^2, \quad \forall t. \quad (2.9)$$

Before the GFC, the cash inflow rate increased significantly as a consequence of rising liquidity. The following assumption is made in order to model the LCR in a stochastic framework.

Assumption 2.3 (cash inflow rate). The cash inflow, u^3 , is predictable with respect to $\{\mathcal{F}_t\}_{t \geq 0}$.

The cash inflow provides us with a means of controlling LCR dynamics. The dynamics of the *cash outflow per unit of the nett cash outflow*, $e : \Omega \times [0, T] \rightarrow \mathbb{R}$, is given by

$$de_t = r_t^e dt + \sigma_t^e dW_t^e, \quad e(t_0) = e_0, \quad (2.10)$$

where e_t is the cash outflow per unit of the nett cash outflow, $r^e : T \rightarrow \mathbb{R}$ is the rate of outflow per unit of the nett cash outflow, the scalar $\sigma^e : T \rightarrow \mathbb{R}$ is the volatility in the outflow per nett cash outflow unit, and $W^e : \Omega \times [0, T] \rightarrow \mathbb{R}$ is standard Brownian motion.

Next, we take $i : \Omega \times [0, T] \rightarrow \mathbb{R}^+$ as the *nett cash outflow increase before cash outflow per monetary unit of the nett cash outflow*, $r^i : T \rightarrow \mathbb{R}^+$ is the rate of increase of nett cash outflows before cash outflow per nett cash outflow unit, the scalar $\sigma^i \in \mathbb{R}$ is the volatility in the increase of nett cash outflows before outflow, and $W^i : \Omega \times [0, T] \rightarrow \mathbb{R}$ represents standard Brownian motion. Then, we set

$$di_t = r_t^i dt + \sigma^i dW_t^i, \quad i(t_0) = i_0. \quad (2.11)$$

The stochastic process i_t in (2.11) may typically originate from nett cash flow volatility that may result from changes in market activity, cash supply, and inflation.

2.3. The Liquidity Coverage Ratio

This section discusses ratio analysis and liquidity coverage ratio dynamics.

2.3.1. Ratio Analysis

Ratio analysis is conducted on the bank's balance sheet composition. In this case, the LCR measures a bank's ability to access funding for a 30-day period of acute market stress. In this paper, as in Basel III, we are interested in the LCR that is defined as the sum of interbank

assets and securities issued by public entities as a percentage of interbank liabilities. The LCR formula is given by

$$\text{Liquidity Coverage Ratio} = \frac{\text{Stock of High Quality Liquid Assets}}{\text{Nett Cash Outflows over a 30-day Period}}. \quad (2.12)$$

This ratio measures the bank system's liquidity position that allows the assessment of a bank's capacity to ensure the coverage of some of its more immediate liabilities with highly available assets. It also identifies the amount of unencumbered, high-quality liquid assets a bank holds that can be used to offset the nett cash outflows it would encounter under a short-term stress scenario specified by supervisors, including both specific and systemic shocks.

2.3.2. Liquidity Coverage Ratio Dynamics

Using the equations for liquid assets x^1 and nett cash outflow x^2 , we have that

$$\begin{aligned} dx_t^1 &= x_t^1 dh_t + x_t^2 u_t^3 dt - x_t^2 de_t \\ &= \left[r^C(t) x_t^1 + x_t^1 \tilde{\pi}_t^T \tilde{r}_t^y + x_t^2 u_t^1 + x_t^2 u_t^2 - x_t^2 r_t^e \right] dt \\ &\quad + \left[x_t^1 \tilde{\pi}_t^T \tilde{\Sigma}_t^y dW_t^y - x_t^2 \sigma^e dW_t^e \right], \\ dx_t^2 &= x_t^2 di_t - x_t^2 de_t \\ &= x_t^2 \left[r_t^i dt + \sigma^i dW_t^i \right] - x_t^2 \left[r_t^e dt + \sigma^e dW_t^e \right] \\ &= x_t^2 \left[r_t^i - r_t^e \right] dt + x_t^2 \left[\sigma^i dW_t^i - \sigma^e dW_t^e \right]. \end{aligned} \quad (2.13)$$

The SDEs (2.13) may be rewritten into matrix-vector form in the following way.

Definition 2.4 (stochastic system for the LCR model). Define the stochastic system for the LCR model as

$$dx_t = A_t x_t dt + N(x_t) u_t dt + a_t dt + S(x_t, u_t) dW_t, \quad (2.14)$$

with the various terms in this stochastic differential equation being

$$\begin{aligned}
 u_t &= \begin{bmatrix} u_t^2 \\ \tilde{\pi}_t \end{bmatrix}, \quad u : \Omega \times [0, T] \longrightarrow \mathbb{R}^{n+1}, \\
 A_t &= \begin{bmatrix} r^C(t) & -r_t^e \\ 0 & r_t^i - r_t^e \end{bmatrix}, \\
 N(x_t) &= \begin{bmatrix} x_t^2 & x_t^1 \tilde{r}_t^{yT} \\ 0 & 0 \end{bmatrix}, \quad a_t = \begin{bmatrix} x_t^2 u_t^1 \\ 0 \end{bmatrix}, \\
 S(x_t, u_t) &= \begin{bmatrix} x_t^1 \tilde{\pi}_t^T \tilde{\Sigma}_t^y & -x_t^2 \sigma^e & 0 \\ 0 & -x_t^2 \sigma^e & x_t^2 \sigma^i \end{bmatrix}, \\
 W_t &= \begin{bmatrix} W_t^y \\ W_t^e \\ W_t^i \end{bmatrix},
 \end{aligned} \tag{2.15}$$

where W_t^y, W_t^e , and W_t^i are mutually (stochastically) independent standard Brownian motions. It is assumed that for all $t \in T$, $\sigma_t^e > 0$, $\sigma_t^i > 0$ and $\tilde{C}_t > 0$. Often the time argument of the functions σ^e and σ^i is omitted.

We can rewrite (2.14) as follows:

$$\begin{aligned}
 N(x_t)u_t &:= \begin{bmatrix} x_t^2 \\ 0 \end{bmatrix} u_t^2 + \begin{bmatrix} x_t^1 \tilde{r}_t^{yT} \\ 0 \end{bmatrix} \tilde{\pi}_t \\
 &:= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x_t u_t^3 + \sum_{m=1}^n \begin{bmatrix} x_t^1 \tilde{r}_t^{y,m} \\ 0 \end{bmatrix} \tilde{\pi}_t^m \\
 &:= B_0 x_t u_t^0 + \sum_{m=1}^n \begin{bmatrix} \tilde{r}_t^{y,m} & 0 \\ 0 & 0 \end{bmatrix} x_t \tilde{\pi}_t^m \\
 &:= \sum_{m=0}^n [B_m x_t] u_t^m, \\
 S(x_t, u_t) dW_t &= \begin{bmatrix} [\tilde{\pi}_t^T \tilde{C}_t \tilde{\pi}_t]^{1/2} & 0 \\ 0 & 0 \end{bmatrix} x_t dW_t^1 \\
 &\quad + \begin{bmatrix} 0 & -\sigma^e \\ 0 & -\sigma^e \end{bmatrix} x_t dW_t^2 + \begin{bmatrix} 0 & 0 \\ 0 & \sigma^i \end{bmatrix} x_t dW_t^3 \\
 &= \sum_{j=1}^3 [M^{jj}(u_t) x_t] dW_t^{jj},
 \end{aligned} \tag{2.16}$$

where B and M are only used for notational purposes to simplify the equations. From the stochastic system given by (2.14) it is clear that $u = (u^2, \tilde{\pi})$ affects only the stochastic differential equation of x_t^1 but not that of x_t^2 . In particular, for (2.14) we have that $\tilde{\pi}$ affects

the variance of x_t^1 and the drift of x_t^1 via the term $x_t^1 \widetilde{r}_t^y \widetilde{\pi}_t$. On the other hand, u^2 affects only the drift of x_t^1 . Then (2.14) becomes

$$dx_t = A_t x_t dt + \sum_{j=0}^n [B^j x_t] u_t^j dt + a_t dt + \sum_{j=1}^3 [M^j(u_t) x_t] dW_t^{jj}. \quad (2.17)$$

The model can be simplified if attention is restricted to the system with the LCR, as stated earlier, denoted in this section by $x_t = x_t^1 / x_t^2$.

Definition 2.5 (stochastic model for a simplified LCR). Define the simplified LCR system by the SDE

$$\begin{aligned} dx_t = x_t & \left[r^C(t) + r_t^e - r_t^i + (\sigma^e)^2 + (\sigma^i)^2 + \widetilde{r}_t^y \widetilde{\pi}_t \right] dt \\ & + [u_t^1 + u_t^2 - r_t^e - (\sigma^e)^2] dt \\ & + \left[(\sigma^e)^2 (1 - x_t)^2 + (\sigma^i)^2 x_t^2 + x_t^2 \widetilde{\pi}_t^T \widetilde{C}_t \widetilde{\pi}_t \right]^{1/2} d\overline{W}_t, \quad x(t_0) = x_0. \end{aligned} \quad (2.18)$$

Note that in the drift of the SDE (2.18), the term

$$-r_t^e + x_t r_t^e = -r_t^e (x_t - 1), \quad (2.19)$$

appears because it models the effect of the decline of both liquid assets and nett cash outflows. Similarly the term $-(\sigma^e)^2 + x_t (\sigma^e)^2 = (\sigma^e)^2 (x_t - 1)$ appears.

3. Bank Liquidity in a Numerical Quantitative Framework

In this section, we discuss bank liquidity in a numerical quantitative framework. Recently the finance literature has devoted more attention to modeling and assessing liquidity risk in a numerical quantitative framework (see, e.g., [5, 8, 9]).

3.1. Bank Liquidity: Numerical Example 1

In this subsection, we use the data supplied in [6] (see, also, Appendices A.1 and A.2) to assess the liquidity of banks. The dataset originates from a supervisory liquidity report for Dutch banks. It covers a detailed breakdown of liquid assets and liabilities including cash in- and outflows of banks (see, also, [5]).

3.1.1. Data Description: Numerical Example 1

The aforementioned supervisory liquidity report includes on- and off-balance sheet items for about 85 Dutch banks (foreign bank subsidiaries included) with a breakdown per item (average granularity of about 7 items per bank). The report presents month end data available for the period October 2003 to March 2009. In this case, supervisory requirements dictate that

actual bank liquidity must exceed required liquidity, at both a one-week and a one-month horizon. Actual liquidity is defined as the stock of liquid assets (weighted for haircuts) and recognized cash inflows (weighted for their liquidity value) during the test period. Required liquidity is defined as the assumed calls on contingent liquidity lines, assumed withdrawals of deposits, drying up of wholesale funding, and liabilities due to derivatives. In this way, the liquidity report comprises a combined stock and cash flow approach, in which respect it is a forward looking concept. The weights, w^i , of the assumed haircuts on liquid assets and run-off rates of liabilities are presented in last two columns of Tables 1 and 2 below. In this regard, the pecking order hypothesis is tested empirically in [5] by classifying the assets and liabilities of the banks in our sample according to the month weights in the liquidity report (as presented in the last column of Tables 1 and 2). In the report, the w^i values are fixed (see, e.g., [6]) and reflect the bank-specific and market-wide situation. The w^i values are based on best practices of values of haircuts on liquid assets and run-off rates of liabilities of the banking industry and credit rating agencies.

The various balance sheet and cash flow items in the prudential report [6] are assumed to reflect the instruments which banks use in liquidity risk management by way of responding to shocks. The instruments are expressed in gross amounts. To enhance the economic interpretation we define coherent groups of instruments and the sum of item amounts per group. The first column of Tables 1 and 2 below provides the group classification. Here, the second columns in these tables describe the particular class of assets and liabilities. For the liquidity test for the full month, a distinction is made between non-scheduled items and scheduled items. By contrast to non-scheduled items, scheduled items are included on the basis of their possible or probable due dates. For the liquidity test for the first week, scheduled items are only included if they are explicitly taken into account in day-to-day liquidity management Treasury operations. In Tables 1 and 2 below, scheduled items are indicated by the letter S.

3.1.2. Data Presentation: Numerical Example 1

In this section, we firstly represent data related to assets and then data related to liabilities.

3.1.3. Data Analysis: Numerical Example 1

From Tables 1 and 2, we have seen that the behavior of banks can be described by rather simple indicators constructed from firm-specific balance sheet data. Although they are descriptive in nature, the measures identify trends in banks behavior that convey forward looking information on market-wide developments. A key insight from the analysis is that while banks usually follow a pecking order in their balance sheet adjustments (by making larger adjustments to the most liquid balance sheet items compared to less liquid items), during the crisis banks were more inclined to a static response. This suggests that they have less room to follow a pecking order in their liquidity risk management in stressed circumstances. It implies that banks responses in crises may have more material effects on the economy, since a static response rule means that banks are more likely to adjust their (less liquid) retail lending and deposits than under normal market conditions. A sufficient stock of liquid buffers could prevent that banks are forced to such detrimental static responses, which lends support to the initiatives of the Basel Committee to tighten liquidity regulation for banks (see, e.g., [1]).

Table 1: Assets for liquidity testing.

Group	Assets	S	Week	Month
	Cash in the form of Banknotes/Coins		100	100
	Receivables from CBs (including ECB)			
1	1		100	100
1	2	S	100	100
1	3	S	100	100
1	4	S	d^*	d^*
	Collection documents			
	1		100	100
	2	S	100	100
	Readily marketable debt instruments/CB eligible receivables			
	<i>Issued by public authorities</i>			
2	1		95**	95**
2	2		85**	85**
2	3		85	85
2	4		95	95
	Zone A			
2	5		70	70
	Zone B			
	<i>Issued by credit institutions</i>			
2	1		90**	90**
2	2		80**	80**
2	3		90	90
2	4		70	70
	<i>Issued by other institutions</i>			
2	1		90**	90**
2	2		80**	80**
2	3		90	90
2	4		70	70
	Amounts receivables			
	<i>Branches and banking subsidiaries not included in the report</i>			
3	1		50	100
3	2	S	100	100
3	3	S	100	90
	<i>Other credit institutions</i>			
3	1		50	100
3	2	S	100	100
3	3	S	100	90
	<i>Public authorities</i>			
3	1		50	100
3	2	S	100	100
3	3	S	100	90

Table 1: Continued.

Group	Assets	S	Week	Month
	<i>Other professional money market players</i>			
3	1		50	100
3	2	S	100	100
3	3	S	100	90
	<i>Other counterparties</i>			
	1		0	0
	2	S	100	90
4	3	S	50	40
	Receivables i.r.o REPO and reverse REPO transactions			
	<i>Reverse repo transactions (other than with CBs)</i>			
5	1	S	100	100
5	2	S	100	100
	<i>Repo Transactions (Other Than with CBs)</i>			
5	1	S	90/d*/**	90/d*/**
5	2	S	70	70
	<i>Securities lending/borrowing transactions</i>			
5	1		100	100
	Lending/borrowing transactions			
5	2		100	100
	Lending/borrowing transactions			
	Other securities and gold			
6	1		70	70
6	2		0	0
2	3	S	100	100
	4		90	90
	Official standby facilities			
14	1		100	100
14	Receivables i.r.o derivatives	S	***	***

*: Less applicable discount.

**: Either at stated percentage or at percentages applicable for ecb/escb collateral purposes.

***: Calculated amount for the period concerned.

90/d*/**: 90% OR less applicable discount (provided the method is consistently applied).

The measures for size and the number of extreme balance sheet adjustments gauge the time dimension of macroprudential risk, and indicators of the dependency and concentration of reactions capture the cross-sectoral dimension. The measures are robust to different specifications and distributions of the data. Applied to Dutch banks, the measures show that the number, size, and similarity of responses substantially changed during the crisis, in particular on certain market segments. They also indicate that the nature of banks behavior is asymmetric, being more intense in busts than in booms. Furthermore, during the crisis the deleveraging of large banks started earlier was more intense and more advanced than the deleveraging of smaller banks.

Given these findings, the indicators are useful for macroprudential analysis, for instance with regard to monitoring frameworks. Our analysis underscores the relevance

Table 2: Liabilities for liquidity testing.

Group		Liabilities	S	Week	Month
		Moneys borrowed from CBs			
7	1	Overdrafts payable within one week		100	100
7	2	Other amounts owed	S	100	100
		Debt instruments issued by the bank itself			
8	1	Issued debt securities	S	100	100
8	2	Subordinate liabilities	S	100	100
		Deposits and fixed term loans			
		<i>Branches and banking subsidiaries not included in the report</i>			
9	1	Amounts owed i.r.o securities transactions	S	100	100
9	2	Deposits and other funding—fixed maturity—plus interest payable	S	100	90
		<i>Other counterparties</i>			
	1	Amounts owed i.r.o securities transactions	S	100	100
10	2	Deposits and other funding—fixed maturity—plus interest payable	S	100	90
10		Fixed-term savings deposits	S	20	20
		Liabilities i.r.o repo and reverse repo transactions			
		<i>Repo transactions (other than with CBs)</i>			
11	1	Amounts owed i.r.o bonds	S	100	100
11	2	Amounts owed i.r.o shares	S	100	100
		<i>Securities lending/borrowing transactions</i>			
11	1	Negative securities stock on account of securities lending/borrowing transactions		100	100
11	2	Securities to be delivered on account of securities lending/borrowing transactions	S	100	100
		Credit balances and other moneys borrowed with an indefinite effective term			
		<i>Branches and banking subsidiaries not included in the report</i>			
12	1	Current account balances and other demand deposits		50	100
		<i>Other credit institutions</i>			
12	1	Balances on vostro accounts of banks		50	50
12	2	Other demand deposits		50	100
		<i>Other professional money market players</i>			
12	1	Demand deposits		50	100
		<i>Savings accounts</i>			
13	1	Savings accounts without a fixed term		2.5	10
		<i>Other</i>			
13	1	Demand deposits and other liabilities		5	20
13	2	Other amounts due and to be accounted for including the balance of forward transactions and amounts due i.r.o. social and provident funds		5	20

Table 2: Continued.

Group		Liabilities	S	Week	Month
		Official standby facilities			
14	1	Official standby facilities granted		100	100
		Liabilities i.r.o. derivatives			
14	1	Known liabilities i.r.o derivatives	S	***	***
14	1	Unknown liabilities i.r.o derivatives		***	***
		Other contingent liabilities and irrevocable credit facilities			
14	1	Unused irrevocable credit facilities, including underwriting of issues		2.5	10
14	2	Bills accepted	S	100	100
14	3	Credit-substitute guarantees		2.5	10
14	4	Non-credit-substitute guarantees		2.5	10
14	5	Other off-balance sheet liabilities		1.25	5

of using several indicators of liquidity risk at the same time, given the different leads and lags of the measures with systemic risk. The empirical results also provide useful information for financial stability models. A better understanding of banks behavior helps to improve the microfoundations of such models, especially with regard to the behavioral assumptions of heterogeneous institutions. Finally, the empirical findings in our study are relevant to understand the role of banks in monetary transmission and to assess the potential demand for CB finance in stress situations. The measures explain developments of financial intermediation channels (wholesale and retail, unsecured, secured, etc.) along the cross-sectional and time dimensions. They also shed more light on the size and number of banks that rely on CB financing.

3.2. Bank Liquidity: Numerical Example 2

In this section, we provide a simulation of the LCR dynamics given in (2.18).

3.2.1. Simulation: Numerical Example 2

In this subsection, we provide parameters and values for a numerical simulation. The parameters and their corresponding values for the simulation are shown in Table 3.

3.2.2. LCR Dynamics: Numerical Example 2

In Figure 1, we provide the LCR dynamics in the form of a trajectory derived from (2.18).

3.2.3. Properties of the LCR Trajectory: Numerical Example 2

Figure 1 shows the simulated trajectory for the LCR of low liquidity assets. Here different values of banking parameters are collected in Table 3. The number of jumps of the trajectory was limited to 1001, with the initial values for l fixed at 20.

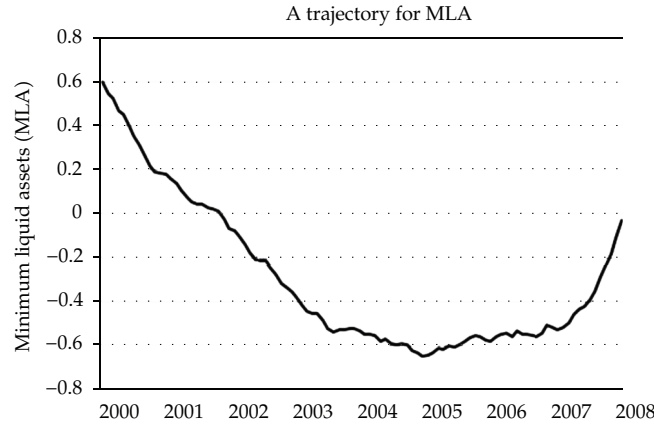


Figure 1: Trajectory of the LCR for low liquidity assets.

Table 3: Choices of liquidity coverage ratio parameters.

Parameter	Value	Parameter	Value	Parameter	Value
C	1 000	r^C	0.06	r^e	0.07
r^i	0.02	σ^e	1.7	σ^i	1.9
\tilde{r}^y	0.05	$\tilde{\pi}$	0.4	u^1	0.03
u^2	0.01	\tilde{C}	750	\bar{W}	0.01

As we know, banks manage their liquidity by offsetting liabilities via assets. It is actually the diversification of the bank's assets and liabilities that expose them to liquidity shocks. Here, we use ratio analysis (in the form of the LCR) to manage liquidity risk relating various components in the bank's balance sheets. In Figure 1, we observe that between $t = 2000$ and $t = 2005$, there was a significant decrease in the trajectory which shows that either liquid assets declined or nett cash outflows increased.

There was also an increase between $t = 2005$ and $t = 2007$ which suggests that either liquid assets increased or nett cash outflows decreased. There was an even sharper increase subsequent to $t = 2007$ which comes as somewhat of a surprise. In order to mitigate the aforementioned increase in liquidity risk, banks can use several facilities such as repurchase agreements to secure more funding. However, a significant increase was recorded between $t = 2005$ and $t = 2008$, with the trend showing that banks have more liquid assets on their books. If $l > 0$, it demonstrated that the banks has kept a high volume of liquid assets which might be stemming from quality liquidity risk management. In order for banks to improve liquidity they may use debt securities that allow savings from nonfinancial private sectors, a good network of branches and other competitive strategies.

The LCR has some limitations regarding the characterization of the banks liquidity position. Therefore, other ratios could be used for a more complete analysis taking into account the structure of the short-term assets and liabilities of residual maturities.

4. Bank Liquidity in a Theoretical Quantitative Framework

In this section, we investigate bank liquidity in a theoretical quantitative framework. In particular, we characterize a liquidity provisioning strategy and discuss residual aggregate risk in order to eventually determine the appropriate value of the price process. In order to model uncertainty, in the sequel, we consider the filter probability space $(\Omega, \mathbb{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$, $T \in \mathbb{R}$ described in Assumption 2.1.

4.1. Preliminaries about the Liquidity Provisioning Strategy

Firstly, we consider a dynamic liquidity provisioning strategy for a risky underlying illiquid asset process, $(\Lambda_t)_{0 \leq t \leq T}$. For purposes of relating the discussion below to the GFC, we choose Λ to be residential mortgage loans hereafter known simply as *mortgages*. Mortgages were very illiquid (nonmarketable) before and during the GFC. In this case, for liquidity provisioning purposes, the more liquid marketable securities, S —judging by their credit rating before and during the GFC—are used as a substitute for mortgages. This was true during the period before and during the GFC, with mortgage-backed securities being traded more easily than the underlying mortgages. Furthermore, we assume that the bank mainly holds illiquid mortgages and marketable securities (compare with the assets presented in Tables 1 and 2) with cash for investment being injected by outside investors. The liquid marketable securities, S , are not completely correlated with the illiquid mortgages, Λ , creating market incompleteness. Under the probability measure, \mathbb{P} , the price of the traded substitute securities and the illiquid underlying mortgages are given by

$$dS_t = S_t [\mu^s dt + \sigma^s dW_t^S], \quad d\Lambda_t = \Lambda_t [\mu^\Lambda dt + \sigma^\Lambda dW_t^\Lambda], \quad (4.1)$$

respectively, where μ and σ are constants. We define the constant market price of risk for securities as

$$\lambda_s = \frac{\mu_s - r}{\sigma}. \quad (4.2)$$

We note that if the market correlation $|\rho|$ between W^S and W^Λ is equal to one, then the securities and mortgages are completely correlated and the market is complete.

Let Θ be a liquidity provisioning strategy for the bank's asset portfolio. The dynamics of its wealth process is given by

$$d\Pi_t = n_t^S dS_t + (\Pi_t - n_t^S S_t) r dt + dC_t, \quad (4.3)$$

where dC_t is an amount of cash infused into the portfolio, n_t^S is the number of shares of securities held in the portfolio at time t , Π_t is the value of the wealth process, and r is the riskless interest rate. The cumulative cost process $C(\Theta)$ associated with the strategy, Θ , is

$$C_t(\Theta) = \tilde{\Pi}_t(\Theta) - \int_0^t n_u^S dS_u, \quad 0 \leq t \leq T. \quad (4.4)$$

The cost process is the total amount of cash that has been injected from date 0 to date t . We determine a provisioning strategy that generate a payoff $(\Lambda_T - K)^+$ at the maturity T . The quantity $\int_t^T \exp\{-r(s-t)\} dC_s$ is the discounted cash amount that needs to be injected into the portfolio between dates t and T . Since $\int_t^T \exp\{-r(s-t)\} dC_s$ is uncertain, the risk-averse agent will focus on minimizing the associated ex-ante aggregate liquidity risk

$$R_t(\Theta) = \mathbb{E}^{\mathbb{P}} \left[\left(\int_t^T \exp\{-r(s-t)\} dC_u \right)^2 \right], \quad 0 \leq t \leq T. \quad (4.5)$$

It is clear that this concept is related to the conditioned expected square value of future costs. The strategy $\Theta, 0 \leq t \leq T$ is mean self-financing if its corresponding cost process $C = (C_t)_{0 \leq t \leq T}$ is a martingale. Furthermore, the strategy Θ is self-financing if and only if

$$\tilde{\Pi}_t(\Theta) = \tilde{\Pi}_0(\Theta) + \int_0^t n_u^S d\tilde{S}_u, \quad 0 \leq t \leq T. \quad (4.6)$$

A strategy $\hat{\Theta}$ is called an admissible continuation of Θ if $\hat{\Theta}$ coincides with Θ at all times before t and $\Pi_t(\Theta) = L, \mathbb{P}$ a.s. Moreover, a provisioning strategy is called *liquidity risk minimizing* if for any $t \in [0, T], \Theta$ minimizes the remaining liquidity risk. In other words, for any admissible continuous $\hat{\Theta}$ of Θ at t we have

$$R_t(\Theta) \leq R_t(\hat{\Theta}). \quad (4.7)$$

Criterion given in (4.5) can be formally rewritten as

$$\forall t \min_{(n^S, \Pi)} R_t, \quad \text{subject to } \Pi_t = (\Lambda_T - K)^+. \quad (4.8)$$

We define the *expected squared error of the cost over the next period* as

$$\mathbb{E}^{\mathbb{P}}[(\Delta C_t)^2] = \mathbb{E}_t \left[\left(\Pi_{t+\Delta t} - \Pi_t - n_t^S (S_{t+\Delta t} - S_t) - \left(\Pi_t - n_t^S S_t \right) (\exp\{r(t+\Delta t)\} - \exp\{rt\}) \right)^2 \right]. \quad (4.9)$$

In the next section, we minimize the above quantity at each date, with respect to $(n_0^S, n_{\Delta}^S, \dots, n_{t+\Delta t}^S)$ and also discuss the notion of a liquidity provisioning strategy.

4.2. Characterization of the Liquidity Provisioning Strategy

During the GFC, liquidity provisioning strategies involved several interesting elements. Firstly, private provisioning of liquidity was provided via the financial system. Secondly, there was a strong connection between financial fragility and cash-in-the-market pricing. Also, contagion and asymmetric information played a major role in the GFC. Finally, much of the debate on liquidity provisioning has been concerned with the provisioning of

liquidity to financial institutions and resulting spillovers to the real economy. The next result characterizes the liquidity provisioning strategy that we study.

Theorem 4.1 (characterization of the provisioning strategy). *The locally liquidity risk minimizing strategy is described by the following.*

(1) *The investment in mortgages is*

$$\hat{n}_t^S = \frac{\sigma^\Lambda \Lambda_t}{\sigma^S S_t} \rho C^\Lambda(t, \Lambda_t) = \frac{\sigma^\Lambda \Lambda_t}{\sigma^S S_t} \rho \exp\left\{\left(\mu^\Lambda - r - \rho \sigma^\Lambda \lambda^S\right)(T-t)\right\} \mathbb{N}(d_{1,t}), \quad (4.10)$$

where λ_s is the Sharpe ratio and $C(t, \Lambda_t)$ is the minimal entropy price

$$\begin{aligned} C(t, \Lambda_t) &= \exp\{-r(T-t)\} \mathbb{E}^\mathbb{Q}[(\Lambda_T - K)^+] \\ &= \exp\left\{\left(\mu^\Lambda - r - \rho \sigma^\Lambda \lambda^S\right)(T-t)\right\} \Lambda_t \mathbb{N}(d_{1,t}) - K \exp\{-r(T-t)\} \mathbb{N}(d_{2,t}), \end{aligned} \quad (4.11)$$

where \mathbb{Q} is the minimal martingale measure defined as

$$\begin{aligned} \frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_t &= \exp\left\{-\frac{1}{2} \Lambda^{S^2}(T-t) - \lambda^S (W_T^S - W_t^S)\right\}, \\ d_{1,t} &= \frac{1}{\sigma^\Lambda \sqrt{T-t}} \left[\ln \frac{\Lambda_t}{K} + \left(\mu^\Lambda - \rho \sigma^\Lambda \lambda^S + \frac{\sigma^{\Lambda^2}}{2} \right) (T-t) \right], \quad d_{2,t} = d_{1,t} - \sigma^\Lambda \sqrt{T-t}. \end{aligned} \quad (4.12)$$

(2) *The cash investment is*

$$C(t, \Lambda_t) - \hat{n}_t^S S_t. \quad (4.13)$$

If the Sharpe ratio, λ_s , of the traded substitute securities is equal to zero, the minimal martingale measure coincides with the original measure \mathbb{P} , and the above strategy is globally liquidity risk minimizing.

Proof. Let $\hat{S}_t \equiv \exp\{-rt\} S_t$ be the discounted value of the traded securities at time t . This process follows a martingale under the martingale measure, \mathbb{Q} , since we have

$$d\hat{S}_t = \hat{S}_t \sigma^S dW_t^{S,\mathbb{Q}}, \quad (4.14)$$

where $dW_t^{S,\mathbb{Q}} \equiv dW_t^S + \lambda^S dt$ is the increment to a \mathbb{Q} -Brownian motion. Hence, we can write the Kunita-Watanabe decomposition of the discounted option payoff under \mathbb{Q} :

$$\exp\{-rt\} (\Lambda_T - K)^+ = H_0 + \int_0^T \zeta_t d\hat{S}_t + V_T^H, \quad (4.15)$$

where V^H is a \mathbb{P} -martingale orthogonal to \hat{S} under \mathbb{Q} . Lévy's Theorem shows that the process A defined by

$$dA_t = \frac{dW_t^\Lambda - \rho dW_t^S}{\sqrt{1 - \rho^2}} \quad (4.16)$$

is a \mathbb{P} -Brownian motion and that it is independent of W^S . Then, by Girsanov's theorem, $(W^{S,\mathbb{Q}}, A)$ also follows two-dimensional \mathbb{Q} -Brownian motion. Since V^H is a martingale under \mathbb{Q} and is orthogonal to F , the martingale representation theorem shows that we have $V_t^H = \psi dA_t$ for some process ψ . In particular, V^H is orthogonal under \mathbb{P} to the martingale part of \hat{S} , where the martingale part of \hat{S} under \mathbb{P} is defined as

$$G_t = \int_0^t \sigma^S S_s dW_s^S. \quad (4.17)$$

Next, we suppose that

$$P_t = \exp\{-r(T-t)\} \mathbb{E}^\mathbb{Q}[(\Lambda_T - K)^+]. \quad (4.18)$$

Using (4.15) we obtain

$$P_t = \exp\{rt\} \left[H_0 + \int_0^t \zeta_s d\hat{S}_s + V_t^H \right]. \quad (4.19)$$

Consider now the non-self-financing strategy with value $\tilde{\Pi}_t = P_t$ and the number of securities given by $\tilde{n}_t^S = \zeta_t$. Given (4.1) and (4.19), we obtain that $d\tilde{C}_t = \exp\{rt\} dV_t^H$. This shows that, V^H, \tilde{C} is a \mathbb{P} -martingale orthogonal to G . We recall that a strategy (Π, n^S) is *locally risk minimizing* if and only if the associated cost process follows a \mathbb{P} -martingale orthogonal to G . Hence the strategy $(\tilde{\Pi}, \tilde{n}^S)$ is locally risk minimizing.

We now prove an explicitly expression for the random variable P_t , which is called the *minimum entropy price*. The Black-Scholes formula implies that

$$\begin{aligned} P_t &= C(t, \Lambda_t) \\ &= \exp\left\{(\mu^\Lambda - r - \sigma^\Lambda \rho \lambda^S)(T-t)\right\} \left[\Lambda_t N(d_1, t) - K \exp\left\{(\mu^\Lambda - \rho \sigma^\Lambda \lambda^S)(T-t)\right\} N(d_2, t) \right], \end{aligned} \quad (4.20)$$

which can be written as a function $C(t, \Lambda_t)$ of t and Λ_t . Using (4.19), we obtain that

$$\zeta_t = \frac{\sigma^\Lambda \Lambda_t}{\sigma^S S_t} \rho C^\Lambda(t, \Lambda_t). \quad (4.21)$$

The required expression for \tilde{n}_t^F follows immediately. □

Our paper addresses the problem of dynamic bank provisioning for (illiquid) nonmarketable mortgages, Λ , for which substitute (liquid) marketable securities, S , is part of the liquidity provisioning strategy. Due to the presence of cross-hedge liquidity risk we operate in an incomplete market setting. In this regard, we employ a non-self-financing strategy to ensure that uncertainty is reduced and trading is conducted in the Treasuries market. Moreover, the strategy is designed to influence a perfect replication at the cost of continuous cash infusion into the replicating bank portfolio. Since the cash infusion is random, the risk-averse agent would require that the total uncertainty involved over the remaining life of the mortgage be minimized. As a consequence, for $0 \leq t \leq T$, the associated *ex-ante aggregated liquidity risk* is given by

$$R_t(\Theta) = \mathbb{E}^{\mathbb{P}} \left[\left(\int_t^T \exp\{-r(s-t)\} dC_u \right)^2 \right], \quad 0 \leq t \leq T. \quad (4.22)$$

Now $\int_t^T \exp\{-r(s-t)\} dC_u$ is stochastic, so we will focus on minimizing the risk in (4.5). We apply a local risk minimization criterion which entails that instead of minimizing the uncertainty with respect to the cash infusion, C_t , over the process, the strategy attempts to minimize, at each date, the uncertainty over the next infinitesimal period. Also, the incompleteness entails the existence of infinitely many equivalent martingale measures. In order to determine the appropriate price of the asset value one should choose an appropriate equivalent martingale measure. In this case, the process is \mathbb{Q} -Brownian motion so that the discounted price process $\exp\{-rt\}S_t$ follows martingale pricing. The equivalent martingale measure will be determined according to the risk minimization criterion in Theorem 4.1. Let us consider the discounted price to be

$$\exp\{-rt\}(\Lambda_T - K)^+. \quad (4.23)$$

Applying the Kunita-Watanabe decomposition for the discounted price under a measure \mathbb{Q} , we get

$$K^\Pi = H_0 + \int_0^T \zeta_t d\hat{S}_u + V_T^H, \quad (4.24)$$

where V^H is a \mathbb{P} -martingale orthogonal to \hat{S} under the measure \mathbb{Q} . Let

$$P_t = \exp\{-r(T-t)\} \mathbb{E}_t^{\mathbb{Q}}[(\Lambda_T - K)^+] \quad (4.25)$$

which can be rewritten as

$$P_t = \exp\{rt\} \left[H_0 + \int_0^t \zeta_u d\hat{S}_u + L_t^H \right]. \quad (4.26)$$

4.3. Residual Aggregate Liquidity Risk

During the GFC, two types of uncertainty concerning liquidity requirements arose. Firstly, each individual bank was faced with idiosyncratic liquidity risk. At any given time its depositors may have more or less liquidity needs. Uncertainty also arose from the fact that banks face aggregate liquidity risk. In some periods aggregate liquidity demand is high while in others it is low. Thus, aggregate risk exposes all banks to the same shock, by increasing or decreasing the demand for liquidity that they face simultaneously. The ability of banks to hedge themselves against these liquidity risks crucially depend on the functioning, or, more precisely, the completeness of financial markets. The next theorem provides an explicit expression for the aggregate liquidity risk when a locally risk minimizing strategy is utilized in an incomplete market.

Theorem 4.2 (residual aggregate liquidity risk). *The aggregate liquidity risk when a locally risk minimizing strategy at time t is implemented is equal to*

$$R_t^{rm} = \sigma^{\Lambda^2} (1 - \rho^2) \int_t^T \exp\{-2r(s-t)\} \mathbb{E}^{\mathbb{P}} \left[\Lambda^{S^2} C^{\Lambda}(s, \Lambda_s)^2 \right] ds. \quad (4.27)$$

This can be approximated by

$$R_t^{rm} \approx \sigma^{\Lambda^2} (1 - \rho)^2 C^{\Lambda}(0, \Lambda_0)^2 \Lambda_0^2 \frac{1 - \exp\{-2r(T-t)\}}{2r}. \quad (4.28)$$

Proof. Let us now assume $\Phi = \Phi^*$ and $\Pi_t = C(t, \Lambda_t)$. Under \mathbb{Q} , the wealth process, Π , evolves as

$$d\Pi_t = r\Pi_t dt + \rho\sigma^{\Lambda}\Lambda_t C^{\Lambda}(t, \Lambda_t) dW_t^{S, \mathbb{Q}} + d\tilde{C}_t. \quad (4.29)$$

In addition, $(\exp\{-rt\}C(t, \Lambda_t))_t$ follows a \mathbb{Q} -martingale, where

$$dC(t, \Lambda_t) = rC(t, \Lambda_t)dt + C^{\Lambda}(t, \Lambda_t)\sigma^{\Lambda}\Lambda_t dW_t^{\Lambda, \mathbb{Q}}, \quad (4.30)$$

$$dW_t^{\Lambda, \mathbb{Q}} = dW_t^{\Lambda} + \rho\lambda^S dt \quad (4.31)$$

defines a \mathbb{Q} -Brownian motion. One can write it as

$$dW_t^{\Lambda, \mathbb{Q}} = dW_t^{\Lambda} - \rho dW_t^S + \rho dW_t^{S, \mathbb{Q}} = \sqrt{1 - \rho^2} dW_t^2 + \rho dW_t^{\Lambda, \mathbb{Q}}. \quad (4.32)$$

Comparing (4.29) and (4.30) we obtain that

$$\exp\{-rt\}dC_t = \exp\{-rt\}C^{\Lambda}(t, \Lambda_t)\sigma^{\Lambda}\Lambda_t\sqrt{1 - \rho^2}dW_t^2, \quad (4.33)$$

hence (4.27).

In what follows, we let δ_t be the delta of the mortgage process at time t that is computed from the minimal entropy price so that $\delta_t = C^\Lambda(t, \Lambda_t)$. We must now compute $\mathbb{E}^\mathbb{P}[\delta_t^2 \Lambda_t^2]$ for all t in $[0, T]$. If $(\delta_t^2, \Lambda_t^2)_{t \geq 0}$ were a martingale, the task would be easy since we would have $\mathbb{E}^\mathbb{P}[\delta_t^2 \Lambda_t^2] = \delta_0^2 \Lambda_0^2$. But $(\delta_t^2 \Lambda_t^2)_{t \geq 0}$ is not a martingale. However, it can be shown that for small $\sigma^{\Lambda^2} T$, the expectation $\mathbb{E}^\mathbb{P}[\delta_t^2 \Lambda_t^2]$ is approximated by the constant $\delta_0^2 \Lambda_0^2$. The formal proof follows from the fact that $\mathbb{E}^\mathbb{P}[\gamma_t \Lambda_t^2] \approx \gamma_0 \Lambda_0^2$, $\gamma_t = C^{\Lambda\Lambda}(t, \Lambda_t)$, denoting the gamma of the value of the asset. Therefore, we finally have that

$$\sigma^{\Lambda^2} (1 - \rho^2) \int_t^T \exp\{-2rs\} \mathbb{E}^\mathbb{P}[\delta_s^2 \Lambda_s^2] ds \approx \sigma^{\Lambda^2} (1 - \rho^2) \delta_0^2 \Lambda_0^2 \frac{1 - \exp\{-2r(T-t)\}}{2r}. \quad (4.34)$$

□

Applying a non-self-financing strategy and considering $\tilde{\Pi}_t = P_t$ and $\dot{n}_t^S = \zeta_t$, we obtain that

$$d\tilde{C}_t = \exp\{rt\} dV_t^H. \quad (4.35)$$

This implies that \tilde{C} is \mathbb{P} -martingale orthogonal to G . In this regard, the strategy (Π, n^S) is locally risk minimizing if and only if the associated cost process $C(\Theta)$ follows a \mathbb{P} -martingale orthogonal to G . This means the strategy minimizes at each date the uncertainty over the next infinitesimal period. In applying the risk-minimization strategy there remains some “residual” aggregate liquidity risk stemming from the imperfection of the Brownian motion processes W^S and W^Λ . After the bank has implemented the locally risk minimizing strategy at time t , the aggregate liquidity risk is

$$R_t^{rm} = \sigma^{\Lambda^2} (1 - \rho^2) \int_t^T \exp\{-2r(s-t)\} \mathbb{E}^\mathbb{P}[\Lambda^2 C^\Lambda(S, \Lambda_u)^2] du. \quad (4.36)$$

For δ_t associated with the value process at time t computed via the minimized entropy price, we now need to compute $\mathbb{E}^\mathbb{P}[\delta_t^2, \Lambda_t^2]$ for all t in $[0, T]$. Since $\mathbb{E}^\mathbb{P}[\delta_t^2, \Lambda_t^2]$ is approximated by the constant $\delta_0^2 \Lambda_0^2$, then $\mathbb{E}^\mathbb{P}[\gamma_t \Lambda_t^2] \approx \gamma_0 \Lambda_0^2$, $\gamma_t = C^\Lambda(t, \Lambda_t)$ which is the gamma of the mortgage value. Therefore, the residual liquidity risk at time t is

$$\sigma^{\Lambda^2} (1 - \rho)^2 \int_t^T \exp\{-2rs\} \mathbb{E}^\mathbb{P}[\delta_u^2 \Lambda_u^2] du \approx \sigma^{\Lambda^2} (1 - \rho)^2 \delta_0^2 \Lambda_0^2 \frac{1 - \exp\{-2r(T-t)\}}{2r}. \quad (4.37)$$

5. Conclusions and Future Directions

In this paper, we discuss liquidity risk management for banks. We investigate the stochastic dynamics of bank items such as loans, reserves, securities, deposits, borrowing and bank capital (compare with Question 1). In accordance with Basel III, our paper proposes that overall liquidity risk is best analyzed using ratio analysis approaches. Here, liquidity risk is measured via the LCR. In this case, we provide numerical results to highlight some important issues. Our numerical quantitative model shows that a low LCR stems from a low level of liquid assets or high nett cash outflows (compare with Question 2). Moreover, we provide a characterization of liquidity risk provisioning by considering an (illiquid) nonmarketable

mortgage as an underlying asset and using (liquid) marketable securities for provisioning. In this case, we use non-self-financing strategy that considers market incompleteness to provision for liquidity risk. Then, we provide a quantitative framework for assessing residual risk stemming from the above strategy (compare with Question 3).

Future research should focus on other features of the GFC that are related to liquidity provisioning. The first involves the decrease in prices of AAA-rated tranches of structured financial products below fundamental values. The second is the effect of the GFC on interbank markets for term funding and collateralized money markets. Thirdly, further investigations should address the fear of contagion should a major institution fail. Finally, the effects on the real economy should be considered. In addition, the stochastic dynamic model we have considered in this paper does not take assets and liabilities with residual maturities into account. Such a model should be developed.

Appendices

A. More about Liquidity Risk

In this section, we provide more information about measures by cash flow, liquidity monitoring approaches, liquidity risk ratings and national approaches to liquidity risk.

A.1. Measures by Cash Flow

Banks use the intensity of the cash flow to predict the level of stress events. In this case, we determine the level of both cash in flows and cash out flows depending on both supply and demand for liquidity in the normal market performance. In this regard, the bank cash flow predicts the level of stress event s . Moreover, the use of proforma is an acceptable standard which determines the uses and sources of funds in the bank. It identifies where the bank funding short fall and liquidity gap lies.

A.2. Liquidity Monitoring Approaches

The BCBS has set international standards for sound management of liquidity risk. In this regard, the monitoring and evaluation of the banks operational activities is an internal control measure. However, the monitoring approach is divided into three levels, that is, the liquid assets approach, the cash flow approach, and a mixture of both. *Liquid asset approach* is mostly appropriate used in the Treasury bond market. In this regard, banks are required to maintain some liquid asset in their balance sheet that could be used during the hard period. Assets such as government securities are appropriate to maintain in the balance sheet because they can easily enable the bank to secure funding through securitization. While *Cash flow matching approach* enable banks to match the cash in flows with the cash out flows from the balance sheet activities.

The monitoring approaches for assessing liquidity risk is divided into three classes, that is, liquid asset approach, the cash flow approach and the combination of both. In the liquid asset approach a bank prescribed to a minimum level of cash or high-quality liquid or marketable assets in relation to the deposits and other sources of funds. While maturity

Table 4

Quantitative indicators
Availability of funds
Diversification of funding sources
Alternative funding sources
Capacity to augment liquidity through asset sales and/or securitization
Volume of wholesale liabilities with embedded options
Vulnerability of a bank to funding difficulties
Support provided by parent company
Qualitative indicators
Effectiveness of a board's policy in response to liquidity risk
Effectiveness process in identifying, measuring, monitoring, and controlling
Interacting of management to changing market conditions
Development of contingency funding plans
Information system management
Comprehensive and effective internal audit coverage

mismatch classify the expected inflows and outflows of funds into time bands of their residual maturity.

A.3. Liquidity Risk Rating

The rating of liquidity risk is categorized into two sets of indicators, that is, the quantitative and qualitative liquidity risk indicators.

Table 4 shows the quantitative and qualitative liquidity risk indicators. In light of the above, the rating for quantitative liquidity risk management is classified into three levels, that is, low, moderate level, and high level of liquidity risk. Therefore, a bank with a full set of all the indicated quantitative indicators has a low level of liquidity risk. Moreover, the rating for qualitative liquidity risk is divided into three levels, that is, strong, satisfactory, and weak quality of management of liquidity risk. In the above, we indicated that rating of liquidity risk is divided into two sets of indicators, namely, the quantitative liquidity risk indicators and qualitative liquidity risk indicators. According to Table 4, a bank with a full set of all the indicated qualitative indicators has a low level of liquidity risk, while a bank with a full set of all indicated qualitative indicators has a higher level of liquidity risk management.

A.4. National Approaches to Liquidity Risk

In this section, we discuss a useful principle which needs to be developed by individual countries to ensure sound management of liquidity risk and appropriate level of liquidity insurance by banks. This principle could be enforced via policies that assess liquidity as an internal measure; stress testing and other scenario analysis which determine the probability of a bank culminating into liquidity crisis; contingency funding to provide reliable sources of

funds to cover the short fall; setting limitations such as holding of liquid assets, minimum liquid assets, limits on maturity mismatches, and limits on a particular funding sources; reporting about liquidity risks and sources of liquidity as well as through public disclosure to enable investors to access bank information.

References

- [1] Basel Committee on Banking Supervision, "Basel III: international framework for liquidity risk measurement, standards and monitoring," Tech. Rep., Bank for International Settlements (BIS), Basel, Switzerland, 2010.
- [2] Basel Committee on Banking Supervision, "Progress report on Basel III implementation," Tech. Rep., Bank for International Settlements (BIS), Basel, Switzerland, 2011.
- [3] Basel Committee on Banking Supervision, "Basel III: a global regulatory framework for more resilient banks and banking systems-revised version," Tech. Rep., Bank for International Settlements (BIS), Basel, Switzerland, 2011.
- [4] Basel Committee on Banking Supervision, "Principles for sound liquidity risk management and supervision," Tech. Rep., Bank for International Settlements (BIS), Basel, Switzerland, 2008.
- [5] J. W. van den End and M. Tabbæ, "When liquidity risk becomes a systemic issue: empirical evidence of bank behaviour," *Journal of Financial Stability*. In press.
- [6] De Nederlandsche Bank, "Credit system supervision manual," 2.3 Liquidity Risk 44–47, 2003.
- [7] F. Gideon, J. Mukuddem-Petersen, M. P. Mulaudzi, and M. A. Petersen, "Optimal provisioning for bank loan losses in a robust control framework," *Optimal Control Applications & Methods*, vol. 30, no. 3, pp. 309–335, 2009.
- [8] E. Loutskina, "The role of securitization in bank liquidity and funding management," *Journal of Financial Economics*, vol. 100, no. 3, pp. 663–684, 2011.
- [9] M. A. Petersen, B. de Waal, J. Mukuddem-Petersen, and M. P. Mulaudzi, "Subprime mortgage funding and liquidity risk," *Quantitative Finance*. In press.
- [10] F. Gideon, J. Mukuddem-Petersen, and M. A. Petersen, "Minimizing banking risk in a Lévy process setting," *Journal of Applied Mathematics*, vol. 2007, Article ID 32824, 25 pages, 2007.

Research Article

Optimal Control for a Class of Chaotic Systems

Jianxiong Zhang and Wansheng Tang

Institute of Systems Engineering, Tianjin University, Tianjin 300072, China

Correspondence should be addressed to Jianxiong Zhang, jxzhang@tju.edu.cn

Received 19 October 2011; Revised 17 January 2012; Accepted 10 February 2012

Academic Editor: Chuanhou Gao

Copyright © 2012 J. Zhang and W. Tang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes the optimal control methods for a class of chaotic systems via state feedback. By converting the chaotic systems to the form of uncertain piecewise linear systems, we can obtain the optimal controller minimizing the upper bound on cost function by virtue of the robust optimal control method of piecewise linear systems, which is cast as an optimization problem under constraints of bilinear matrix inequalities (BMIs). In addition, the lower bound on cost function can be achieved by solving a semidefinite programming (SDP). Finally, numerical examples are given to illustrate the results.

1. Introduction

As a very interesting nonlinear phenomenon, chaos has been widely applied in many areas, such as secure communication, signal generator design, biology, economics, and many other engineering systems, which has been researched thoroughly over the past two decades [1]. Recently, chaos control of chaotic systems has become an active research topic [2]. In general, there are several schemes to achieve the control of continuous time chaotic systems, such as OGY method [3], parametric resonance method [4], adaptive feedback method [5, 6], delay feedback method [7], backstepping design method [8], fractional controller design method [9], sliding mode control method [10, 11], internal model approach [12], impulsive control approach [13], as well as linear and nonlinear feedback control methods [14–17]. However, most of the existing methods were used to achieve chaos control either by employing the linearization scheme in the neighborhood of the objective point which is difficult to accomplish the global analysis, or by applying the nonlinear feedback controller which often limits practical applications. Based on the fuzzy control theory, Tanaka et al. [18] studied the feedback control of chaotic systems. The result formulated in terms of linear matrix inequalities (LMIs, [19]) was convenient to solve, but the controller design for the associated fuzzy systems was fulfilled by virtue of global quadratic Lyapunov function which is conservative in the control synthesis.

As pointed out in [20], piecewise linear systems, which can approximate general nonlinear systems to any degree of accuracy, can be analyzed based on piecewise quadratic Lyapunov function technique that introduces more flexibility than the classical global quadratic Lyapunov function technique. Thus, the piecewise linear systems provide a powerful way of analysis and synthesis for nonlinear systems. Chaotic systems belong to complex nonlinear systems. In fact, it is significant to design a practicable piecewise linear feedback controller to stabilize globally a chaotic system with a performance measure for the control synthesis. We recently [21] proposed a new chaotic system and designed a piecewise linear feedback controller to stabilize globally the new system based on piecewise linear systems method. So far, there have been very few results dealing with the optimal control for chaotic systems. In this paper, we investigate the problem of designing piecewise linear feedback controller to stabilize a class of chaotic systems, and meanwhile minimize a quadratic cost function for the closed-loop systems. Particularly, in this paper, a class of chaotic systems are converted to uncertain piecewise linear systems. Then, based on piecewise quadratic Lyapunov function technique and Hamilton-Jacobi-Bellman (HJB) inequality method, the optimal chaos control via piecewise linear state feedback controller is studied. It is shown that the optimal controller minimizing the upper bound on cost function can be obtained by solving an optimization problem under constraints of bilinear matrix inequalities (BMIs). The lower bound on cost function can be attained by solving a semidefinite programming (SDP). If the upper and lower bounds obtained are sufficiently tight, it is concluded that the associated solutions achieve or get close to optimality.

This paper is organized as follows. In Section 2, the optimal control problem of chaotic systems is introduced. In Section 3, the optimal control for a class of chaotic systems via piecewise linear state feedback controller is proposed. The upper bound and lower bound on cost function are designed. Illustrative examples are given in Section 4, and the conclusion is drawn in Section 5.

Throughout this paper, a real symmetric matrix $P > 0$ (≥ 0 , ≤ 0) denotes P being a positive definite (positive semidefinite, or negative semidefinite) matrix, and $A > B$ means $A - B > 0$. I denotes an identity matrix of appropriate dimension. The superscript “ T ” represents the transpose of a matrix. Matrices, if their dimensions are not explicitly stated, are assumed to have compatible dimensions for algebraic operations.

2. Problem Formulation

Consider the chaotic system of the form:

$$\dot{\mathbf{x}} = A\mathbf{x} + F(\mathbf{x}) + B\mathbf{u}, \quad (2.1)$$

where A and B are constant matrices, $\mathbf{x} \in \mathbb{R}^n$ is the state vector, $\mathbf{u} \in \mathbb{R}^m$ ($m \leq n$) is the control input variable, and the nonlinear term $F(\mathbf{x}) \in \mathbb{R}^n$ is assumed to satisfy Lipschitz continuity condition, uniform or local, and $F(0) = 0$.

Associated with this system is the cost function:

$$J = \int_0^\infty (\mathbf{x}^T(t)Q\mathbf{x}(t) + \mathbf{u}^T(t)R\mathbf{u}(t))dt, \quad (2.2)$$

where $Q > 0$, $R > 0$ are given weighting matrices.

The goal of this paper is to design a state feedback law $\mathbf{u}(t)$ stabilizing the chaotic system (2.1) and meanwhile minimizing the cost function (2.2).

It is known that the control law $\mathbf{u}(t)$ can be derived from the solution to the associated HJB equation. However, generally speaking, the HJB equation corresponding to a general nonlinear system is notoriously hard to solve. Many numerical methods have been devised for the solution of optimal control problems but tended to suffer from combinatorial explosion. Piecewise linear systems, which can approximate nonlinear systems to any degree of accuracy, provide a powerful means of analysis for nonlinear systems. By virtue of HJB inequalities rather than equations, the authors in [20, 22] have investigated the state feedback optimal control of piecewise linear systems. It was shown that the upper bound on piecewise quadratic cost function can be obtained by solving a nonconvex BMIs problem, and the lower bound on cost function can be obtained by solving an SDP. Motivated by this, we first convert the chaotic system (2.1) to the form of uncertain piecewise linear systems and then extend the corresponding results of optimal control for the ordinary piecewise linear systems in [20] to the case of uncertain piecewise linear systems. Thus, we can achieve the optimal control for the original chaotic system.

Note that the nonlinear term $F(\mathbf{x})$ in system (2.1) can be approximated by a piecewise linear function as follows:

$$F(\mathbf{x}) = K_i \mathbf{x} + a_i + \Delta_i(\mathbf{x}), \quad \mathbf{x} \in X_i, \quad i \in \mathbb{I}, \quad (2.3)$$

where $K_i \in \mathbb{R}^{n \times n}$, $a_i \in \mathbb{R}^n$ are some given parameters, $\{X_i\}_{i \in \mathbb{I}} \subseteq \mathbb{R}^n$ denotes a partition of the state space of chaotic system, \mathbb{I} is the index set, and $\Delta_i(\mathbf{x})$ is the approximation error, which can be regarded as uncertainties in the system. Then, it is obvious that system (2.1) can be converted to the uncertain piecewise linear system:

$$\dot{\mathbf{x}} = (A + K_i)\mathbf{x} + a_i + \Delta_i(\mathbf{x}) + B\mathbf{u}, \quad \mathbf{x} \in X_i, \quad i \in \mathbb{I}. \quad (2.4)$$

It is worth mentioning that system (2.1) can represent a large class of chaotic systems such as Genesio-Tesi chaotic system [23], Couillet chaotic system [24], Chua's Circuit system [25], and the new chaotic systems presented in [21, 26]. A simple but typical case is the three-dimensional chaotic system with the nonlinear term $F(\mathbf{x})$ taking the following form:

$$F(\mathbf{x}) = [0, 0, f(x_1)]^T, \quad (2.5)$$

where $f(x_1)$ is the nonlinear term in the 3rd dimension of the system and can be approximated by a piecewise linear function as

$$f(x_1) = k_i x_1 + l_i + \delta_i(x_1), \quad \mathbf{x} \in X_i, \quad i \in \mathbb{I}, \quad (2.6)$$

where $k_i, l_i \in \mathbb{R}$ are some given parameters, $\delta_i(x_1)$ is the approximation error. Then, system (2.1) with the nonlinear term (2.5) can be converted to the form of the uncertain piecewise linear system (2.4) as

$$\dot{\mathbf{x}} = A_i \mathbf{x} + a_i + \Delta_i + B\mathbf{u}, \quad \mathbf{x} \in X_i, \quad i \in \mathbb{I} \quad (2.7)$$

with

$$A_i = A + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ k_i & 0 & 0 \end{bmatrix}, \quad a_i = \begin{bmatrix} 0 \\ 0 \\ l_i \end{bmatrix}, \quad \Delta_i = \begin{bmatrix} 0 \\ 0 \\ \delta_i(x_1) \end{bmatrix}. \quad (2.8)$$

3. State Feedback Optimal Control of Systems

Without loss of generality, consider the uncertain piecewise linear system of the form

$$\dot{\mathbf{x}}(t) = (A_i + \Delta A_i)\mathbf{x}(t) + (B_i + \Delta B_i)\mathbf{u}(t) + a_i + \Delta a_i \quad (3.1)$$

for $\mathbf{x}(t) \in X_i$, where $\{X_i\}_{i \in \mathbb{I}} \subseteq \mathbb{R}^n$ denotes a partition of the state space into a number of polyhedral cells, \mathbb{I} is the index set of the cells, (A_i, B_i, a_i) is the i th nominal local model of the system, a_i is the offset term. ΔA_i , ΔB_i , and Δa_i represent parametric perturbations in the system state matrix, input matrix, and offset term of the i th nominal local model, respectively, and are assumed to be of the following form:

$$[\Delta A_i, \Delta B_i, \Delta a_i] = M_i H [N_{A_i}, N_{B_i}, N_{a_i}], \quad (3.2)$$

where $H \in \mathbb{R}^{i \times j}$ is an uncertain matrix bounded by $H^T H \leq I$, and $M_i, N_{A_i}, N_{B_i}, N_{a_i}$ are known constant matrices of appropriate dimensions which specify how the elements of the nominal matrices A_i, B_i , and a_i are affected by the uncertain parameters in H .

Define $\mathbb{I}_0 \subseteq \mathbb{I}$ as the set of indices for cells that contain origin and $\mathbb{I}_1 \subseteq \mathbb{I}$ the set of indices for cells that do not contain the origin. It is assumed that $a_i = \Delta a_i = 0$ for all $i \in \mathbb{I}_0$.

For any given initial condition $\mathbf{x}(0) = \mathbf{x}_0$, and input signals \mathbf{u} , it is assumed that system (3.1) has a unique solution, and there is no sliding mode. Note that with possible discontinuities in $A_i \mathbf{x}$ across the boundaries of the partitions, the solution of system (3.1) may be just continuous and piecewise C^1 . For a definition of the state trajectory of the system in (3.1) refer to [20] for details.

For convenience, the following notations are introduced:

$$\begin{aligned} \bar{\mathbf{x}} &= \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}, \quad \bar{A}_i = \begin{bmatrix} A_i & a_i \\ 0 & 0 \end{bmatrix}, \quad \bar{B}_i = \begin{bmatrix} B_i \\ 0 \end{bmatrix}, \quad \bar{M}_i = \begin{bmatrix} M_i \\ 0 \end{bmatrix}, \quad \bar{N}_{A_i} = [N_{A_i}, N_{a_i}], \\ \Delta \bar{A}_i &= \begin{bmatrix} \Delta A_i & \Delta a_i \\ 0 & 0 \end{bmatrix} = \bar{M}_i H \bar{N}_{A_i}, \quad \Delta \bar{B}_i = \begin{bmatrix} \Delta B_i \\ 0 \end{bmatrix} = \bar{M}_i H N_{B_i}, \end{aligned} \quad (3.3)$$

then system (3.1) can be expressed as

$$\dot{\bar{\mathbf{x}}}(t) = (\bar{A}_i + \Delta \bar{A}_i)\bar{\mathbf{x}}(t) + (\bar{B}_i + \Delta \bar{B}_i)\mathbf{u}(t), \quad i \in \mathbb{I}. \quad (3.4)$$

Associated with this system is the following cost function:

$$J = \int_0^\infty \left(\mathbf{x}^T(t) Q_i \mathbf{x}(t) + \mathbf{u}^T(t) R_i \mathbf{u}(t) \right) dt, \quad (3.5)$$

where i is defined so that $\mathbf{x}(t) \in X_i$, and $Q_i > 0$, $R_i > 0$ are given weighting matrices.

Note that if Q_i , R_i in (3.5) are set to be the same, respectively, for every $i \in \mathbb{I}$, the cost function (3.5) will reduce to (2.2). In addition, the matrix $\bar{Q}_i = \text{diag}\{Q_i, 0\} \in \mathbb{R}^{(n+1) \times (n+1)}$ is introduced, which will be used in the sequel.

As noted in [20], to find a piecewise Lyapunov function that is continuous across region boundaries, the matrices $\bar{F}_i = [F_i, f_i]$, $i \in \mathbb{I}$ with $f_i = 0$ for $i \in \mathbb{I}_0$ should be constructed, which are used to characterize the boundaries between the regions:

$$\bar{F}_i \bar{\mathbf{x}} = \bar{F}_j \bar{\mathbf{x}}, \quad \mathbf{x} \in X_i \cap X_j, \quad i, j \in \mathbb{I}. \quad (3.6)$$

Then, the piecewise Lyapunov function candidates that are continuous across the region boundaries can be parameterized as

$$V(\mathbf{x}) = \begin{cases} \mathbf{x}^T P_i \mathbf{x}, & \mathbf{x} \in X_i, \quad i \in \mathbb{I}_0, \\ \bar{\mathbf{x}}^T \bar{P}_i \bar{\mathbf{x}}, & \mathbf{x} \in X_i, \quad i \in \mathbb{I}_1, \end{cases} \quad (3.7)$$

with $P_i = F_i^T S F_i$ and $\bar{P}_i = \bar{F}_i^T S \bar{F}_i$, where S is a symmetric matrix which characterizes the free parameters of the Lyapunov function candidates.

Note the form of \bar{P}_i and the characteristics of the matrices \bar{F}_i . The continuity of the Lyapunov function $V(\mathbf{x})$ across the partition boundaries is ensured from (3.6) and (3.7).

The S -procedure has been used in [20, 22] to reduce the conservatism of the stability result. Specifically, the matrices $\bar{E}_i = [E_i, e_i]$, $i \in \mathbb{I}$ with $e_i = 0$ for $i \in \mathbb{I}_0$, such that

$$\bar{E}_i \bar{\mathbf{x}} \geq 0, \quad \mathbf{x} \in X_i, \quad i \in \mathbb{I}, \quad (3.8)$$

should be constructed to verify the positivity of a piecewise quadratic function of the form (3.7) on a polyhedral partition. It should be noted that the above vector inequalities imply that each entry of the vector is nonnegative.

A systematic procedure for constructing the matrices \bar{E}_i , \bar{F}_i for a given piecewise linear system was suggested in [20].

Consider the following piecewise linear feedback control law:

$$\mathbf{u} = -L_i \mathbf{x} - l_i := -\bar{L}_i \bar{\mathbf{x}}, \quad \mathbf{x} \in X_i, \quad i \in \mathbb{I}, \quad (3.9)$$

with $l_i = 0$ for $i \in \mathbb{I}_0$.

In general, the control law of form (3.9) will bring more flexibility in stability analysis than that of the ordinary linear feedback form. However, this control law may be discontinuous and give rise to sliding modes [20]. To avoid this case, we should construct

the control law continuously across subspace boundaries and take the feedback gain matrix \bar{L}_i as follows

$$\bar{L}_i = L^T \bar{F}_i, \quad i \in \mathbb{I}, \quad (3.10)$$

where L is a parameter matrix characterizing the free parameters of the state feedback controller, and \bar{F}_i is the matrix defined in (3.6). It should be pointed out that the gain matrix \bar{L}_i should take the form of $L_i = L^T F_i$ for $i \in \mathbb{I}_0$.

Substituting the control law (3.9) into system (3.4), we can get the following closed-loop system:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= (A_i + \Delta A_i - (B_i + \Delta B_i)L_i)\mathbf{x}(t), \quad \text{for } i \in \mathbb{I}_0, \\ \dot{\bar{\mathbf{x}}}(t) &= (\bar{A}_i + \Delta \bar{A}_i - (\bar{B}_i + \Delta \bar{B}_i)\bar{L}_i)\bar{\mathbf{x}}(t), \quad \text{for } i \in \mathbb{I}_1. \end{aligned} \quad (3.11)$$

Our goal in this section is to find a parameter matrix L to stabilize system (3.11) and meanwhile minimize the cost function (3.5). Before presenting the main results of this paper, we introduce the following lemmas.

Lemma 3.1 (Johansson and Rantzer [22]). *Consider symmetric matrices S , U_i , and W_i such that U_i and W_i have nonnegative entries, while $P_i = F_i^T S F_i$, $i \in \mathbb{I}_0$ and $\bar{P}_i = \bar{F}_i^T S \bar{F}_i$, $i \in \mathbb{I}_1$, satisfy*

$$\begin{aligned} A_i^T P_i + P_i A_i + E_i^T U_i E_i &< 0, \\ E_i^T W_i E_i &< P_i, \end{aligned} \quad (3.12)$$

for $i \in \mathbb{I}_0$, and

$$\begin{aligned} \bar{A}_i^T \bar{P}_i + \bar{P}_i \bar{A}_i + \bar{E}_i^T U_i \bar{E}_i &< 0, \\ \bar{E}_i^T W_i \bar{E}_i &< \bar{P}_i, \end{aligned} \quad (3.13)$$

for $i \in \mathbb{I}_1$, then every continuous and piecewise C^1 trajectory $\mathbf{x}(t)$ of system (3.4) with $\Delta A_i = 0$, $\Delta a_i = 0$ and $u = 0$ for all $t > 0$ tends to zero exponentially.

Lemma 3.2 (Xie [27]). *Given matrices G , M , and N of appropriate dimensions with G symmetric, then $G + MHN + N^T H^T M^T < 0$ for all matrices H satisfying $H^T H \leq I$, if and only if there exists some $\varepsilon > 0$ such that*

$$G + \varepsilon^{-1} M M^T + \varepsilon N^T N < 0. \quad (3.14)$$

Motivated by the result in [20], we can get the upper bound on the cost function (3.5) for uncertain piecewise linear systems based on the HJB inequality method. The result is presented as follows.

Theorem 3.3. Consider the closed-loop uncertain system (3.11) with $\mathbf{x}_0 \in X_{i_0}$. If there exist a set of constants $\varepsilon_i > 0$ and symmetric matrices S , U_i , and W_i such that U_i and W_i have nonnegative entries, while $P_i = F_i^T S F_i$, $i \in \mathbb{I}_0$, and $\bar{P}_i = \bar{F}_i^T S \bar{F}_i$, $i \in \mathbb{I}_1$, satisfy

$$\begin{bmatrix} \Phi_i & \varepsilon_i(N_{A_i} - N_{B_i}L_i)^T & P_i M_i & L_i^T \\ \varepsilon_i(N_{A_i} - N_{B_i}L_i) & -\varepsilon_i I & 0 & 0 \\ M_i^T P_i & 0 & -\varepsilon_i I & 0 \\ L_i & 0 & 0 & -R_i^{-1} \end{bmatrix} < 0, \quad (3.15)$$

$$E_i^T W_i E_i < P_i$$

for $i \in \mathbb{I}_0$,

$$\begin{bmatrix} \bar{\Phi}_i & \varepsilon_i(\bar{N}_{A_i} - N_{B_i}\bar{L}_i)^T & \bar{P}_i \bar{M}_i & \bar{L}_i^T \\ \varepsilon_i(\bar{N}_{A_i} - N_{B_i}\bar{L}_i) & -\varepsilon_i I & 0 & 0 \\ \bar{M}_i^T \bar{P}_i & 0 & -\varepsilon_i I & 0 \\ \bar{L}_i & 0 & 0 & -R_i^{-1} \end{bmatrix} < 0 \quad (3.16)$$

$$\bar{E}_i^T W_i \bar{E}_i < \bar{P}_i$$

for $i \in \mathbb{I}_1$, where

$$\begin{aligned} \Phi_i &:= (A_i - B_i L_i)^T P_i + P_i (A_i - B_i L_i) + E_i^T U_i E_i + Q_i, \\ \bar{\Phi}_i &:= (\bar{A}_i - \bar{B}_i \bar{L}_i)^T \bar{P}_i + \bar{P}_i (\bar{A}_i - \bar{B}_i \bar{L}_i) + \bar{E}_i^T U_i \bar{E}_i + \bar{Q}_i, \end{aligned} \quad (3.17)$$

then the closed-loop system is globally exponentially stable, and the cost function (3.5) satisfies

$$J \leq \inf_{S, U_i, W_i, \varepsilon_i} \bar{\mathbf{x}}_0^T \bar{P}_{i_0} \bar{\mathbf{x}}_0. \quad (3.18)$$

Proof. By Schur complement [19], the first inequality of (3.15) is equivalent to

$$\Phi_i + L_i^T R_i L_i + \varepsilon_i^{-1} P_i M_i M_i^T P_i + \varepsilon_i (N_{A_i} - N_{B_i} L_i)^T (N_{A_i} - N_{B_i} L_i) < 0. \quad (3.19)$$

Note the definitions of (3.3) and (3.17). By virtue of Lemma 3.2, inequality (3.19) is equivalent to

$$(A_i + \Delta A_i - (B_i + \Delta B_i)L_i)^T P_i + P_i (A_i + \Delta A_i - (B_i + \Delta B_i)L_i) + E_i^T U_i E_i + Q_i + L_i^T R_i L_i < 0. \quad (3.20)$$

Along a similar proof technique as used above, it can also be shown that the first inequality of (3.16) is equivalent to

$$\left(\bar{A}_i + \Delta \bar{A}_i - (\bar{B}_i + \Delta \bar{B}_i) \bar{L}_i\right)^T \bar{P}_i + \bar{P}_i \left(\bar{A}_i + \Delta \bar{A}_i - (\bar{B}_i + \Delta \bar{B}_i) \bar{L}_i\right) + \bar{E}_i^T U_i \bar{E}_i + \bar{Q}_i + \bar{L}_i^T R_i \bar{L}_i < 0, \quad (3.21)$$

where $\bar{Q}_i = \text{diag}\{Q_i, 0\}$. Note that $Q_i > 0$ and $R_i > 0$. By Lemma 3.1, it is obviously shown from inequalities (3.20), (3.21), and the second inequalities of (3.15) and (3.16) that the closed-loop system (3.11) is stable.

In addition, it can be seen from inequalities (3.20) and (3.21) that

$$\begin{aligned} & \left(\bar{A}_i + \Delta \bar{A}_i - (\bar{B}_i + \Delta \bar{B}_i) \bar{L}_i\right)^T \bar{P}_i + \bar{P}_i \left(\bar{A}_i + \Delta \bar{A}_i - (\bar{B}_i + \Delta \bar{B}_i) \bar{L}_i\right) \\ & + \bar{E}_i^T U_i \bar{E}_i + \bar{Q}_i + \bar{L}_i^T R_i \bar{L}_i \leq 0, \quad i \in \mathbb{I}. \end{aligned} \quad (3.22)$$

Multiplying from left and right by \bar{x}^T and \bar{x} , respectively, and removing the nonnegative term $\bar{x}^T \bar{E}_i^T U_i \bar{E}_i \bar{x}$ render

$$\frac{d}{dt} \left(\bar{x}^T \bar{P}_i \bar{x} \right) + \bar{x}^T \bar{Q}_i \bar{x} + \mathbf{u}^T R_i \mathbf{u} \geq 0. \quad (3.23)$$

Integration from 0 to ∞ , and noticing the global stability of closed-loop system (3.11), gives the result of (3.18). The proof is thus completed. \square

It is shown that the matrix inequalities (3.15) and (3.16) are BMIs due to the bilinear forms of $\bar{P}_i \bar{B}_i \bar{L}_i$ and $\varepsilon_i \bar{L}_i$ when both the Lyapunov matrix \bar{P}_i and the feedback gain matrix \bar{L}_i become the variables to be determined. Our interest is to find a parameter matrix L to minimize the upper bound $\bar{x}_0^T \bar{P}_{i_0} \bar{x}_0$ on the cost function (3.5) for the state feedback closed-loop system (3.11). Then, the optimization problem can be formulated as

$$\begin{aligned} & \min_{L, S, U_i, W_i, \varepsilon_i} \bar{x}_0^T \bar{P}_{i_0} \bar{x}_0 \\ & \text{s.t.} \quad \begin{cases} \bar{L}_i \in \mathbb{L} \\ (3.15)-(3.16), \end{cases} \end{aligned} \quad (3.24)$$

where $i \in \mathbb{I}$, and \mathbb{L} is the set of admissible values for the state feedback gain matrix \bar{L}_i , bounded by practical design constraints.

Remark 3.4. It should be noted that the optimization problem (3.24) is a nonconvex optimization problem with the BMIs constraints of (3.15) and (3.16). For BMIs problem, we [28] recently have already designed a mixed algorithm combining genetic algorithm (GA) and interior point method to solve it. Here, we can use the mixed algorithm proposed in [28] to obtain the optimal controller parameter matrix L and the corresponding objective $\bar{x}_0^T \bar{P}_{i_0} \bar{x}_0$.

In general, one can set the parameter matrix L to be the decision variables searched by GA. For a given chromosome corresponding to L , the nonconvex problem (3.24) reduces to an SDP involving LMIs which can be solved efficiently by Matlab LMI toolbox.

Remark 3.5. It should be pointed out that when solving the BMIs problem which is an NP hard problem in essence, the mixed algorithm combining GA with the interior point method may suffer from long computational time, especially for high-dimensional systems. Therefore, the optimal control problem can only be solved offline. In addition, the approximation error introduced by the linearization procedure for the chaotic system in Section 2 may adversely impact the stability analysis of the closed-loop system. To overcome this negative impact, one can divide the state space into a more sophisticated partition, but this will also increase the computational burden. Thus, one should seek a balance between the solution accuracy and the computational burden. On the other hand, for the chaotic systems there exists at least a bounded attractor. Due to the boundedness of the chaotic attractor, a relatively fine partition can be achieved to reduce the approximation error in the piecewise linearization procedure, which leads to a controller with a good performance.

To tell if the solutions obtained above are close to optimality or not, we must set up a lower bound on cost function (3.5). The result is presented as follows.

Theorem 3.6. *If there exist a set of constants $\varepsilon_i > 0$ and symmetric matrices S and U_i such that U_i have nonnegative entries, while $P_i = F_i^T S F_i$, $i \in \mathbb{I}_0$ and $\bar{P}_i = \bar{F}_i^T S \bar{F}_i$, $i \in \mathbb{I}_1$ satisfy*

$$\begin{bmatrix} \Psi_i & P_i B_i - \varepsilon_i N_{A_i}^T N_{B_i} & P_i M_i \\ B_i^T P_i - \varepsilon_i N_{B_i}^T N_{A_i} & R_i - \varepsilon_i N_{B_i}^T N_{B_i} & 0 \\ M_i^T P_i & 0 & \varepsilon_i I \end{bmatrix} > 0, \quad (3.25)$$

for $i \in \mathbb{I}_0$,

$$\begin{bmatrix} \bar{\Psi}_i & \bar{P}_i \bar{B}_i - \varepsilon_i \bar{N}_{A_i}^T \bar{N}_{B_i} & \bar{P}_i \bar{M}_i \\ \bar{B}_i^T \bar{P}_i - \varepsilon_i \bar{N}_{B_i}^T \bar{N}_{A_i} & R_i - \varepsilon_i \bar{N}_{B_i}^T \bar{N}_{B_i} & 0 \\ \bar{M}_i^T \bar{P}_i & 0 & \varepsilon_i I \end{bmatrix} > 0, \quad (3.26)$$

for $i \in \mathbb{I}_1$, where

$$\begin{aligned} \Psi_i &:= A_i^T P_i + P_i A_i + Q_i - E_i^T U_i E_i - \varepsilon_i N_{A_i}^T N_{A_i}, \\ \bar{\Psi}_i &:= \bar{A}_i^T \bar{P}_i + \bar{P}_i \bar{A}_i + \bar{Q}_i - \bar{E}_i^T U_i \bar{E}_i - \varepsilon_i \bar{N}_{A_i}^T \bar{N}_{A_i}, \end{aligned} \quad (3.27)$$

then for every trajectory $\mathbf{x}(t)$ of the uncertain system (3.4) with $\mathbf{x}(\infty) = 0$, $\mathbf{x}(0) = \mathbf{x}_0 \in X_{i_0}$, the cost function (3.5) satisfies

$$J \geq \sup_{S, U_i, \varepsilon_i} \bar{\mathbf{x}}_0^T \bar{P}_{i_0} \bar{\mathbf{x}}_0. \quad (3.28)$$

Proof. We will first show the conditions for the cost function (3.5) satisfying the lower bound (3.28) can be guaranteed by

$$\begin{bmatrix} (A_i + \Delta A_i)^T P_i + P_i(A_i + \Delta A_i) + Q_i - E_i^T U_i E_i & P_i(B_i + \Delta B_i) \\ (B_i + \Delta B_i)^T P_i & R_i \end{bmatrix} > 0, \quad (3.29)$$

for $i \in \mathbb{I}_0$, and

$$\begin{bmatrix} (\bar{A}_i + \Delta \bar{A}_i)^T \bar{P}_i + \bar{P}_i(\bar{A}_i + \Delta \bar{A}_i) + \bar{Q}_i - \bar{E}_i^T U_i \bar{E}_i & \bar{P}_i(\bar{B}_i + \Delta \bar{B}_i) \\ (\bar{B}_i + \Delta \bar{B}_i)^T \bar{P}_i & \bar{R}_i \end{bmatrix} > 0, \quad (3.30)$$

for $i \in \mathbb{I}_1$.

Actually, for $i \in \mathbb{I}$, we can get from (3.29) and (3.30) that

$$\begin{bmatrix} (\bar{A}_i + \Delta \bar{A}_i)^T \bar{P}_i + \bar{P}_i(\bar{A}_i + \Delta \bar{A}_i) + \bar{Q}_i - \bar{E}_i^T U_i \bar{E}_i & \bar{P}_i(\bar{B}_i + \Delta \bar{B}_i) \\ (\bar{B}_i + \Delta \bar{B}_i)^T \bar{P}_i & \bar{R}_i \end{bmatrix} \geq 0. \quad (3.31)$$

Multiplying from left and right by $[\bar{\mathbf{x}}^T, \mathbf{u}^T]$ and $[\bar{\mathbf{x}}^T, \mathbf{u}^T]^T$, respectively, and removing the nonnegative term $\bar{\mathbf{x}}^T \bar{E}_i^T U_i \bar{E}_i \bar{\mathbf{x}}$ yield

$$\begin{aligned} 0 &\leq 2\bar{\mathbf{x}}^T \bar{P}_i \left((\bar{A}_i + \Delta \bar{A}_i) \bar{\mathbf{x}} + (\bar{B}_i + \Delta \bar{B}_i) \mathbf{u} \right) + \bar{\mathbf{x}}^T \bar{Q}_i \bar{\mathbf{x}} + \mathbf{u}^T R_i \mathbf{u} \\ &= \frac{d}{dt} \left(\bar{\mathbf{x}}^T \bar{P}_i \bar{\mathbf{x}} \right) + \bar{\mathbf{x}}^T \bar{Q}_i \bar{\mathbf{x}} + \mathbf{u}^T R_i \mathbf{u}. \end{aligned} \quad (3.32)$$

Integration from 0 to ∞ , and noticing $\mathbf{x}(\infty) = 0$, gives the result of (3.28).

Next, we will show that inequality (3.29) is equivalent to (3.25). For simplifying the presentation, denote

$$G := \begin{bmatrix} A_i^T P_i + P_i A_i + Q_i - E_i^T U_i E_i & P_i B_i \\ B_i^T P_i & R_i \end{bmatrix}. \quad (3.33)$$

Note the uncertain form (3.2). Then, inequality (3.29) can be written as

$$G + \begin{bmatrix} P_i M_i \\ 0 \end{bmatrix} H [N_{A_i}, N_{B_i}] + [N_{A_i}, N_{B_i}]^T H^T \begin{bmatrix} P_i M_i \\ 0 \end{bmatrix}^T > 0. \quad (3.34)$$

By Lemma 3.2, inequality (3.34) is equivalent to the existence of some $\varepsilon_i > 0$ such that

$$G - \varepsilon_i^{-1} \begin{bmatrix} P_i M_i \\ 0 \end{bmatrix} \begin{bmatrix} P_i M_i \\ 0 \end{bmatrix}^T - \varepsilon_i [N_{A_i}, N_{B_i}]^T [N_{A_i}, N_{B_i}] > 0, \quad (3.35)$$

that is,

$$\begin{bmatrix} A_i^T P_i + P_i A_i + Q_i - E_i^T U_i E_i - \varepsilon_i N_{A_i}^T N_{A_i} - \varepsilon_i^{-1} P_i M_i M_i^T P_i & P_i B_i - \varepsilon_i N_{A_i}^T N_{B_i} \\ B_i^T P_i - \varepsilon_i N_{B_i}^T N_{A_i} & R_i - \varepsilon_i N_{B_i}^T N_{B_i} \end{bmatrix} > 0, \quad (3.36)$$

which, by Schur complement, is equivalent to inequality (3.25). By similar techniques, it can also be shown that inequality (3.30) is equivalent to inequality (3.26). The proof is complete. \square

Remark 3.7. It is shown that inequalities (3.25) and (3.26) are LMIs about the variables P_i , \bar{P}_i , and ε_i . So the problem of maximizing the lower bound (3.28) can be cast as an SDP with LMIs constraints of (3.25) and (3.26), and solved numerically effectively.

Remark 3.8. In the above analysis, it is assumed that the initial condition \mathbf{x}_0 is given or known in advance. Note that the bounds in (3.18) and (3.28) depend on the initial state \mathbf{x}_0 . To remove this dependence on the initial state, we can use the techniques developed in [28] and extend the corresponding results to the case where the initial condition \mathbf{x}_0 is a random variable subjected to uniform distribution on a certain bounded region X_0 . For further details, please refer to [28].

The global quadratic Lyapunov function technique is often applied in the control synthesis of dynamical systems [26]. In the following, by virtue of the global quadratic Lyapunov function technique and linear feedback control law, we present an optimal guaranteed cost control method for the chaotic system (2.1) associated with the cost function (2.2), which with the comparisons in the simulation results will show advantages of the obtained results in Theorems 3.3 and 3.6.

Consider the following linear feedback control law:

$$\mathbf{u} = -\tilde{L}\mathbf{x}. \quad (3.37)$$

Substituting the control law (3.37) into system (2.1), we can get the following closed-loop system:

$$\dot{\mathbf{x}}(t) = (A - B\tilde{L})\mathbf{x}(t) + F(\mathbf{x}). \quad (3.38)$$

Additionally, note the boundedness of the chaotic attractor and the Lipschitz continuity condition for the nonlinear term $F(\mathbf{x})$. There exist some matrix $\Gamma \geq 0$ and a bounded set Ω which bounds the chaotic attractor, such that

$$F^T(\mathbf{x})F(\mathbf{x}) \leq \mathbf{x}^T \Gamma^2 \mathbf{x}, \quad \forall \mathbf{x} \in \Omega. \quad (3.39)$$

The upper bound on the cost function (2.2) for the chaotic system (2.1) by applying linear feedback control law (3.37) is presented as follows.

Theorem 3.9. Consider system (2.1) with the initial condition $\mathbf{x}_0 \in \Omega$. If there exist positive constants α, β , positive definite matrix Y , and any matrix Z with appropriate dimensions such that

$$\begin{bmatrix} AY + YA^T - BZ - Z^T B^T + \alpha I & Y\Gamma & Y & Z^T \\ & \Gamma Y & -\alpha I & 0 & 0 \\ & Y & 0 & -Q^{-1} & 0 \\ & Z & 0 & 0 & -R^{-1} \end{bmatrix} < 0, \quad (3.40)$$

$$\begin{bmatrix} -\beta & \mathbf{x}_0^T \\ \mathbf{x}_0 & -Y \end{bmatrix} < 0,$$

then the closed-loop system (3.38) is globally exponentially stable, and the cost function (2.2) satisfies

$$J < \beta. \quad (3.41)$$

Furthermore, the corresponding control law can be obtained as $\mathbf{u} = -ZY^{-1}\mathbf{x}$.

Proof. Denote $P = Y^{-1} > 0$. Construct the Lyapunov function candidate as

$$V(\mathbf{x}) = \mathbf{x}^T P \mathbf{x}. \quad (3.42)$$

By virtue of the fact that $M^T N + N^T M \leq \alpha^{-1} M^T M + \alpha N^T N$, for all $\alpha > 0$, and matrices M and N with appropriate dimensions, calculating the time derivative of $V(\mathbf{x})$ along the trajectory of the closed-loop system (3.38) and noticing (3.39) yield

$$\begin{aligned} \frac{dV(\mathbf{x})}{dt} &= \mathbf{x}^T \left((A - B\tilde{L})^T P + P(A - B\tilde{L}) \right) \mathbf{x} + F^T(\mathbf{x}) P \mathbf{x} + \mathbf{x}^T P F(\mathbf{x}) \\ &\leq \mathbf{x}^T \left((A - B\tilde{L})^T P + P(A - B\tilde{L}) + \alpha P^2 \right) \mathbf{x} + \alpha^{-1} F^T(\mathbf{x}) F(\mathbf{x}) \\ &\leq \mathbf{x}^T \left((A - B\tilde{L})^T P + P(A - B\tilde{L}) + \alpha P^2 + \alpha^{-1} \Gamma^2 \right) \mathbf{x}. \end{aligned} \quad (3.43)$$

On the other hand, by Schur complement, the first inequality of (3.40) is equivalent to

$$AY + YA^T - BZ - Z^T B^T + \alpha I + \alpha^{-1} Y \Gamma^2 Y + Y Q Y + Z^T R Z < 0. \quad (3.44)$$

Noticing $Y P = I$, $\tilde{L} = Z P$, pre- and postmultiplying both sides of (3.44) by P implies

$$(A - B\tilde{L})^T P + P(A - B\tilde{L}) + \alpha P^2 + \alpha^{-1} \Gamma^2 + Q + \tilde{L}^T R \tilde{L} < 0. \quad (3.45)$$

Thus, it follows from (3.43) and (3.45) that

$$\frac{dV(\mathbf{x})}{dt} + \mathbf{x}^T Q \mathbf{x} + \mathbf{x}^T \tilde{L}^T R \tilde{L} \mathbf{x} < 0. \quad (3.46)$$

Note that $Q > 0$ and $R > 0$. It is obvious that $dV(\mathbf{x})/dt < 0$ which guarantees the global stability of closed-loop system (3.38), that is, $\mathbf{x}(\infty) = 0$.

Integration both sides of (3.46) from 0 to ∞ , and noticing $V(\mathbf{x}(\infty)) = 0$, renders

$$J < V(\mathbf{x}_0) = \mathbf{x}_0^T Y^{-1} \mathbf{x}_0, \quad (3.47)$$

with which combining the second inequality of (3.40) shows the result of (3.41). The proof is complete. \square

Remark 3.10. It is shown that the inequalities in (3.40) are LMIs in the variables Y , Z , α , β . So the problem of minimizing the upper bound (3.41) can be cast as an SDP with LMIs constraints of (3.40) and can be solved numerically effectively. On the other hand, it should be pointed out that the control synthesis methods based on the global quadratic Lyapunov function (3.42) and linear feedback control law (3.37) are conservative in practice compared with those in Theorems 3.3 and 3.6, which will be shown in illustrative examples.

4. Illustrative Examples

In this section, we will give two examples to illustrate the effectiveness of the proposed methods.

4.1. Genesio—Tesi Chaotic System

Consider the Genesio—Tesi chaotic system presented in [23], and the controlled system is described as follows:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -p_1 & -p_2 & -p_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ x_1^2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} u, \quad (4.1)$$

where $p_1 = 6$, $p_2 = 2.92$, $p_3 = 1.2$.

Denote that $c^T = [1, 0, 0]$ and $\mathbf{x}^T = [x_1, x_2, x_3]$. Note the boundedness of the chaotic attractor shown in [23]. The state space can be confined to $X := \{\mathbf{x} \mid -6 \leq c^T \mathbf{x} \leq 6\}$ by simulation. The partition of state space is set to be

$$\begin{aligned} X_1 &= \{\mathbf{x} \mid c^T \mathbf{x} \in [-1, 1)\}, & X_2 &= \{\mathbf{x} \mid c^T \mathbf{x} \in [1, 3)\}, & X_3 &= \{\mathbf{x} \mid c^T \mathbf{x} \in [3, 6]\}, \\ X_4 &= \{\mathbf{x} \mid c^T \mathbf{x} \in [-3, -1)\}, & X_5 &= \{\mathbf{x} \mid c^T \mathbf{x} \in [-6, -3)\}. \end{aligned} \quad (4.2)$$

Then, the nonlinear term x_1^2 can be described as

$$x_1^2 = k_i x_1 + l_i + \delta_i(x_1), \quad x \in X_i, \quad i = 1, 2, 3, 4, 5, \quad (4.3)$$

where δ_i denotes the approximation error. Taking $k_1 = 0$, $k_2 = -k_4 = 4.5$, $k_3 = -k_5 = 9$, $l_1 = 0$, $l_2 = l_4 = -4.5$, $l_3 = l_5 = -18$, one can obtain that

$$\begin{aligned} |\delta_1(x_1)| &\leq |x_1|, & |\delta_2(x_1)| &\leq 1, & |\delta_3(x_1)| &\leq 2.25, \\ |\delta_4(x_1)| &\leq 1, & |\delta_5(x_1)| &\leq 2.25. \end{aligned} \quad (4.4)$$

Note the expressions (4.3) and (4.4). System (4.1) can be converted to the piecewise linear system (3.1) with

$$\begin{aligned} A_1 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -2.92 & -1.2 \end{bmatrix}, & A_2 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1.5 & -2.92 & -1.2 \end{bmatrix}, & A_3 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 3 & -2.92 & -1.2 \end{bmatrix}, \\ A_4 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -10.5 & -2.92 & -1.2 \end{bmatrix}, & A_5 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -15 & -2.92 & -1.2 \end{bmatrix}, & B_i &= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \\ a_1 &= 0, & a_2 = a_4 &= [0 \ 0 \ -4.5]^T, & a_3 = a_5 &= [0 \ 0 \ -18]^T, \\ \Delta A_1 &= M_1 H N_{A_1}, & \Delta A_2 = \Delta A_3 = \Delta A_4 = \Delta A_5 &= 0, & \Delta B_i &= 0, \\ \Delta a_1 &= 0, & \Delta a_2 = M_2 H N_{a_2}, & \Delta a_3 = M_3 H N_{a_3}, \\ \Delta a_4 &= M_4 H N_{a_4}, & \Delta a_5 &= M_5 H N_{a_5}, \\ M_i &= \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}, & N_{A_1} &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0.5 & 0 & 0 \end{bmatrix}, \\ N_{a_2} = N_{a_4} &= \begin{bmatrix} 0 \\ 0 \\ 0.5 \end{bmatrix}, & N_{a_3} = N_{a_5} &= \begin{bmatrix} 0 \\ 0 \\ 1.125 \end{bmatrix}, \end{aligned} \quad (4.5)$$

where $i = 1, \dots, 5$, and H is an uncertain matrix bounded by $H^T H \leq I$.

It is worthwhile to mention that the nominal autonomous piecewise linear system (3.1) with parameters (4.5), that is, $u \equiv 0$, $\Delta A_i = 0$, $\Delta a_i = 0$, can exhibit chaotic dynamics, and the strange attractor is depicted in Figure 1. It is shown from Figure 1 that the system (3.1) with parameters (4.5) evolves to a single-scroll chaotic attractor, which is similar to the Genesio-Tesi chaotic attractor. Thus, it is indicated that the piecewise linear system approximating a chaotic system can preserve the complex dynamic behaviors of the original system.

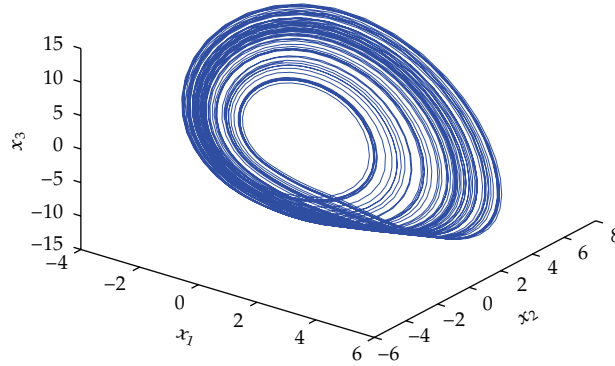


Figure 1: Phase portraits of the nominal autonomous system (3.1) with parameters (4.5).

Consider the cost function (2.2) with $Q = \text{diag}\{1, 1, 1\}$, $R = 1$, and the initial value $\mathbf{x}_0 = [-1.8, -1, 1]^T$ of system (4.1). The matrices \bar{E}_i and \bar{F}_i can be constructed by virtue of the method proposed in [20]. Assume that the feedback gain matrix \bar{L}_i is bounded by $\|\bar{L}_i\|_\infty < 12$, where $\|\bar{L}_i\|_\infty$ denotes the largest absolute value among all the entries of vector \bar{L}_i . Then, applying the mixed algorithm provided in [28], we solve the BMIs problem (3.24) based on Theorem 3.3 with the code written in MATLAB 7.0 and get the optimal upper bound on J , denoted as \bar{J}^* , and the corresponding optimal parameter matrix L^* as follows:

$$\begin{aligned} \bar{J}^* &= 17.7528, \\ L^* &= [-3.3088, -0.3345, 2.7111, 1.2363, -0.0548]^T. \end{aligned} \quad (4.6)$$

According to the expression of (3.10), we can get the following state feedback gain matrices:

$$\begin{aligned} L_1 &= [2.7111, 1.2363, -0.0548], \\ \bar{L}_2 &= [2.3765, 1.2363, -0.0548, 0.3345], \\ \bar{L}_3 &= [6.0198, 1.2363, -0.0548, -10.5954], \\ \bar{L}_4 &= [6.0198, 1.2363, -0.0548, 3.3088], \\ \bar{L}_5 &= [2.3765, 1.2363, -0.0548, -7.6212], \end{aligned} \quad (4.7)$$

with which the optimal control u taking the form of piecewise linear feedback control law (3.9) can be obtained.

Actually, the cost function (2.2) for the closed-loop system (4.1) with above controller gain matrices is computed as $J = 13.1623$. The numerical simulation of system (4.1) with the piecewise linear state feedback control is shown in Figure 2.

In addition, according to Theorem 3.6, the maximal lower bound on J , denoted as \underline{J}^* , can be obtained by solving the corresponding SDP with the LMI toolbox in MATLAB 7.0 as follows:

$$\underline{J}^* = 10.2047. \quad (4.8)$$

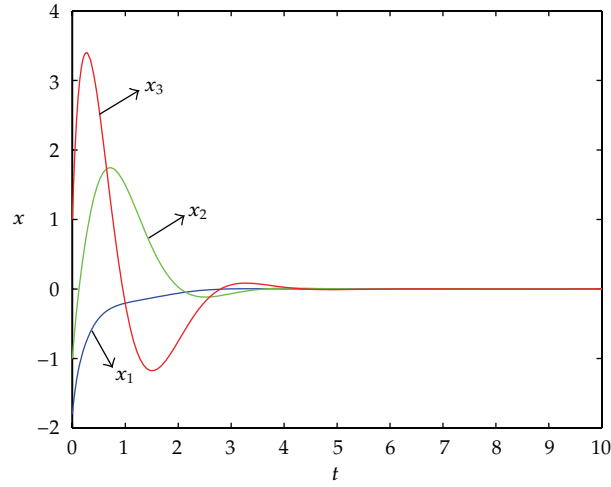


Figure 2: Time response of system (4.1) with the piecewise linear state feedback.

On the other hand, note that $-6 \leq c^T \mathbf{x} \leq 6$. The matrix Γ in (3.39) can be obtained as $\Gamma = \text{diag}\{6, 0, 0\}$. According to Theorem 3.9, we solve the corresponding SDP, and obtain the optimal gain matrix \tilde{L}^* in (3.37) and upper bound β^* as follows:

$$\tilde{L}^* = ZY^{-1} = [14.1843, 1.1514, -0.3980], \quad \beta^* = 53.0699, \quad (4.9)$$

which shows a fact that the optimal control methods based on the global quadratic Lyapunov function are conservative compared with those in Theorem 3.3.

4.2. A New Chaotic System

Consider the new chaotic system presented in [26], and the controlled system is described as follows:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -p_1 & -p_1 & -p_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ p_2 \tanh(x_1) \end{bmatrix} + \begin{bmatrix} 0.5 \\ 1 \\ 1 \end{bmatrix} u, \quad (4.10)$$

where $p_1 = 0.5$, $p_2 = 5$, and the hyperbolic function $\tanh(x) = (\exp(x) - \exp(-x)) / (\exp(x) + \exp(-x))$. The strange attractor of the autonomous system (4.10) with $u \equiv 0$ is shown in Figure 3, which is a double-scroll chaotic attractor.

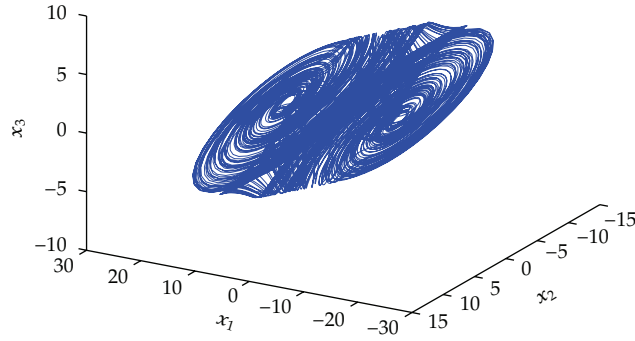


Figure 3: Phase portraits of the autonomous system (4.10).

Note the boundedness of the chaotic attractor shown in Figure 3. The state space can be confined to $X := \{\mathbf{x} \mid -23.3 \leq c^T \mathbf{x} \leq 23.3\}$ by simulation. The partition of state space is set to be

$$\begin{aligned} X_1 &= \{\mathbf{x} \mid c^T \mathbf{x} \in [-23.3, -1.18)\}, & X_2 &= \{\mathbf{x} \mid c^T \mathbf{x} \in [-1.18, 1.18)\}, \\ X_3 &= \{\mathbf{x} \mid c^T \mathbf{x} \in [1.18, 23.3]\}. \end{aligned} \quad (4.11)$$

Then, the nonlinear term $\tanh(x_1)$ can be described as

$$\tanh(x_1) = \begin{cases} k_1 x_1 + l_1 + \delta_1(x_1), & \mathbf{x} \in X_1, \\ k_2 x_1 + l_2 + \delta_2(x_1), & \mathbf{x} \in X_2, \\ k_3 x_1 + l_3 + \delta_3(x_1), & \mathbf{x} \in X_3, \end{cases} \quad (4.12)$$

where δ_i denotes the approximation error. Taking $k_1 = 0$, $k_2 = 0.85$, $k_3 = 0$, $l_1 = -1$, $l_2 = 0$, $l_3 = 1$, one can obtain that

$$|\delta_1(x_1)| \leq 0.17, \quad |\delta_2(x_1)| \leq 0.15|x_1|, \quad |\delta_3(x_1)| \leq 0.17. \quad (4.13)$$

Note the expressions (4.12) and (4.13). System (4.10) can be converted to the piecewise linear system (3.1) with

$$\begin{aligned} A_1 = A_3 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -p_1 & -p_1 & -p_1 \end{bmatrix}, & A_2 &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -p_1 + 0.85p_2 & -p_1 & -p_1 \end{bmatrix}, \\ B_1 = B_2 = B_3 &= \begin{bmatrix} 0.5 \\ 1 \\ 1 \end{bmatrix}, & a_1 = -a_3 &= \begin{bmatrix} 0 \\ 0 \\ -p_2 \end{bmatrix}, & a_2 &= 0, \end{aligned}$$

$$\begin{aligned}
\Delta A_1 = \Delta A_3 = 0, \quad \Delta A_2 = M_2 H N_{A_2}, \quad \Delta B_1 = \Delta B_2 = \Delta B_3 = 0, \\
\Delta a_1 = M_1 H N_{a_1}, \quad \Delta a_2 = 0, \quad \Delta a_3 = M_3 H N_{a_3}, \\
M_1 = M_2 = M_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad N_{A_2} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0.15p_2 & 0 & 0 \end{bmatrix}, \quad N_{a_1} = N_{a_3} = \begin{bmatrix} 0 \\ 0 \\ 0.17p_2 \end{bmatrix}.
\end{aligned} \tag{4.14}$$

Consider the cost function (2.2) with $Q = \text{diag}\{0.8, 0.8, 0.8\}$, $R = 1.2$, and the system initial value $x_0 = [1.4, 1, -0.6]^T$. Assume that the feedback gain matrix \bar{L}_i is bounded by $\|\bar{L}_i\|_\infty < 8$. Then, similarly to the above subsection, we get the maximal lower bound \underline{J}^* , the optimal upper bound \bar{J}^* , and the corresponding optimal parameter matrix L^* as follows:

$$\begin{aligned}
\underline{J}^* = 5.5117, \quad \bar{J}^* = 9.7024, \\
L^* = [-4.3655, 1.2292, 2.3555, 1.3460, 1.4357]^T.
\end{aligned} \tag{4.15}$$

According to the expression of (3.10), we can get the following state feedback gain matrices:

$$\begin{aligned}
\bar{L}_1 = [6.7210, 1.3460, 1.4357, 5.1513], \quad L_2 = [2.3555, 1.3460, 1.4357], \\
\bar{L}_3 = [3.5847, 1.3460, 1.4357, -1.4505],
\end{aligned} \tag{4.16}$$

with which the optimal control u taking the form of (3.9) is obtained.

Additionally, the cost function (2.2) for the closed-loop system (4.10) with above controller gain matrices is computed as $J = 7.8725$. The numerical simulation of system (4.10) with piecewise linear state feedback control is shown in Figure 4.

Furthermore, note that $\tanh^2(x_1) \leq x_1^2$. The matrix Γ in (3.39) can be obtained as $\Gamma = \text{diag}\{p_2, 0, 0\}$. According to Theorem 3.9, we solve the corresponding SDP, and obtain the optimal gain matrix \tilde{L}^* in (3.37) and upper bound β^* as follows:

$$\tilde{L}^* = ZY^{-1} = [19.2415, 2.5432, -0.2071], \quad \beta^* = 98.965, \tag{4.17}$$

which is significantly greater than the optimal upper bound \bar{J}^* obtained from Theorem 3.3.

It is obviously shown from the above examples that the optimal upper bounds \bar{J}^* obtained above get close to the corresponding lower bounds \underline{J}^* , respectively. This implies that we have achieved or got close to the optimal control for the chaotic systems. Additionally, it should be pointed out that the newly reported chaotic system (4.10) is topologically not equivalent to the Genesio-Tesi chaotic system (4.1). However, by virtue of the optimal control methods proposed in this paper, both the different chaotic systems (4.1) and (4.10) can be optimally stabilized. The examples show the effectiveness of the proposed results.

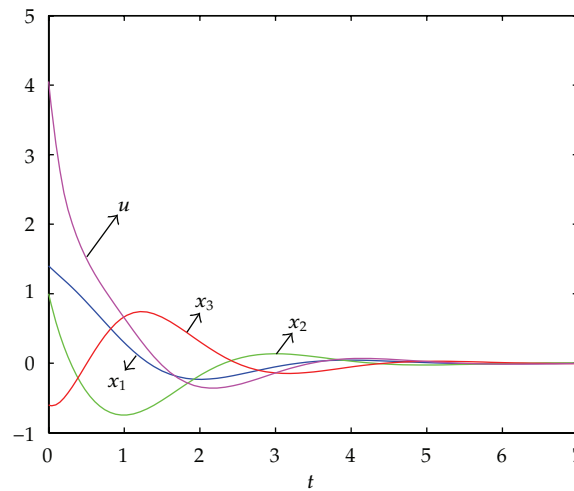


Figure 4: The control law and time response of the controlled system (4.10).

5. Conclusion

In this paper, we first convert a class of chaotic systems to the form of uncertain piecewise linear systems then investigate the optimal control for the chaotic systems where the piecewise linear state feedback optimal controller can be obtained by solving an optimization problem with BMIs constraints. The performance of the controller can be evaluated by the upper and lower bounds on the cost function. The optimal chaos synchronization for this class of chaotic systems will be studied in the near future.

Acknowledgment

The authors thank the anonymous referees and editor for their valuable comments and suggestions. This work was supported by the National Natural Science Foundation of China (no. 61004015), the Research Fund for the Doctoral Programme of Higher Education of China (no. 20090032120034), the Program for New Century Excellent Talents in Universities of China, and the Program for Changjiang Scholars and Innovative Research Team in University of China (no. IRT1028).

References

- [1] G. Chen and X. Dong, *From Chaos to Order*, vol. 24 of *Methodologies, Perspectives and Applications*, World Scientific, Singapore, 1998.
- [2] A. L. Fradkov and R. J. Evans, "Control of chaos: methods and applications in engineering," *Annual Reviews in Control*, vol. 29, no. 1, pp. 33–56, 2005.
- [3] E. Ott, C. Grebogi, and J. A. Yorke, "Controlling chaos," *Physical Review Letters*, vol. 64, no. 11, pp. 1196–1199, 1990.
- [4] C. Hernandez-Tenorio, T. L. Belyaeva, and V. N. Serkin, "Parametric resonance for solitons in the nonlinear Schrödinger equation model with time-dependent harmonic oscillator potential," *Physica B*, vol. 398, no. 2, pp. 460–463, 2007.
- [5] H. Sun and H. Cao, "Chaos control and synchronization of a modified chaotic system," *Chaos, Solitons and Fractals*, vol. 37, no. 5, pp. 1442–1455, 2008.

- [6] G. Chen, "A simple adaptive feedback control method for chaos and hyper-chaos control," *Applied Mathematics and Computation*, vol. 217, no. 17, pp. 7258–7264, 2011.
- [7] B. Niu and J. Wei, "Stability and bifurcation analysis in an amplitude equation with delayed feedback," *Chaos, Solitons & Fractals*, vol. 37, no. 5, pp. 1362–1371, 2008.
- [8] M. T. Yassen, "Chaos control of chaotic dynamical systems using backstepping design," *Chaos, Solitons and Fractals*, vol. 27, no. 2, pp. 537–548, 2006.
- [9] M. S. Tavazoei, M. Haeri, and S. Jafari, "Fractional controller to stabilize fixed points of uncertain chaotic systems: Theoretical and experimental study," *Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering*, vol. 222, no. 3, pp. 175–184, 2008.
- [10] J. M. Nazzal and A. N. Natsheh, "Chaos control using sliding-mode theory," *Chaos, Solitons and Fractals*, vol. 33, no. 2, pp. 695–702, 2007.
- [11] M. Roopaei, B. Ranjbar Sahraei, and T.-C. Lin, "Adaptive sliding mode control in a novel class of chaotic systems," *Communications in Nonlinear Science and Numerical Simulation*, vol. 15, no. 12, pp. 4158–4170, 2010.
- [12] W. J. Sun, "A global asymptotic synchronization problem via internal model approach," *International Journal of Control, Automation, and Systems*, vol. 8, no. 5, pp. 1153–1158, 2010.
- [13] J. Wang, W. Tang, and J. Zhang, "Approximate synchronization of two non-linear systems via impulsive control," *Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering*, vol. 226, no. 13, pp. 338–347, 2012.
- [14] X. R. Chen and C. X. Liu, "Chaos synchronization of fractional order unified chaotic system via non-linear control," *International Journal of Modern Physics B*, vol. 25, pp. 407–415, 2011.
- [15] A. Uçar, K. E. Lonngren, and E. W. Bai, "Synchronization of the unified chaotic systems via active control," *Chaos, Solitons and Fractals*, vol. 27, no. 5, pp. 1292–1297, 2006.
- [16] Q. Jia, "Chaos control and synchronization of the Newton-Leipnik chaotic system," *Chaos, Solitons and Fractals*, vol. 35, no. 4, pp. 814–824, 2008.
- [17] G. M. Mahmoud, T. Bountis, G. M. Abdel-Latif, and E. E. Mahmoud, "Chaos synchronization of two different chaotic complex Chen and Lü systems," *Nonlinear Dynamics*, vol. 55, no. 1-2, pp. 43–53, 2009.
- [18] K. Tanaka, T. Ikeda, and H. O. Wang, "A unified approach to controlling chaos via an LMI-based fuzzy control system design," *IEEE Transactions on Circuits and Systems*, vol. 45, no. 10, pp. 1021–1040, 1998.
- [19] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, vol. 15, Society for Industrial and Applied Mathematics SIAM, Philadelphia, Pa, USA, 1994.
- [20] M. Johansson, *Piecewise Linear Control Systems*, vol. 284, Springer, Berlin, Germany, 2003.
- [21] J. Zhang and W. Tang, "Analysis and control for a new chaotic system via piecewise linear feedback," *Chaos, Solitons and Fractals*, vol. 42, no. 4, pp. 2181–2190, 2009.
- [22] M. Johansson and A. Rantzer, "Computation of piecewise quadratic Lyapunov functions for hybrid systems," *IEEE Transactions on Automatic Control*, vol. 43, no. 4, pp. 555–559, 1998.
- [23] R. Genesio and A. Tesi, "Harmonic balance methods for the analysis of chaotic dynamics in nonlinear systems," *Automatica*, vol. 28, no. 3, pp. 531–548, 1992.
- [24] A. Arneodo, P. Couillet, and C. Tresser, "A possible new mechanism for the onset of turbulence," *Physics Letters A*, vol. 81, no. 4, pp. 197–201, 1981.
- [25] A. Tsuneda, "A gallery of attractors from smooth Chua's equation," *International Journal of Bifurcation and Chaos*, vol. 15, no. 1, pp. 1–49, 2005.
- [26] J. Zhang and W. Tang, "Control and synchronization for a class of new chaotic systems via linear feedback," *Nonlinear Dynamics*, vol. 58, no. 4, pp. 675–686, 2009.
- [27] L. Xie, "Output feedback H_∞ control of systems with parameter uncertainty," *International Journal of Control*, vol. 63, no. 4, pp. 741–750, 1996.
- [28] J. Zhang and W. Tang, "Output feedback optimal guaranteed cost control of uncertain piecewise linear systems," *International Journal of Robust and Nonlinear Control*, vol. 19, no. 5, pp. 569–590, 2009.

Research Article

Bifurcation Analysis for a Kind of Nonlinear Finance System with Delayed Feedback and Its Application to Control of Chaos

Rongyan Zhang

Department of Electronic Information Engineering, Huanghe Science and Technology College, Henan, Zhengzhou 450063, China

Correspondence should be addressed to Rongyan Zhang, hebgydx2006@163.com

Received 14 October 2011; Revised 15 January 2012; Accepted 23 January 2012

Academic Editor: Chuanhou Gao

Copyright © 2012 Rongyan Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A kind of nonlinear finance system with time-delayed feedback is considered. Firstly, by employing the polynomial theorem to analyze the distribution of the roots to the associate characteristic equation, the conditions of ensuring the existence of Hopf bifurcation are given. Secondly, by using the normal form theory and center manifold argument, we derive the explicit formulas determining the stability, direction, and other properties of bifurcating periodic solutions. Finally, we give several numerical simulations, which indicate that when the delay passes through certain critical values, chaotic oscillation is converted into a stable steady state or a stable periodic orbit.

1. Introduction

Since the chaotic phenomenon in economics was first found in 1985, great impact has been imposed on the prominent western economics at present, because the chaotic phenomenon occurring in the economic system means that the macroeconomic operation has in itself the inherent indefiniteness. Although the government can adopt such macrocontrol measures as the financial policies or the monetary policies to interfere, the effectiveness of the interference is very limited. The instability and complexity make the precise economic prediction greatly limited, and the reasonable prediction behavior has become complicated as well. In the fields of finance, stocks, and social economics, because of the interaction between nonlinear factors, with all kinds of economic problems being more and more complicated and with the evolution process from low dimensions to high dimensions, the diversity and complexity have manifested themselves in the internal structure of the system and there exists extremely

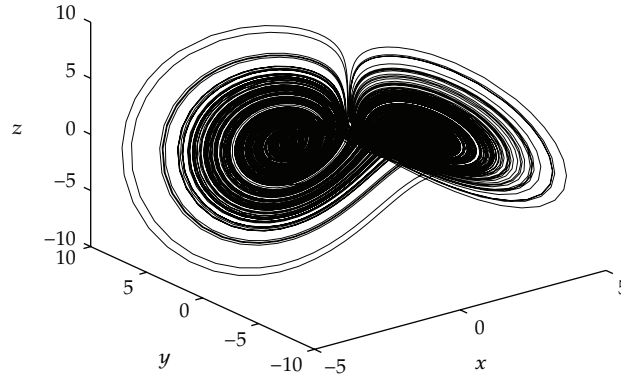


Figure 1: Strange attractor of finance system (1.1).

complicated phenomenon and external characteristics in such a kind of system. So it has become more and more important to study the control of the complicated continuous economic system and stabilize the instable periodic or stationary solutions, in order to make the precise economic prediction possible [1, 2].

Recent works [1, 2] have reported a dynamic model of finance, composed of three first-order differential equations. The model describes the time variations of three state variables: the interest rate x , the investment demand y , and the price index z . By choosing an appropriate coordinate system and setting appropriate dimensions for each state variable, [1, 2] offer the simplified finance system as

$$\begin{aligned}\dot{x}(t) &= -a(x(t) + y(t)), \\ \dot{y}(t) &= -y(t) - ax(t)z(t), \\ \dot{z}(t) &= b + ax(t)y(t),\end{aligned}\tag{1.1}$$

which is chaotic when $a = 1.69$, $b = 4$ (see Figure 1).

Over the last years, [3, 4] studied impulsive control and state feedback control of the finance system (1.1). In this paper, we are interesting in delayed feedback control of the finance system (1.1). The effects of the time-delayed feedback on the finance system have long been investigated [5–8].

Recently, different techniques and methods have been proposed to achieve chaos control. The existing control methods can be classified, mainly, into two categories. The first one, developed by Ott et al. [9] is based on the invariant manifold structure of unstable orbits. It is theoretically well understood but difficult to apply to fast experimental systems. The second, proposed by Pyragas [10], uses time-delayed controlling forces. In contrast to the former one, it is simple and convenient method of controlling chaos in continuous dynamical system. Thus, we adopt the second one in the present paper.

For predigesting the investigation, here we only put time delay on investment demand y . By adding a time-delayed force $K(y(t) - y(t - \tau))$ to the second equation of finance system (1.1), we obtain the following new system

$$\begin{aligned}\dot{x}(t) &= -a(x(t) + y(t)), \\ \dot{y}(t) &= -y(t) - ax(t)z(t) + K(y(t) - y(t - \tau)), \\ \dot{z}(t) &= b + ax(t)y(t).\end{aligned}\quad (1.2)$$

Here we assume that $(C_1) a, b, \tau \in (0, \infty)$ and $K \in \mathbb{R}$. The time delay τ is taken as the bifurcation parameter and we show that when τ passes through some certain critical values, the equilibrium will lose its stability and hopf bifurcation will take place; by adjusting K values, we achieve the purpose of chaos control. The research of this paper is a new investigation about the hopf bifurcation and chaos control on the finance system and has important theoretical and practical value.

2. Stability of Steady States and Bifurcations of Periodic Solutions

In this section, we investigate the effect of delay on the dynamic behavior of system (1.2). Obviously, when $\tau = 0$, system (1.2) becomes the system (1.1). First, we introduce the following several lemmas in [1, 2] for T's system(1.1).

We know that under the assumption (C_1) , the system (1.1) has two equilibrium points:

$$S_1 = \left(\frac{\sqrt{ab}}{b}, -\frac{\sqrt{ab}}{b}, \frac{1}{a} \right), \quad S_2 = \left(-\frac{\sqrt{ab}}{b}, \frac{\sqrt{ab}}{b}, \frac{1}{a} \right). \quad (2.1)$$

The characteristic equation of the system (1.1) at $S_1(S_2)$ is

$$\lambda^3 + (1 + a)\lambda^2 + ab\lambda + 2a^2b = 0. \quad (2.2)$$

By analyzing the characteristic equation (2.2) and the Routh-Hurwitz criteria, we get the following.

Lemma 2.1. *For $a < 1$, the characteristic equation (2.2) has three eigenvalues with negative real parts, so two equilibrium points S_1, S_2 of the system (1.1) are asymptotic stable.*

Lemma 2.2. *For $a = 1$, the characteristic equation (2.2) has a pair of purely imaginary eigenvalues $\lambda_{1,2} = \pm i\omega_0$ ($\omega_0 = \sqrt{b}$) and a negative real eigenvalue $\lambda_3 = -2$, and*

$$\dot{\lambda}(a = 1) = \frac{5b}{2b + 8} > 0. \quad (2.3)$$

According to the hopf bifurcation theorem [11], a hopf bifurcation of the system (1.1) occurs at $a = 1$.

Lemma 2.3. *For $a > 1$, the characteristic equation (2.2) has one negative real root and one pair of conjugate complex roots with positive real parts, so two equilibrium points S_1, S_2 of the system (1.1) are unstable.*

Clearly, the delayed feedback control system (1.2) has the same equilibria to the corresponding system (1.1). In this section, we analyze the effect of delay on the stability of these steady states. Due to the symmetry of S_1 and S_2 , it is sufficient to analyze the stability of S_1 . By the linear transform

$$\begin{aligned}x_1(t) &= x(t) - \frac{\sqrt{ab}}{b}, \\y_1(t) &= y(t) + \frac{\sqrt{ab}}{b}, \\z_1(t) &= z(t) - \frac{1}{a},\end{aligned}\tag{2.4}$$

system (1.2) becomes

$$\begin{aligned}\dot{x}_1(t) &= -a(x_1(t) + y_1(t)), \\ \dot{y}_1(t) &= -x_1(t) - y_1(t) - \sqrt{ab}z_1(t) - ax_1(t)z_1(t) + K(y_1(t) - y_1(t - \tau)), \\ \dot{z}_1(t) &= -\sqrt{ab}x_1(t) + \sqrt{ab}y_1(t) + ax_1(t)y_1(t).\end{aligned}\tag{2.5}$$

It is easy to see that the origin $S_0(0, 0, 0)$ is the equilibrium of system (2.5). The associated characteristic equation of system (2.5) at $S_0(0, 0, 0)$ is

$$\det \begin{pmatrix} \lambda + a & a & 0 \\ 1 & \lambda + 1 - K + Ke^{-\lambda\tau} & \sqrt{ab} \\ \sqrt{ab} & -\sqrt{ab} & \lambda \end{pmatrix} = 0.\tag{2.6}$$

Expanding (2.6), we have

$$\lambda^3 + (1 + a - K)\lambda^2 + (ab - aK)\lambda + 2a^2b + K(\lambda^2 + a\lambda)e^{-\lambda\tau} = 0.\tag{2.7}$$

Thus, we need to study the distribution of the roots of the third-degree exponential polynomial equation:

$$\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0 + (b_2\lambda^2 + b_1\lambda + b_0)e^{-\lambda\tau} = 0,\tag{2.8}$$

where $a_i, b_i \in \mathbb{R}$ ($i = 0, 1, 2$) and $\sum_{i=0}^2 b_i^2 \neq 0$. We first introduce the following simple result which was proved by Ruan and Wei [12] using Rouché's theorem.

Lemma 2.4. *Consider the exponential polynomial*

$$\begin{aligned}P(\lambda, e^{-\lambda\tau_1}, \dots, e^{-\lambda\tau_m}) \\ = \lambda^n + p_1^{(0)}\lambda^{n-1} + \dots + p_{n-1}^{(0)}\lambda + p_n^{(0)} + [p_1^{(1)}\lambda^{n-1} + \dots + p_{n-1}^{(1)}\lambda + p_n^{(1)}]e^{-\lambda\tau_1} \\ + \dots + [p_1^{(m)}\lambda^{n-1} + \dots + p_{n-1}^{(m)}\lambda + p_n^{(m)}]e^{-\lambda\tau_m},\end{aligned}\tag{2.9}$$

where $\tau_i \geq 0$ ($i = 1, 2, \dots, m$) and $p_j^{(i)}$ ($i = 0, 1, \dots, m, j = 1, 2, \dots, n$) are constants. As $(\tau_1, \tau_2, \dots, \tau_m)$ vary, the sum of the order of the zeros of $P(\lambda, e^{-\lambda\tau_1}, \dots, e^{-\lambda\tau_m})$ on the open right half plane can change only if a zero appears on or crosses the imaginary axis.

Obviously, $i\omega$ ($\omega > 0$) is a root of (2.8) if and only if ω satisfies

$$-i\omega^3 - a_2\omega^2 + a_1\omega i + a_0 + (-b_2\omega^2 + b_1\omega i + b_0)(\cos \omega\tau - i \sin \omega\tau) = 0. \quad (2.10)$$

Separating the real and imaginary parts, we have

$$\begin{aligned} a_2\omega^2 - a_0 &= (b_0 - b_2\omega^2) \cos \omega\tau + b_1\omega \sin \omega\tau, \\ -\omega^3 + a_1\omega &= (b_0 - b_2\omega^2) \sin \omega\tau - b_1\omega \cos \omega\tau, \end{aligned} \quad (2.11)$$

which is equivalent to

$$\omega^6 + (a_2^2 - b_2^2 - 2a_1)\omega^4 + (a_1^2 - 2a_0a_2 - b_1^2 + 2b_0b_2)\omega^2 + a_0^2 - b_0^2 = 0. \quad (2.12)$$

Let $z = \omega^2$ and denote $p = a_2^2 - b_2^2 - 2a_1$, $q = a_1^2 - 2a_0a_2 - b_1^2 + 2b_0b_2$, $r = a_0^2 - b_0^2$, then (2.12) becomes

$$z^3 + pz^2 + qz + r = 0. \quad (2.13)$$

In the following, we need to seek conditions under which (2.12) has at least one positive root. Denote

$$h(z) = z^3 + pz^2 + qz + r. \quad (2.14)$$

Therefor, applying [13], we obtain the following lemma.

Lemma 2.5. *For the polynomial equation (2.13), one has the following results.*

- (i) *If $r < 0$, then (2.13) has at least one positive root.*
- (ii) *If $r \geq 0$ and $\Delta = p^2 - 3q \leq 0$, then (2.13) has no positive roots.*
- (iii) *If $r \geq 0$ and $\Delta = p^2 - 3q > 0$, then (2.13) has positive roots if and only if $z_1^* = (1/3)(-p + \sqrt{\Delta}) > 0$ and $h(z_1^*) \leq 0$.*

Suppose that (2.13) has positive roots. Without loss of generality, we assume that it has three positive roots, defined by z_1 , z_2 , and z_3 , respectively. Then (2.12) has three positive roots:

$$\omega_1 = \sqrt{z_1}, \quad \omega_2 = \sqrt{z_2}, \quad \omega_3 = \sqrt{z_3}. \quad (2.15)$$

From (2.11), we have

$$\cos \omega \tau = \frac{b_1 \omega^2 (\omega^2 - a_1) - (a_2 \omega^2 - a_0)(b_2 \omega^2 - b_0)}{(b_2 \omega^2 - b_0)^2 + b_1^2 \omega^2}. \quad (2.16)$$

Thus, if we denote

$$\tau_k^{(j)} = \frac{1}{\omega_k} \left[\cos^{-1} \left(\frac{b_1 \omega_k^2 (\omega_k^2 - a_1) - (a_2 \omega_k^2 - a_0)(b_2 \omega_k^2 - b_0)}{(b_2 \omega_k^2 - b_0)^2 + b_1^2 \omega_k^2} \right) + 2j\pi \right], \quad (2.17)$$

where $k = 1, 2, 3$; $j = 0, 1, 2, \dots$, then $\pm i\omega_k$ is a pair of purely imaginary roots of (2.8) with $\tau_k^{(j)}$. Define

$$\tau_0 = \tau_{k_0}^{(0)} = \min_{k \in 1, 2, 3} \tau_k^{(0)}, \quad \omega_0 = \omega_{k_0}. \quad (2.18)$$

Note that when $\tau = 0$, (2.8) becomes

$$\lambda^3 + (a_2 + b_2)\lambda^2 + (a_1 + b_1)\lambda + a_0 + b_0 = 0. \quad (2.19)$$

Therefor, applying Lemmas 2.4 and 2.5 to (2.8), we get the following lemma.

Lemma 2.6. *For (2.8), one has*

- (i) *if $r \geq 0$ and $\Delta = p^2 - 3q \leq 0$, then all roots with positive real parts of (2.8) have the same sum to those of the polynomial equation (2.19) for all $\tau \geq 0$.*
- (ii) *if either $r < 0$ or $r \geq 0$, $\Delta = p^2 - 3q > 0$, $z_1^* = (1/3)(-p + \sqrt{\Delta}) > 0$ and $h(z_1^*) \leq 0$, then all roots with positive real parts of (2.8) have the same sum to those of the polynomial equation (2.19) for $\tau \in [0, \tau_0]$.*

Let

$$\lambda(\tau) = \alpha(\tau) + i\omega(\tau) \quad (2.20)$$

be the root of (2.8) near $\tau = \tau_k^{(j)}$ satisfying

$$\alpha(\tau_k^{(j)}) = 0, \quad \omega(\tau_k^{(j)}) = \omega_k. \quad (2.21)$$

Then by [13], we have the following transversality condition.

Lemma 2.7. Suppose that $z_k = \omega_k^2$ and $h'(z_k) \neq 0$. Then

$$\frac{R\lambda(\tau_k^{(j)})}{d\tau} \neq 0, \quad (2.22)$$

and $R\lambda(\tau_k^{(j)})/d\tau$ and $h'(z_k)$ have the same sign.

Now, we study the characteristic equation (2.7) of the system (2.5). Comparing (2.7) with (2.8), we know that

$$a_2 = 1 + a - K, \quad a_1 = ab - aK, \quad a_0 = 2a^2b, \quad b_2 = K, \quad b_1 = aK, \quad b_0 = 0. \quad (2.23)$$

Thus,

$$\begin{aligned} p &= a_2^2 - b_2^2 - 2a_1 = a^2 + 2a + 1 - 2ab - 2K, \\ q &= a_1^2 - 2a_0a_2 - b_1^2 + 2b_0b_2 = a^2b^2 + 2a^2bK - 4a^2b - 4a^3b, \\ r &= a_0^2 - b_0^2 = 4a^4b^2 > 0, \end{aligned} \quad (2.24)$$

and then we can compute

$$\Delta = p^2 - 3q, \quad h(z) = z^3 + pz^2 + qz + r, \quad z_1^* = \frac{1}{3}(-p + \sqrt{\Delta}). \quad (2.25)$$

When $\tau = 0$, (2.7) becomes (2.2)

$$\lambda^3 + (1 + a)\lambda^2 + ab\lambda + 2a^2b = 0. \quad (2.26)$$

Applying Lemmas 2.1, 2.2, 2.6, and 2.7 to (2.7), we have the following theorems.

Theorem 2.8. Let $\tau_k^{(j)}$ and τ_0 be defined by (2.17) and (2.18). Suppose that conditions (C_1) and $a < 1$ hold.

- (i) If $\Delta \leq 0$, then (2.7) had all roots with negative real parts for all $\tau \geq 0$, and the equilibrium S_1 (or S_2) of the system (1.2) is stable.
- (ii) If $\Delta > 0$, $z_1^* > 0$ and $h(z_1^*) \leq 0$, (2.7) had all roots with negative real parts for $\tau \in [0, \tau_0)$, and the equilibrium S_1 (or S_2) of the system (1.2) is stable.
- (iii) If the conditions of (ii) are satisfied, and $h'(z_k) \neq 0$, then system (1.2) exhibits the Hopf bifurcation at the equilibrium S_1 (or S_2) for $\tau = \tau_k^{(j)}$.

Theorem 2.9. Let $\tau_k^{(j)}$ and τ_0 be defined by (2.17) and (2.18). Suppose that conditions (C_1) and $a > 1$ hold.

- (i) If $\Delta \leq 0$, then (2.7) had two roots with positive real parts for all $\tau \geq 0$, and the equilibrium S_1 (or S_2) of the system (1.2) is unstable.

- (ii) If $\Delta > 0$, $z_1^* > 0$ and $h(z_1^*) \leq 0$, (2.7) has two roots with positive real parts for $\tau \in [0, \tau_0)$, and the equilibrium S_1 (or S_2) of the system (1.2) is unstable.
- (iii) If the conditions of (ii) are satisfied, and $h'(z_k) \neq 0$, then system (1.2) exhibits the Hopf bifurcation at the equilibrium S_1 (or S_2) for $\tau = \tau_k^{(j)}$.

3. Direction and Stability of the Hopf Bifurcation

In the Section 2, we obtained some conditions which guarantee that the system (1.2) undergoes the Hopf bifurcation at a sequence values of τ . In this section, we shall study the direction and stability of the Hopf bifurcation. The method we used is based on the normal form theory and the center manifold theorem introduced by Hassard et al. [14]. Throughout this section, we always assume that system (1.2) undergoes Hopf bifurcations at the steady state (x_*, y_*, z_*) for $\tau = \tau_k$ and then $\pm i\omega_k$ is corresponding purely imaginary roots of the characteristic equation at the steady state (x_*, y_*, z_*) .

Letting $x_1 = x - x_*$, $x_2 = y - y_*$, $x_3 = z - z_*$, $\bar{x}_i(t) = x_i(\tau t)$, $\tau = \tau_k + \mu$ and dropping the bars for simplification of notations, system (1.2) is transformed into an FDE in $C = C([-1, 0], R^3)$ as

$$\dot{x}(t) = L_\mu(x_t) + f(\mu, x_t), \quad (3.1)$$

where $x(t) = (x_1(t), x_2(t), x_3(t))^T \in R^3$, and $L_\mu : C \rightarrow R$, $f : R \times C \rightarrow R$ are given, respectively, by

$$\begin{aligned} L_\mu(\phi) &= (\tau_k + \mu) \begin{pmatrix} -a & -a & 0 \\ -az_* & K-1 & -ax_* \\ ay_* & ax_* & 0 \end{pmatrix} \begin{pmatrix} \phi_1(0) \\ \phi_2(0) \\ \phi_3(0) \end{pmatrix} + (\tau_k + \mu) \begin{pmatrix} 0 & 0 & 0 \\ 0 & -K & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \phi_1(-1) \\ \phi_2(-1) \\ \phi_3(-1) \end{pmatrix}, \\ f(\mu, \phi) &= (\tau_k + \mu) \begin{pmatrix} 0 \\ -a\phi_1(0)\phi_3(0) \\ a\phi_1(0)\phi_2(0) \end{pmatrix}. \end{aligned} \quad (3.2)$$

By the Riesz representation theorem, there exists a function $\eta(\theta, \mu)$ of bounded variation for $\theta \in [-1, 0]$, such that

$$L_\mu \phi = \int_{-1}^0 d\eta(\theta, 0) \phi(\theta) \quad (3.3)$$

for $\phi \in C[-1, 0]$.

In fact, we can choose

$$\eta(\theta, \mu) = (\tau_k + \mu) \begin{pmatrix} -a & -a & 0 \\ -az_* & K-1 & -ax_* \\ ay_* & ax_* & 0 \end{pmatrix} \delta(\theta) - (\tau_k + \mu) \begin{pmatrix} 0 & 0 & 0 \\ 0 & -K & 0 \\ 0 & 0 & 0 \end{pmatrix} \delta(\theta + 1), \quad (3.4)$$

where δ is the Dirac delta function. For $\phi \in C^1([-1, 0], (R^3)^*)$, define

$$A(\mu)\phi = \begin{cases} \frac{d\phi(\theta)}{d\theta}, & \theta \in [-1, 0), \\ \int_{-1}^0 d\eta(\mu, s)\phi(s), & \theta = 0, \end{cases} \quad (3.5)$$

$$R(\mu)\phi = \begin{cases} 0, & \theta \in [-1, 0), \\ f(\mu, \phi), & \theta = 0. \end{cases}$$

Then system (3.1) is equivalent to

$$\dot{x}_t = A(\mu)x_t + R(\mu)x_t, \quad (3.6)$$

where $x_t(\theta) = x(t + \theta)$ for $\theta \in [-1, 0]$.

For $\psi \in C^1([0, 1], R^3)$, define

$$A^*\psi(s) = \begin{cases} -\frac{d\psi(s)}{ds}, & s \in (0, 1], \\ \int_{-1}^0 d\eta^T(t, 0)\psi(-t), & s = 0, \end{cases} \quad (3.7)$$

and a bilinear inner product

$$\langle \psi(s), \phi(\theta) \rangle = \bar{\psi}(0) \cdot \phi(0) - \int_{\theta=-1}^0 \int_{\xi=0}^{\theta} \bar{\psi}^T(\xi - \theta) d\eta(\theta)\phi(\xi) d\xi, \quad (3.8)$$

where $\eta(\theta) = \eta(\theta, 0)$. Then $A = A(0)$ and $A^* = A^*(0)$ are adjoints operators.

By the discussion in Section 2, we know that $\pm i\omega_k \tau_k$ are eigenvalues of A , thus they are also eigenvalues of A^* .

By direct computation, we obtain that $q(\theta) = q_0 e^{i\theta\omega_k \tau_k}$, with

$$q(0) = (1, \alpha, \beta)^T = \left(1, -\frac{a + i\omega_k}{a}, \frac{a(y_* - x_*) - i\omega_k x_*}{i\omega_k} \right)^T, \quad (3.9)$$

is the eigenvector of A corresponding to $i\omega_k \tau_k$, and $q^*(s) = Dq_0^* e^{is\omega_k \tau_k}$, with

$$q_0^* = (1, \alpha^*, \beta^*)^T = \left(1, \frac{i\omega_k(a - i\omega_k)}{a^2 x_* y_* - i\omega_k a z_*}, \frac{(a - i\omega_k) a x_*}{a^2 x_* y_* - i\omega_k a z_*} \right)^T, \quad (3.10)$$

is the eigenvector of A^* corresponding to $-i\omega_k \tau_k$, where

$$D = \frac{1}{1 + \bar{\alpha}\alpha^* + \bar{\beta}\beta^* - K\tau_k \bar{\alpha}\alpha^* e^{i\omega_k \tau_k}}. \quad (3.11)$$

Using the same notation as in [14], we compute the coordinates to describe the center manifold C_0 at $\mu = 0$. Let x_t be the solution of (3.1) when $\mu = 0$. Define

$$z(t) = \langle q^*, x_t \rangle, \quad W(t, \theta) = x_t(\theta) - 2 \operatorname{Re}\{z(t)q(\theta)\}. \quad (3.12)$$

On the center manifold C_0 , we have

$$W(t, \theta) = W(z(t), \bar{z}(t), \theta), \quad (3.13)$$

where

$$W(z, \bar{z}, \theta) = W_{20}(\theta) \frac{z^2}{2} + W_{11}(\theta) z\bar{z} + W_{02}(\theta) \frac{\bar{z}^2}{2} + W_{30}(\theta) \frac{z^3}{6} + \dots, \quad (3.14)$$

z and \bar{z} are local coordinates for center manifold C_0 in the direction of q^* and \bar{q}^* . Note that W is real if x_t is real. We consider only real solutions. For the solution $x_t \in C_0$ of (3.1), since $\mu = 0$, we have

$$\begin{aligned} \dot{z}(t) &= i\tau_k \omega_k z + \bar{q}^*(\theta) f(0, W(z, \bar{z}, \theta) + 2 \operatorname{Re}\{zq(\theta)\}) \\ &= i\tau_k \omega_k z + \bar{q}^*(0) f(0, W(z, \bar{z}, 0) + 2 \operatorname{Re}\{zq(0)\}) \stackrel{\text{def}}{=} i\tau_k \omega_k z + \bar{q}^*(0) f_0(z, \bar{z}). \end{aligned} \quad (3.15)$$

We rewrite this equation as

$$\dot{z}(t) = i\tau_k \omega_k z(t) + g(z, \bar{z}), \quad (3.16)$$

where

$$g(z, \bar{z}) = \bar{q}^*(0) f_0(z, \bar{z}) = g_{20} \frac{z^2}{2} + g_{11} z\bar{z} + g_{02} \frac{\bar{z}^2}{2} + g_{21} \frac{z^2 \bar{z}}{2} + \dots. \quad (3.17)$$

Noticing

$$\begin{aligned} x_t(\theta) &= (x_{1t}(\theta), x_{2t}(\theta), x_{3t}(\theta)) = W(t, \theta) + zq(\theta) + \bar{z}\bar{q}(\theta), \\ q(\theta) &= (1, \alpha, \beta)^T e^{i\theta\omega_k\tau_k}, \end{aligned} \quad (3.18)$$

we have

$$\begin{aligned} x_{1t}(0) &= z + \bar{z} + W_{20}^{(1)}(0) \frac{z^2}{2} + W_{11}^{(1)}(0) z\bar{z} + W_{02}^{(1)}(0) \frac{\bar{z}^2}{2} + O(|(z, \bar{z})|^3), \\ x_{2t}(0) &= \alpha z + \bar{\alpha} \bar{z} + W_{20}^{(2)}(0) \frac{z^2}{2} + W_{11}^{(2)}(0) z\bar{z} + W_{02}^{(2)}(0) \frac{\bar{z}^2}{2} + O(|(z, \bar{z})|^3), \\ x_{3t}(0) &= \beta z + \bar{\beta} \bar{z} + W_{20}^{(3)}(0) \frac{z^2}{2} + W_{11}^{(3)}(0) z\bar{z} + W_{02}^{(3)}(0) \frac{\bar{z}^2}{2} + O(|(z, \bar{z})|^3). \end{aligned} \quad (3.19)$$

Thus, from (3.17), we have

$$\begin{aligned}
 g(z, \bar{z}) &= \bar{q}^*(0) f_0(z, \bar{z}) = \bar{D}\tau_k \left(1, \bar{\alpha}^*, \bar{\beta}^* \right) \begin{pmatrix} 0 \\ -ax_{1t}(0)x_{3t}(0) \\ ax_{1t}(0)x_{2t}(0) \end{pmatrix} \\
 &= -a\bar{D}\tau_k \bar{\alpha}^* \left[z + \bar{z} + W_{20}^{(1)}(0) \frac{z^2}{2} + W_{11}^{(1)}(0) z\bar{z} + W_{02}^{(1)}(0) \frac{\bar{z}^2}{2} + O(|(z, \bar{z})|^3) \right] \\
 &\quad \left[\beta z + \bar{\beta} \bar{z} + W_{20}^{(3)}(0) \frac{z^2}{2} + W_{11}^{(3)}(0) z\bar{z} + W_{02}^{(3)}(0) \frac{\bar{z}^2}{2} + O(|(z, \bar{z})|^3) \right] \\
 &\quad + a\bar{D}\tau_k \bar{\beta}^* \left[z + \bar{z} + W_{20}^{(1)}(0) \frac{z^2}{2} + W_{11}^{(1)}(0) z\bar{z} + W_{02}^{(1)}(0) \frac{\bar{z}^2}{2} + O(|(z, \bar{z})|^3) \right] \\
 &\quad \left[\alpha z + \bar{\alpha} \bar{z} + W_{20}^{(2)}(0) \frac{z^2}{2} + W_{11}^{(2)}(0) z\bar{z} + W_{02}^{(2)}(0) \frac{\bar{z}^2}{2} + O(|(z, \bar{z})|^3) \right].
 \end{aligned} \tag{3.20}$$

Comparing the coefficients of (3.17), we get

$$\begin{aligned}
 g_{20} &= -2a\bar{D}\tau_k (\bar{\beta}^* \alpha - \bar{\alpha}^* \beta), \\
 g_{11} &= 2a\bar{D}\tau_k (\bar{\beta}^* \operatorname{Re}(\alpha) - \bar{\alpha}^* \operatorname{Re}(\beta)), \\
 g_{02} &= -2a\bar{D}\tau_k (\bar{\beta}^* \bar{\alpha} - \bar{\alpha}^* \bar{\beta}), \\
 g_{21} &= -a\bar{D}\tau_k \bar{\alpha}^* \left[2W_{11}^{(3)}(0) + W_{20}^{(3)}(0) + 2\beta W_{11}^{(1)}(0) + \bar{\beta} W_{20}^{(1)}(0) \right] \\
 &\quad + a\bar{D}\tau_k \bar{\beta}^* \left[2W_{11}^{(2)}(0) + W_{20}^{(2)}(0) + 2\alpha W_{11}^{(1)}(0) + \bar{\alpha} W_{20}^{(1)}(0) \right].
 \end{aligned} \tag{3.21}$$

Since there are $W_{20}(\theta)$ and $W_{11}(\theta)$ in g_{21} , we need to compute them.

From (3.6) and (3.12), we have

$$\begin{aligned}
 \dot{W} = \dot{x}_t - \dot{z}q - \dot{\bar{z}}\bar{q} &= \begin{cases} AW - 2\operatorname{Re}\{\bar{q}^*(0)f_0q(\theta)\}, & \theta \in [-1, 0), \\ AW - 2\operatorname{Re}\{\bar{q}^*(0)f_0q(0)\} + f_0, & \theta = 0. \end{cases} \\
 &\stackrel{\text{def}}{=} AW + H(z, \bar{z}, \theta),
 \end{aligned} \tag{3.22}$$

where

$$H(z, \bar{z}, \theta) = H_{20}(\theta) \frac{z^2}{2} + H_{11}(\theta) z\bar{z} + H_{02}(\theta) \frac{\bar{z}^2}{2} + \dots \tag{3.23}$$

Expanding the above series and comparing the corresponding coefficients, we obtain

$$(A - 2i\tau_k \omega_k)W_{20}(\theta) = -H_{20}(\theta), \quad AW_{11}(\theta) = -H_{11}(\theta), \dots \tag{3.24}$$

From (3.22), we know that for $\theta \in [-1, 0)$,

$$H(z, \bar{z}, \theta) = -\bar{q}^*(0)f_0q(\theta) - q^*(0)\bar{f}_0\bar{q}(0) = -gq(\theta) + \bar{g}\bar{q}(\theta). \quad (3.25)$$

Comparing the coefficients with (3.23) gives that

$$H_{20}(\theta) = -g_{20}q(\theta) - \bar{g}_{02}\bar{q}(\theta), \quad (3.26)$$

$$H_{11}(\theta) = -g_{11}q(\theta) - \bar{g}_{11}\bar{q}(\theta). \quad (3.27)$$

From (3.24), (3.26) and the definition of A , it follows that

$$\dot{W}_{20}(\theta) = 2i\tau_k\omega_k W_{20}(\theta) + g_{20}q(\theta) + \bar{g}_{02}\bar{q}(\theta). \quad (3.28)$$

Notice that $q(\theta) = (1, \alpha, \beta)^T e^{i\theta\omega_k\tau_k}$, hence

$$W_{20}(\theta) = -\frac{ig_{20}}{\omega_k\tau_k}q(0)e^{i\theta\omega_k\tau_k} + \frac{i\bar{g}_{02}}{3\omega_k\tau_k}\bar{q}(0)e^{-i\theta\omega_k\tau_k} + E_1e^{2i\theta\omega_k\tau_k}, \quad (3.29)$$

where $E_1 = (E_1^{(1)}, E_1^{(2)}, E_1^{(3)})^T \in R^3$ is a constant vector.

Similarly, from (3.24) and (3.27), we can obtain

$$W_{11}(\theta) = -\frac{ig_{11}}{\omega_k\tau_k}q(0)e^{i\theta\omega_k\tau_k} + \frac{i\bar{g}_{11}}{\omega_k\tau_k}\bar{q}(0)e^{-i\theta\omega_k\tau_k} + E_2, \quad (3.30)$$

where $E_2 = (E_2^{(1)}, E_2^{(2)}, E_2^{(3)})^T \in R^3$ is also a constant vector.

In what follows, we shall seek appropriate E_1 and E_2 . From the definition of A and (3.24), we obtain

$$\int_{-1}^0 d\eta(\theta)W_{20}(\theta) = 2i\omega_k\tau_k W_{20}(0) - H_{20}(0), \quad (3.31)$$

$$\int_{-1}^0 d\eta(\theta)W_{11}(\theta) = -H_{11}(0), \quad (3.32)$$

where $\eta(\theta) = \eta(\theta, 0)$. By (3.22), we have

$$H_{20}(0) = -g_{20}q(0) - \bar{g}_{02}\bar{q}(0) + 2\tau_k \begin{pmatrix} 0 \\ -\beta \\ \alpha \end{pmatrix}, \quad (3.33)$$

$$H_{11}(0) = -g_{11}q(0) - \bar{g}_{11}\bar{q}(0) + 2\tau_k \begin{pmatrix} 0 \\ -\operatorname{Re}(\beta) \\ \operatorname{Re}(\alpha) \end{pmatrix}. \quad (3.34)$$

Substituting (3.29) and (3.33) into (3.31), we obtain

$$\left(2i\omega_k \tau_k I - \int_{-1}^0 e^{2i\theta\omega_k \tau_k} d\eta(\theta)\right) E_1 = 2\tau_k \begin{pmatrix} 0 \\ -\beta \\ \alpha \end{pmatrix}, \quad (3.35)$$

which leads to

$$\begin{aligned} \begin{pmatrix} 2i\omega_k + a & a & 0 \\ az_* & 2i\omega_k - K + 1 + Ke^{-2\omega_k \tau_k} & ax_* \\ -ay_* & -ax_* & 2i\omega_k \end{pmatrix} E_1 &= 2 \begin{pmatrix} 0 \\ -\beta \\ \alpha \end{pmatrix}, \\ E_1^{(1)} &= \frac{2a(2i\omega_k \beta + a\alpha x_*)}{A}, \\ E_1^{(2)} &= \frac{-2(2i\omega_k + a)(2i\omega_k \beta + a\alpha x_*)}{A}, \\ E_1^{(3)} &= \frac{2}{A} A_1, \end{aligned} \quad (3.36)$$

where

$$\begin{aligned} A &= \det \begin{pmatrix} 2i\omega_k + a & a & 0 \\ az_* & 2i\omega_k - K + 1 + Ke^{-2\omega_k \tau_k} & ax_* \\ -ay_* & -ax_* & 2i\omega_k \end{pmatrix}, \\ A_1 &= \det \begin{pmatrix} 2i\omega_k + a & a & 0 \\ az_* & 2i\omega_k - K + 1 + Ke^{-2\omega_k \tau_k} & -\beta \\ -ay_* & -ax_* & \alpha \end{pmatrix}. \end{aligned} \quad (3.37)$$

Similarly, substituting (3.30) and (3.34) into (3.32), we can get

$$\begin{aligned} \begin{pmatrix} a & a & 0 \\ az_* & 1 & ax_* \\ -ay_* & -ax_* & 0 \end{pmatrix} E_2 &= 2 \begin{pmatrix} 0 \\ -\operatorname{Re}(\beta) \\ \operatorname{Re}(\alpha) \end{pmatrix}, \\ E_2^{(1)} &= \frac{2a^2 x_* \operatorname{Re}(\alpha)}{B}, \\ E_2^{(2)} &= \frac{-2a^2 x_* \operatorname{Re}(\alpha)}{B}, \\ E_2^{(3)} &= \frac{2}{B} B_1, \end{aligned} \quad (3.38)$$

where

$$B = \det \begin{pmatrix} a & a & 0 \\ az_* & 1 & ax_* \\ -ay_* & -ax_* & 0 \end{pmatrix}, \quad B_1 = \det \begin{pmatrix} a & a & 0 \\ az_* & 1 & -\operatorname{Re}(\beta) \\ -ay_* & -ax_* & \operatorname{Re}(\alpha) \end{pmatrix}. \quad (3.39)$$

Thus, we can determine $W_{20}(0)$ and $W_{11}(0)$ from (3.29) and (3.30). Furthermore, we can determine g_{21} . Therefore, each g_{ij} in (3.21) is determined by the parameters and delay in (3.1). Thus, we can compute the following values:

$$\begin{aligned} c_1(0) &= \frac{i}{2\tau_k\omega_k} \left(g_{11}g_{20} - 2|g_{11}|^2 - \frac{|g_{02}|^2}{3} \right) + \frac{g_{21}}{2}, \\ \mu_2 &= -\frac{\operatorname{Re}(c_1(0))}{\operatorname{Re}(\lambda'(\tau_k))}, \\ \beta_2 &= 2\operatorname{Re}(c_1(0)), \\ T_2 &= -\frac{\operatorname{Im}(c_1(0)) + \mu_2\operatorname{Im}(\lambda'(\tau_k))}{\tau_k\omega_k}, \end{aligned} \quad (3.40)$$

which determine the quantities of bifurcating periodic solutions in the center manifold at the critical value τ_k , that is, μ_2 determines the directions of the Hopf bifurcation: if $\mu_2 > 0$ ($\mu_2 < 0$) then the Hopf bifurcation is supercritical (subcritical) and the bifurcating periodic solutions exist for $\tau > \tau_k$ ($\tau < \tau_k$); β_2 determines the stability of the bifurcating periodic solutions: the bifurcating periodic solutions are stable (unstable) if $\beta_2 < 0$ ($\beta_2 > 0$); and T_2 determines the period of the bifurcating periodic solutions: the period increases (decreases) if $T_2 > 0$ ($T_2 < 0$).

4. Application to Control Chaos

In the present section, we apply the results in the previous sections to system (1.2) for the purpose of control of chaos. From Section 2, we know that under certain conditions, a family of periodic solutions bifurcate from the steady states of system (1.2) at some critical values of τ and the stability of the steady state maybe change along with increase of τ . If the bifurcating periodic solution is orbitally asymptotically stable or some steady state becomes local stable, then chaos may vanish. Following this ideal, we consider the following delayed feedback control system:

$$\begin{aligned} \dot{x}(t) &= -1.69(x(t) + y(t)), \\ \dot{y}(t) &= -y(t) - 1.69x(t)z(t) + K(y(t) - y(t - \tau)), \\ \dot{z}(t) &= 4 + 1.69x(t)y(t), \end{aligned} \quad (4.1)$$

which has two steady states $S_+ \triangleq (0.65, -0.65, 0.5917)$, $S_- \triangleq (-0.65, 0.65, 0.5917)$. Clearly, when $\tau = 0$ or $K = 0$, system (3.1) is chaotic (as depicted in Figure 1).

For the steady state S_+ or S_- , we have the corresponding characteristic equation of system (4.1) as follows:

$$\lambda^3 + (2.69 - K)\lambda^2 + (6.76 - 1.69K)\lambda + 22.8488 + K(\lambda^2 + 1.69\lambda)e^{-\lambda\tau} = 0. \quad (4.2)$$

Clearly, when $\tau = 0$, (4.2) has a negative root and a pair of complex roots with positive real parts. Following Section 2, we can obtain $p = -2K - 6.2839$, $q = 22.8488K - 77.2289$, $r = 522.0677 > 0$, $\Delta = p^2 - 3q > 0$, and $z_1^* = (1/3)(-p + \sqrt{\Delta}) > 0$ for all $K \in \mathbb{R}$. When

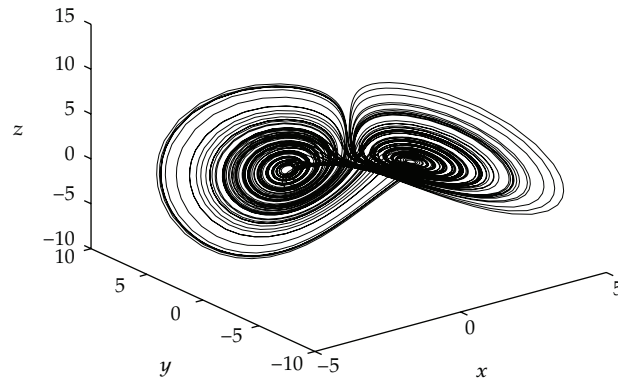


Figure 2: Chaos still exists for $K = -1$, $\tau = 0.2$.

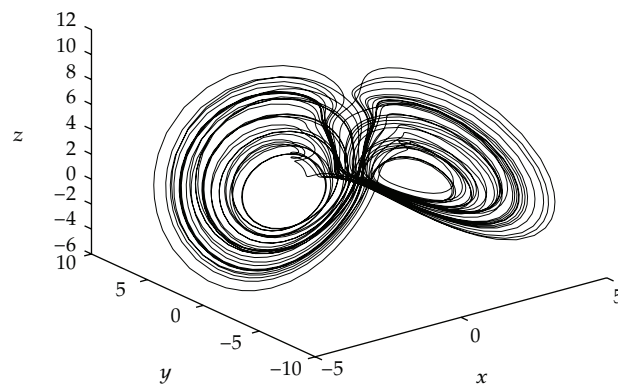


Figure 3: Chaos still exists for $K = -1$, $\tau = 2.5$.

$K < -0.1907$ or $K > 12.107$, $h(z^*) < 0$. Thus, from Lemma 2.6 and Theorem 2.9, we know that (4.2) has roots with positive real parts. In particular, we have $K = -1$, that is,

$$\begin{aligned} \dot{x}(t) &= -1.69(x(t) + y(t)), \\ \dot{y}(t) &= -y(t) - 1.69x(t)z(t) - (y(t) - y(t - \tau)), \\ \dot{z}(t) &= 4 + 1.69x(t)y(t). \end{aligned} \quad (4.3)$$

In this case, we can compute

$$\begin{aligned} p &\doteq -4.2839, & q &\doteq -100.077, & r &\doteq 522.0677, & \Delta &\doteq 318.5850, \\ z_1 &\doteq 8.9478, & \omega_1 &= 2.9913, & \tau_1^{(j)} &\doteq 0.2205 + \frac{2j\pi}{\omega_1}, & h'(z_1) &\doteq 63.4487, \\ z_2 &\doteq 5.6545, & \omega_2 &= 2.3779, & \tau_2^{(j)} &\doteq 0.5227 + \frac{2j\pi}{\omega_2}, & h'(z_2) &\doteq -52.6043. \end{aligned} \quad (4.4)$$

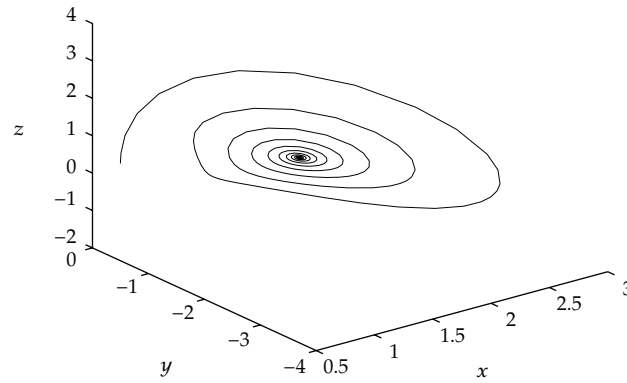


Figure 4: When $K = -1$, $\tau = 0.8$, chaos vanishes, and S_1 becomes local stable. Here initial value is $(0.5, -0.5, 0.6)$.

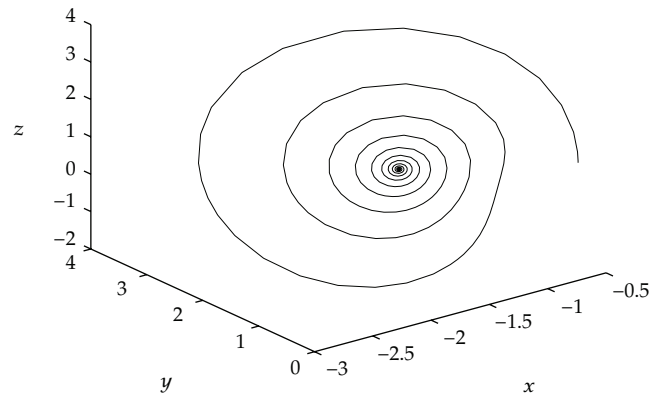


Figure 5: When $K = -1$, $\tau = 0.8$, chaos vanishes, and S_2 becomes local stable. Here initial value is $(-0.5, 0.5, 0.6)$.

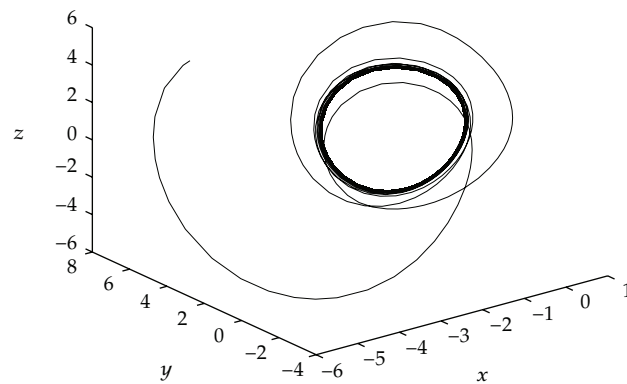


Figure 6: When $K = -1$, $\tau = 2$, chaos vanishes, and S_2 becomes a stable periodic solution. Here initial value is $(5, -5, 5)$.

Thus, from Lemma 2.7, we have $\operatorname{Re} \lambda(\tau_1^{(j)})/d\tau > 0$ and $\operatorname{Re} \lambda(\tau_2^{(j)})/d\tau < 0$. In addition, notice that

$$\tau_1^{(0)} \doteq 0.2205 < \tau_2^{(0)} \doteq 0.5227 < \tau_1^{(1)} \doteq 2.3210 < \tau_2^{(1)} \doteq 3.1650. \quad (4.5)$$

Thus, from Theorem 2.8, we have the following conclusion about the stability of the steady states of system (4.3) and Hopf bifurcation.

5. Conclusion

Suppose that $\tau_k^{(j)}$, $k = 1, 2$; $j = 0, 1, 2, \dots$ is defined by (4.4).

- (i) When $\tau \in [0, \tau_2^{(0)}) \cup (\tau_1^{(1)}, \infty)$, the steady states S_1 and S_2 of the system (4.1) are unstable (see Figures 2 and 3).
- (ii) When $\tau \in (\tau_2^{(0)}, \tau_1^{(1)})$, the steady states S_1 and S_2 of the system (4.1) are asymptotically (see Figures 4 and 5).
- (iii) When $\tau = \tau_k^{(j)}$, system (4.1) undergoes a Hopf bifurcation at the steady states S_1 and S_2 .

The above simulations indicate that when the steady state is stable or the bifurcating periodic solutions are orbitally asymptotically stable, chaos vanishes (see Figures 4–6).

References

- [1] J. H. Ma and Y. S. Chen, "Study for the bifurcation topological structure and the global complicated character of a kind of nonlinear finance system. I," *Applied Mathematics and Mechanics*, vol. 22, no. 11, pp. 1240–1251, 2001.
- [2] J. H. Ma and Y. S. Chen, "Study for the bifurcation topological structure and the global complicated character of a kind of nonlinear finance system. II," *Applied Mathematics and Mechanics*, vol. 22, no. 12, pp. 1375–1382, 2001.
- [3] J. H. Ma, B. Ren, and Y. S. Chen, "Impulsive control of chaotic attractors in nonlinear chaotic systems," *Applied Mathematics and Mechanics*, vol. 25, no. 9, pp. 889–894, 2004.
- [4] J. Ding, W. Yang, and H. Yao, "A new modified hyperchaotic finance system and its control," *International Journal of Nonlinear Science*, vol. 8, no. 1, pp. 59–66, 2009.
- [5] W.-C. Chen, "Dynamics and control of a financial system with time-delayed feedbacks," *Chaos, Solitons and Fractals*, vol. 37, no. 4, pp. 1198–1207, 2008.
- [6] Q. Gao and J. Ma, "Chaos and Hopf bifurcation of a finance system," *Nonlinear Dynamics*, vol. 58, no. 1-2, pp. 209–216, 2009.
- [7] Y. Wang, Y. H. Zhai, and J. Wang, "Chaos and Hopf bifurcation of a finance system with distributed time delay," *International Journal of Applied Mathematics and Mechanics*, vol. 6, pp. 1–13, 2010.
- [8] W.-S. Son and Y.-J. Park, "Delayed feedback on the dynamical model of a financial system," *Chaos, Solitons & Fractals*, vol. 44, no. 4-5, pp. 208–217, 2011.
- [9] E. Ott, C. Grebogi, and J. A. Yorke, "Controlling chaos," *Physical Review Letters*, vol. 64, no. 11, pp. 1196–1199, 1990.
- [10] K. Pyragas, "Continuous control of chaos by self-controlling feedback," *Physics Letters A*, vol. 170, no. 6, pp. 421–428, 1992.
- [11] S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, vol. 2 of *Texts in Applied Mathematics*, Springer, New York, NY, USA, 2nd edition, 2003.
- [12] S. Ruan and J. Wei, "On the zeros of transcendental functions with applications to stability of delay differential equations with two delays," *Dynamics of Continuous, Discrete & Impulsive Systems A*, vol. 10, no. 6, pp. 863–874, 2003.

- [13] Y. Song and J. Wei, "Bifurcation analysis for Chen's system with delayed feedback and its application to control of chaos," *Chaos, Solitons and Fractals*, vol. 22, no. 1, pp. 75–91, 2004.
- [14] B. D. Hassard, N. D. Kazarinoff, and Y. H. Wan, *Theory and Applications of Hopf Bifurcation*, vol. 41 of *London Mathematical Society Lecture Note Series*, Cambridge University Press, Cambridge, UK, 1981.

Research Article

Modeling Optimal Scheduling for Pumping System to Minimize Operation Cost and Enhance Operation Reliability

**Yin Luo, Shouqi Yuan, Yue Tang, Jianping Yuan,
and Jinfeng Zhang**

*Research Center of Fluid Machinery Engineering and Technology, Jiangsu University,
Zhenjiang 212013, China*

Correspondence should be addressed to Yin Luo, luoyin6@gmail.com

Received 20 October 2011; Revised 18 January 2012; Accepted 22 January 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 Yin Luo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Traditional pump scheduling models neglect the operation reliability which directly relates with the unscheduled maintenance cost and the wear cost during the operation. Just for this, based on the assumption that the vibration directly relates with the operation reliability and the degree of wear, it could express the operation reliability as the normalization of the vibration level. The characteristic of the vibration with the operation point was studied, it could be concluded that idealized flow versus vibration plot should be a distinct bathtub shape. There is a narrow sweet spot (80 to 100 percent BEP) to obtain low vibration levels in this shape, and the vibration also follows similar law with the square of the rotation speed without resonance phenomena. Then, the operation reliability could be modeled as the function of the capacity and rotation speed of the pump and add this function to the traditional model to form the new. And contrast with the tradition method, the result shown that the new model could fix the result produced by the traditional, make the pump operate in low vibration, then the operation reliability could increase and the maintenance cost could decrease.

1. Introduction

As important aspects in engineering industries, low cost and high reliability are the focus of the operation control in pumping system [1, 2].

The purpose of pump scheduling function is to schedule the operation of N pumps over a time period to meet consumer demands, and optimizing this function has been proven to be a practical and highly effective method in reducing operation costs without altering the actual infrastructure of the whole system. Thus, this issue naturally draws the attention of researchers [2].

Pump system scheduling should be robust with any operation scenarios and should deal virtually with all operation factors, such as variable speeds, constant speeds, and switched-off pumps, in relation with operation constraints relative to power, head, flow, and speed. It is a very complex problem.

Many researchers have developed optimal control concepts to minimize operating costs associated with water-supply pumping systems. Mays [3] listed and classified various algorithms that have been developed to solve the associated control problem. In earlier studies, linear, nonlinear, integer, dynamic, mixed, and other kinds of programming were used to optimize a single objective: reduction of the electric energy cost. A detailed review of these works can also be found in [4]. Later, Lansey et al. [5] introduced the number of pump switches as an alternative to evaluate the pumps' maintenance costs, which became the second objective considered until that time. This method also proved that one Gulf Coast refinery project [6] can save about \$2 million per year. In this way, as far as pump operation is concerned, the basic optimal scheduling model to reduce electrical and maintenance costs was formulated. The development of the pump scheduling model was later modified by forming or adding some optimization objectives, such as reservoir level variation and power peak, according to specific conditions.

However, in contrast to the estimation of electrical cost, where the computation is straightforward, computing for maintenance costs using the number of pump switches has some limitations.

According to some researchers [7–10], the cost of unscheduled maintenance which was not considered in former model may be the largest contributor to operation cost in process plants, and although with the mainly aspect of the wear and tear in switch course, the operation course should not been ignored.

At the same time, Bloch and Geitner [7] published three hydraulic factors, namely, rotation speed, impeller diameter (tip clearance), and operation point that can affect operation reliability which affects the unscheduled maintenance directly.

Subsequently, further improvements can be realized if these hydraulic factors could be quantified to develop an objective model in pump scheduling to enable the pump to operate in a high-reliability mode, then the reliability incidents would be reduced. As a result, maintenance cost can be reduced [10].

Vibration is one of the most vexing problems of pumping machineries, and it is the primary cause of considerable altercations and litigations. Excessive vibration of pumps and piping can destroy parts of the equipment (such as drive shafts, bearings, and seals). It can affect the reliability and life of the equipment and is often assumed as direct sign of reliability and health of the machine. One major end user puts a great deal of emphasis in reducing pump vibration through precise maintenance program, and some researchers adopted this index to indicate the degree of health and reliability of pumps. For example, Perez proposed a new method based on vibration called Nelson plot to assess the risk of low-flow operation in centrifugal pumps [11].

The main purpose of the current study is to set the vibration level as quantification of the hydraulic factors that influence operation reliability and the wear and tear in operation. Based on analysis of the vibration characteristics, an objective model is formulated to describe the relative reliability during operation, which presents a quantitative approach to evaluate alternative operating conditions. This new objective is added to the traditional model to form a new scheduling model. Subsequently, the operating conditions of the pump can be improved by making the pump operate at low vibration, the maintenance cost will be reduced, and the operation reliability will be enhanced to a certain degree.

2. Vibration Characteristics of Pumps

For centrifugal pumps, the sources or causes of vibration are mainly composed of two types: mechanical and hydraulic.

For the mechanical cause, some imbalance and shaft misalignment always exist. This can cause vibration of the pump, and the intensity of this kind of vibration is related to the excitation force caused by the imbalance of the load. The load is directly related to the operation point.

Hydraulic vibration is caused by the reaction of the impeller vane as it passes the casing cutwater. Pump operation offers the best efficiency point (BEP) and thus creates eddies within the pump and some flow instabilities, such as cavitation. All these factors directly correlate with the operation point.

As throttle and speed controls are the most common control methods in centrifugal pumping system, the vibration characteristics under these two control operating modes were analyzed.

2.1. Vibration Characteristics of the Pump under Throttle Control

Basically, three types of vibrations should be distinguished: free vibrations, forced vibrations, and self-excited vibrations. Whereas free vibrations are rarely significant in pump operations, forced and self-excited vibrations frequently cause problems. For the forced vibrations, the force is the cause of the vibration, and self-excited vibration occurs when the exciting frequency is close to the natural frequency; it is an abnormal condition that should be avoided in system design and installation. Therefore, the vibration characteristics can be obtained by studying the mechanical and hydraulic force characteristics.

A force curve is shown in Figure 1 [12], summarized by Vorhoeven through test results from 47 pumps. The horizontal axis represents the BEP multiples, and the vertical axis represents the dimensionless number of the forces, expressed as

$$K_R = \frac{F}{\rho g H D_2 b_2}, \quad (2.1)$$

where F is the force, ρ is the fluid density, H is the pump head, D_2 is the outer diameter of the impeller, and b_2 is the outlet width of the blade.

From the curve, the force caused by imbalance has a remarkable characteristic: it does not become small because of the BEP, and it is related to the load. When the load exceeds the BEP, the force significantly increases, and this force is relatively stable and small when the pump operates with partial load.

For the force generated by the flow, some features can be found. The minimum excitation is in the range of 0.9 BEP to BEP, and the excitation increases with partial load and at high flow condition. These curves are displayed in distinct bathtub shape. The stronger excitation caused by the interaction has a large slope at the right of the BEP, whereas the other has smaller slope at the left of the BEP.

From the above discussion, conclusion can be drawn that the idealized flow versus vibration plot should have a distinctive bathtub shape. For this bathtub shape, a narrow sweet spot (80% to 100% BEP) exists that can be used to obtain low vibration levels. The

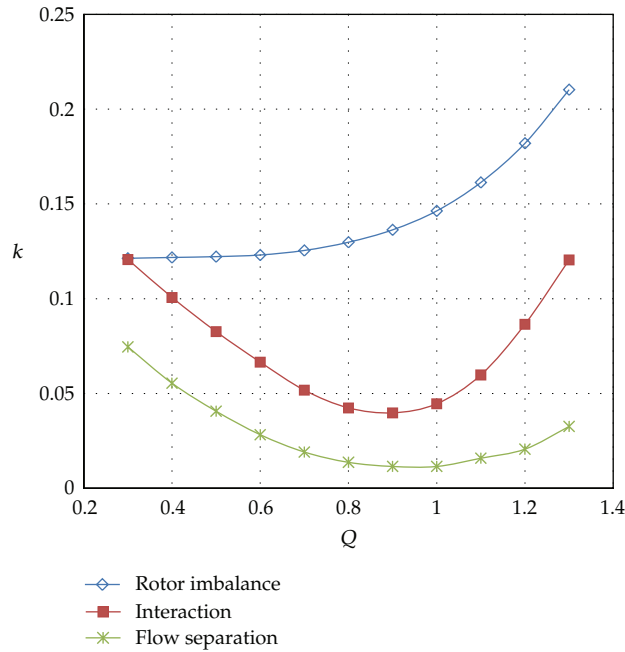


Figure 1: Forces in the impeller.

capacities move to the left and right of the low vibration level point, the vibration levels begin to increase, and the slope at the right area is larger than that at the left.

This bathtub-shaped flow versus vibration plot also can be found in some research works [9–11].

Figure 2 shows the test result of a well-designed centrifugal pump obtained by Ni et al. [13]. The pump was precision manufactured and installed. In Figure 2, BX, BY, and BZ are three direction vibration levels of the bearing house, whereas pc is the vibration level of the pump cover. The flow versus vibration plot in that test is exactly as the expression, and similar result can also be found in Perez's work [11].

2.2. Vibration Characteristics of the Pump under Speed Control

Mechanical and hydraulic excitations are the main causes of vibration.

According to the affinity law in pumps, when a pump is operating at two different speeds, the flow condition in the pump is homologous. Thus, when the pump operates at different rotation speeds, the variable tendency of hydraulic excitation versus the flow is still the same as that at different speeds, but the degree of some flow phenomena, such as instability interaction, will change under different speeds, especially at low speed.

The mechanical excitation is mostly related to the balance degree and the load. The load still has some homologous regularities expressed as the variable tendency of hydraulic excitation versus flow, which is almost the same at different speeds. In addition, once the speed drops below the nominal motor speed, the mechanical excitation is reduced quite obviously.

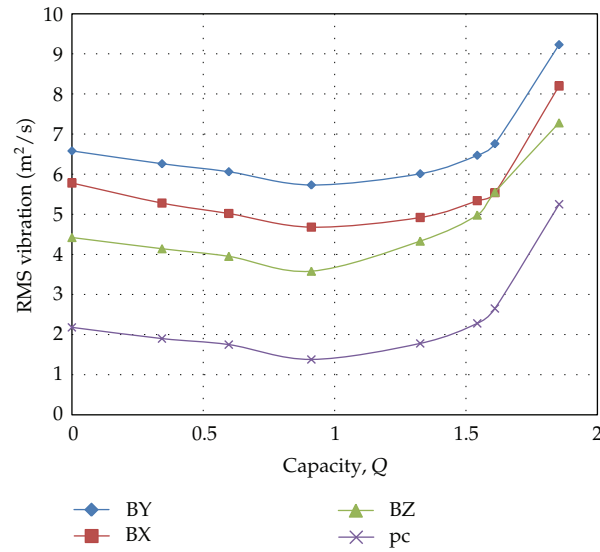


Figure 2: Vibration level in a well-designed centrifugal pump by Ni.

Based on the above results, the excitation versus flow at different rotor speeds is noted to have similar varying tendency to a certain extent, but the degree of excitation is different. Therefore, the vibrations among different rotor speeds from this point have similar varying tendency to a certain extent.

Figure 3 [10] shows the test vibration level dates of a pump in a chemical process with model $1.5 \times 3-13$ ANSI B 73.1. The measured point of vibration is at the thrust bearing horizontal plane, and the date was adopted with root mean square (RMS) expression. The vibrations among different rotor speeds have similar varying tendency to a certain extent.

However, in centrifugal pumps, the vibration spectrum usually contains only the fundamental shaft frequency, the blade passage frequency, and one or two harmonics. All these characteristic frequencies are directly decided by the rotor speed. Therefore, potential dangers of resonance exist, and when resonance occurs at the characteristic frequency close to the natural frequency, the similar varying tendencies are gone. Thus, in practice, lockout speed range is recommended.

As all forces are proportional to the product of pressure \times area, $F \sim n^2 \times d^4$ also applies.

For a given pump, the hydraulic excitation forces increase with the square of the tip speed and the density, that is, $\rho \times u^2$. The bearing housing and shaft vibrations of the pump would also increase with $\rho \times u^2$ if it were not for resonance phenomena.

In Stavale's work [10], the overall vibration was measured for a variable-speed test and compared with a constant-speed system with throttle valve, as shown in Figure 4. During the variable-speed test, the system was fixed, that is, there was no throttling of the control valve to initiate changes in the flow. Prior to the variable-speed test, the control valve was opened wide, and the backpressure valve was adjusted to obtain maximum flow at maximum speed; the backpressure was approximately zero.

Figure 4 shows that in the course of the variable-speed operation, the pump can operate by following the same operating condition line for the approximately zero backpressure. The vibration level of the square of the flow under the same operating

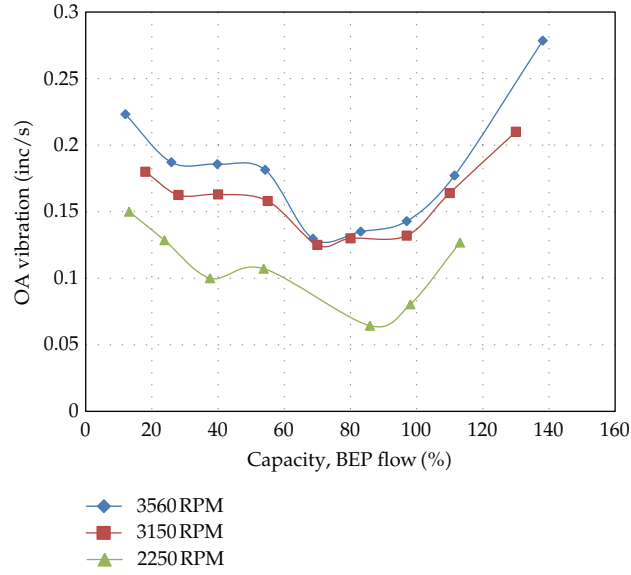


Figure 3: Test vibration level dates in different pump obtained by Stavale.

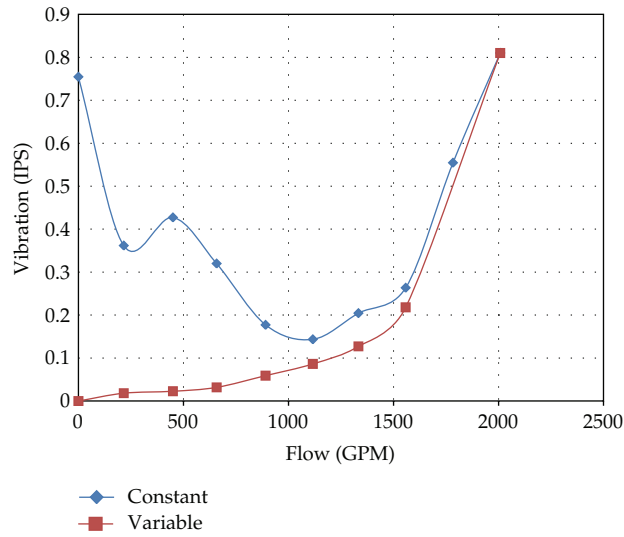


Figure 4: Vibration levels for different operating models of the same pump.

condition line can be clearly seen, so the vibration also follows the same law for the resonance phenomena of a given pump $V_e \sim n^2$, expressed as

$$\frac{V_e(n_1)}{V_e(n_2)} = \frac{n_1^2}{n_2^2}, \quad (2.2)$$

where n_1 and n_2 represent two different operating pump speeds.

3. Maintenance Cost Model in Pump Operation

Maintenance costs cannot be easily estimated, however, the wear of pumps is mainly caused by frequent switching them on and off. Formally, a pump switch is defined as turning on a pump which was previously off. Therefore, minimizing the number of pump switches will result in minimization of maintenance costs. As a result, the traditional method assumes that it increases as the number of “pump switched” increases, and its model can be built according to Lansey and Awumah [5] and Vladimir et al. [14].

In many applications, the cost of unscheduled maintenance which was not considered in former model may be the largest contributor to operation cost in process plants, and although with the mainly aspect of the wear in switch course, the operation course should not been ignored.

The life of the mechanical seal is directly related to shaft movement. Vibration can cause carbon face chipping and seal face opening. Drive lugs will wear, and metal bellows seals will fatigue. In some instances, the shaft movement can cause the rotating seal components to contact the inside of the stuffing box, or some other stationary object, causing the seal faces to open and allowing solids to penetrate between the lapped faces. Vibration is also a major cause of set screws becoming loose and slipping on the shaft, causing the lapped seal faces to open. The vibration would also cause denting of the bearing races for no design for the bearings to handle both a radial and axial load.

Critical dimensions and tolerances such as wear ring clearance and impeller setting will be affected by vibration.

Bearing seals are very sensitive to shaft radial movement. Shaft damage will increase and the seals will fail prematurely. Labyrinth seals operate with a very close tolerance. Excessive movement can damage these tolerances also.

Vibration can affect the reliability and life of the equipment and is often assumed to be the direct as direct sign of the reliability and health of the machine.

As vibration can affect the reliability and life of the equipment, indicators must be set up to consider vibration data. Maintenance cost is further assumed to increase as the reliability indicators reduce. In addition, maintenance cost can be significantly reduced with operation in a high-reliability mode, thus reducing wear and reliability incidents.

3.1. Indicators of Operation Reliability

Vibration data can be normalized based on the following equation:

$$R = \left(1 - \left(\frac{V}{V_{\max}} \right) \right) + C, \quad (3.1)$$

where V is the vibration level at a certain point, V_{\max} is the maximum data value, C is a constant added to set the peak value of R (equal to one) equal to V_{\min}/V_{\max} , and R is the relative reliability indicator.

The R value in (3.1) is a relative reliability number between zero and one. A zero value does not necessarily indicate zero reliability but is rather intended to discourage running the pump at these conditions. Similarly, a value of one does not indicate infinite reliability but is intended to be a relative indicator of the best operating conditions for a given pump. As the mechanical design of a pump can also affect reliability, these values should not be used to

compare pumps of different designs or manufacturers. It is intended to compare the effect of alternative operating conditions on reliability.

3.2. Maintenance Cost Model during Operation Based on Reliability Indicators

3.2.1. Single-Pump Model

For a fixed-speed pump, the mathematical model can be expressed as a set of static relationships between the vibration level and flow rate. The pump vibration level is referred to as vibration velocity or vibration displacement, expressed as RMS or peak-to-peak values. The vibration level data are divided by the lowest of all the flows for simplification and normalization. The flow rate is also referred to as capability. Then, a single model can be expressed as a cubic or fourth-order polynomial equation for the bathtub-shaped curve:

$$V = \bar{v}_0 + \bar{v}_1 Q + \bar{v}_2 Q^2 + \bar{v}_3 Q^3, \quad (3.2)$$

where V is the normalized vibration level at a certain point, Q is the flow rate, and system parameters v_i are determined by specific pump vibration characteristics and can be identified by test data.

With regard to the variable-speed pump, the relationship between (3.2) and its parameters is motor-speed dependent. The affinity law in pump theory states the following:

$$\frac{Q(n_1)}{Q(n_2)} = \frac{n_1}{n_2}, \quad \frac{H(n_1)}{H(n_2)} = \frac{n_1^2}{n_2^2}, \quad \frac{P(n_1)}{P(n_2)} = \frac{n_1^3}{n_2^3}, \quad (3.3)$$

where n_1 and n_2 represent two different operating pump speeds.

Assuming that the pump model (3.2) for a VSD at a special speed n_0 is obtained and its indicators are v_0 , v_1 , v_2 , and v_3 , the pump model of the considered VSD for a given limited speed n has the property defined by (2.2) and (3.2), that is,

$$\begin{aligned} \frac{V(n)}{V(n_0)} &= \frac{V(n)}{v_0 + v_1 Q(n_0) + v_2 Q(n_0)^2 + v_3 Q(n_0)^3} = \frac{n^2}{n_0^2}, \\ V(n) &= \frac{n^2}{n_0^2} \left[v_0 + v_1 Q(n_0) + v_2 Q(n_0)^2 + v_3 Q(n_0)^3 \right]. \end{aligned} \quad (3.4)$$

Then, according to (3.3), with $k = n/n_0$, this part can be transformed into

$$\begin{aligned} V(n) &= \frac{n^2}{n_0^2} \left[v_0 + v_1 \left(Q(n) * \frac{n_0}{n_1} \right) + v_2 \left(Q(n) * \frac{n_0}{n_1} \right)^2 + v_3 \left(Q(n) * \frac{n_0}{n_1} \right)^3 \right] \\ &= k^2 v_0 + k v_1 Q(n) + v_2 Q^2(n) + k^{-1} v_3 Q^3(n). \end{aligned} \quad (3.5)$$

Subsequently, according to (3.1), the operation reliability of the signal pump can be expressed as

$$R_{\text{sig}} = \left(1 - \left(\frac{V(n, Q)}{V_{\text{max}}}\right)\right) + \frac{V_{\text{min}}}{V_{\text{max}}}. \quad (3.6)$$

R_{sig} is a function of Q (flow rate) and n (rotation speed). It is directly related to the operating conditions and reflects the relationship between the maintenance cost and the operating conditions. The higher R_{sig} is, the lower is the maintenance cost in the course of operation.

3.2.2. MultiPump Model

The general multiple pump systems consist of some special characteristics for operating maintenance cost analysis.

- (1) The model should only reflect the pumps that are in operation. Pumps that are not running should not be considered.
- (2) As maintenance cost is not the same for different pumps, the model should consider the difference.
- (3) The relative normalization number is also between zero and one, similar to the single-pump model.

Considering the above-mentioned factors, the following methods are adopted.

- (a) A swift variable vector is added to represent the operating condition of the pump group. Then, only the pump in operation is considered; hence, the program (2.1) can be solved.
- (b) As the cost of pump is usually related to the installed capacity, assumption is made that the bigger the installed capacity is, the higher is the maintenance cost. Thus, a weight constructed with the ratio of single installed capacity to total installed capacity is adopted to represent the different maintenance costs.

From the above discussion, conclusion is made that the maintenance cost model for multipump system with k pumps can be expressed as follows:

$$R = \sum_{i=1}^k \frac{w_i \varphi_i R_{\text{sig}i}}{w_i \varphi_i}, \quad (3.7)$$

where

$$w_i \in \{0, 1\}, \quad i = 1, 2, \dots, k. \quad (3.8)$$

This equation is the swift variable vector. “1” shows that the pump is working, and “0” indicates that the pump is off:

$$\varphi_i = \frac{N_{di}}{N_{dT}}, \quad i = 1, 2, \dots, k. \quad (3.9)$$

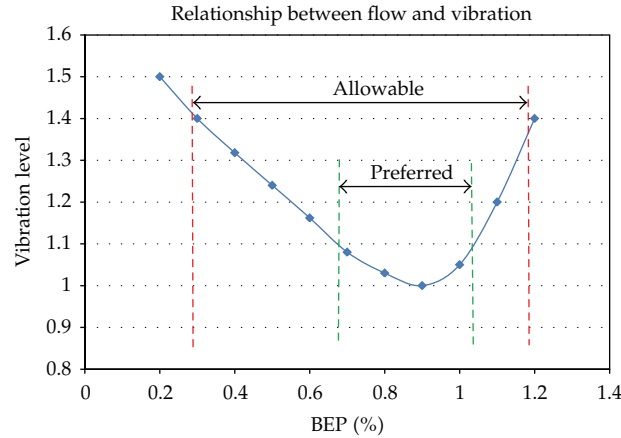


Figure 5: Idealized flow versus vibration plot recommended by API 610.

N_{dT} is the total installed capacity of the multipump system, and N_{di} is single installed capacity of the i th pump.

3.2.3. Modeling Method of the Maintenance Cost without Test Data

As environment and reliability requirements increase, the pumps are built and tested to standards and specifications that define the maximum allowable vibration amplitudes (process pumps and almost all large pumps) in many applications. However, unlike the pump characteristics, some installations still do not have such vibration data because no mandatory standard is required. Thus, the maintenance cost model may not work. Hence, other modeling methods must be developed to deal with this situation.

Fortunately, although specific vibration data may not be known, the allowable vibration amplitudes defined in standards, such as API 610 [15], are clear. The standards define different vibration limits for a “preferred operation range” and an “allowed operation range.” In the preferred operation range, the maximum increment of vibration is less than 10%. On the other hand, the maximum increment of vibration is less than 30% for the allowed operation range. In a certain pumping system, if the preferred and the allowed operation ranges are known, assumption can be made that all systems meet the standard. Therefore, a maintenance cost model can be developed.

Even if the preferred and allowed operation ranges are unknown, API 610 still provides an idealized flow versus vibration plot [11], as shown in Figure 5. Maintenance cost model can also be developed based on the assumption that the pump follows this law.

4. Optimization Model

Generally, the optimal policy should result in the lowest total operating cost for a given set of boundary and system constraints. Thus, objective function and constraints are needed for optimal scheduling model.

4.1. Objective Function

In a typical pumping system, the operation cost mostly comprises energy-consumption cost and maintenance cost. Thus, the objective function also comprises these two factors [16–19].

4.1.1. Energy Cost

Pump scheduling is used for dealing with the following two situations.

One is the selection process on which available pumps are to be used and for what period of time (e.g., a day) the pumps should operate.

The other is referred to as (real-time) control problem. The optimal strategy is concerned with which available pumps must be operated and when they should be operated according to fluctuations in demands and/or operating conditions.

For the first situation [14], the objective is the determination of energy consumption of all pumps in the pumping station during the optimization period. The charging structure used by the electric utility is the factor that must be considered in analyzing electric energy cost. Then, the objective function is mathematically expressed as

$$E = \sum_{i=1}^I \sum_{t=1}^T (ER_{it}) C Q_{it} (HS_{it}), \quad (4.1)$$

where E is the total energy cost to be minimized, I is number of pump systems, and T is number of time intervals that constitute the operating horizon. ER_{it} is the electrical rating of pump i during time period t , C is the conversion coefficient, and Q_{it} and HS_{it} are the discharge and pressure head, respectively, of pump i during time period t .

For the second situation [16], the real scheme is to consider the total input electrical power as the objective function. This method is much closer to the real situation. However, the calculation time will be long, which may not be suitable for real-time control. Thus, the modeling method commonly used is by considering the shaft horsepower of the pumps as objective function because the energy consumption of the motor and the inverter is very much less than that of the pump. Then, the objective function can be mathematically expressed as

$$f1 = \sum_{i=1}^I w_i P_i(Q_i, k_i), \quad (4.2)$$

where $f1$ is the total shaft horsepower to be minimized, I is number of pumps, and Q_i and k_i are the discharge and pressure head, respectively, of pump i under certain operating conditions. P_i is the shaft horsepower of pump i under this condition, which can be expressed as (3.9) according to the characteristic of the pump and the affinity law, and p_i is the conversion coefficient:

$$P(Q, k) = p_0 k^3 + p_1 k^2 Q + p_2 k Q^2 + p_3 Q^3. \quad (4.3)$$

4.1.2. Pump Maintenance Cost

The pump maintenance cost can be as important as the electric energy cost or even more relevant; however, this cost cannot be quantified easily, but it can be described through other correlative factors indirectly. Thus, the number of “pump switched” and the operation relative reliability indicators are chosen to describe this factor.

(1) The Number of “Pump Switched”

Due to the difficulty in starting-up the pumps and the significant increase in frequency during their switching, the maintenance cost increases as the number of “pump switches” increases, which can be expressed as follows when real control is used for the pump schedule:

$$d_H(w, w') = \sum_{i=1}^n |w_i - w'_i|, \quad (4.4)$$

where I is the number of pumps, w_i is the operating status of pump i for the next step, w'_i is the present operating status, w_i and w'_i are both switch variables, 1 indicates that the pump is in operation, and 0 shows that the pump is off. When the pump schedule is used for this scheme for a period of time, the objective function becomes the total switch number at this time, expressed as

$$N_s = \sum_{t=1}^T \sum_{i=1}^n |w_i^t - w_i^{t-1}|. \quad (4.5)$$

(2) Operation Relative Reliability Indicators

R in (3.6) and (3.7) is a relative reliability number between zero and one. As R approaches one, the higher the reliability is, the lower is the pump wear and the longer is the MTBR; subsequently, maintenance cost will be reduced. For the real-time control, (3.7) can be used directly to express the scheme in a period of time. The summation of the time periods can be used as the objective function.

4.2. Constraints

4.2.1. Water-Supply Demand Constraints

For the pump system to meet some of the water-supply requirements [17], the desired water-supply index, which is known for the pump system, can be expressed as (H_{ST}, Q_e) , expressed in a mathematic model.

For parallel-connected pump systems,

$$\begin{aligned} Q_e &= \sum_{i=1}^n Q_i, \\ H_{ST} &= H_i. \end{aligned} \quad (4.6)$$

For series-connected pump systems,

$$Q_e = Q_i, \quad (4.7)$$

$$H_{ST} = \sum_{i=1}^n H_i. \quad (4.8)$$

4.2.2. Operation Constraints

Obviously, a pump should be selected so that it operates predominantly close to the BEP in the so-called “preferred operation range.” This mode of operation is apt to produce the lowest energy and maintenance costs and minimize the risk of system problems. However, off-design operation for limited periods cannot be avoided. Rules are needed to define the allowable ranges and modes of operation to reduce the risk of damage and excessive wear. To this effect, limits must be defined for continuous and for short-term operations at maximum and minimum flow [16].

The range of preferred continuous operation can be defined, for instance, by the requirement that the efficiency must not fall below 80% to 85% of the maximum efficiency of the pump, and allowable ranges can be defined so that the efficiency must not fall below 70%.

This constraint can be expressed as

$$\min(Q_{POPi}) \leq Q_i \leq \max(Q_{POPi}), \quad (4.9)$$

$$\min(Q_{AOPi}) \leq Q_i \leq \max(Q_{AOPi}). \quad (4.10)$$

For optimal operation, the constraint should be in POP, as shown in (4.9). However, for some situations, the optimization model may have no solution. In this case, the constraint is adjusted to AOP; then, (4.10) can be applied.

At the same time, when the pump is operated by VSD, the speed regulation range should also be constrained. The operation efficiency will decrease because of a very wide speed range, and reliability will reduce. Therefore, the factors for operation should be considered. The speed regulation range is constrained as $[k_{\min} \ 1] \cdot k_{\min}$ is determined by several factors, such as the rotation speed at which the pump is no longer able to maintain discharge against the static head from the demand side, the rotation speed to avoid system resonance, and the rotation speed to ensure that the pump operating economically.

As the pump operates under speed-control model, the range of the continuous and short-term operations will constantly change.

According to the affinity law in pump theory, l_1 and l_2 (Figure 6) are set as similar lines, and A, B, C, and D are the boundaries of the operation zone; then, l_1 and l_2 can be expressed as

$$H_{l1} = \frac{H_A}{Q_A} Q_{l1}^2, \quad H_{l2} = \frac{H_B}{Q_B} Q_{l2}^2. \quad (4.11)$$

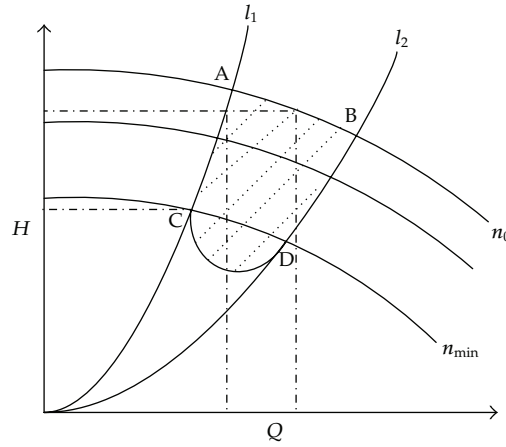


Figure 6: Efficient operation area for the variable-speed pump.

If head H_e is needed for this water supply system, then the boundary could be changed:

$$Q_{\min} = \begin{cases} Q_A \sqrt{\frac{H_e}{H_A}} & H_e \geq H_C, \\ Q_c & H_e < H_C, \end{cases} \quad (4.12)$$

$$Q_{\max} = \begin{cases} Q_B \sqrt{\frac{H_e}{H_B}} & H_e < H_B, \\ Q_B & H_e \geq H_B. \end{cases}$$

5. Application

5.1. Profiles of the Pump Station

A sample model is a circulating water pumping station, one of the most important facilities in an alumina plant used for the mother liquor evaporation process. It uses almost 17% of the electricity consumption in the plant, and the reliability of this system is very important for the whole plant.

There are five pumps of single- and double-stage suction in the pumping system, and the pump model is shown in Table 1.

Variable-speed control is adopted for 1[#] and 2[#], and the minimum speed regulation ranges (k_{\min}) are 0.7 and 0.75 in this system.

The operation characteristics of these pumps are shown in Table 2.

For this pumping system, the basic demand characteristic obeys (4.10) and can be calculated based on process requirements and pipeline characteristics.

$$H_{\text{dem}} = 42 + 3.51 \times 10^{-7} Q^2. \quad (5.1)$$

Table 1: The configuration and the parameters of the pump.

Serial number	Pump model	Q_{den} (m3/h)	H_{den} (m)	Q_{min} (m3/h)	Q_{max} (m3/h)	N (RPM)
1 [#]	14SH-9B	1425	58	855	1853	1450
2 [#]	20SA-10	2850	58	1710	3848	960
3 [#]	20SA-10	2850	58	1710	3848	960
4 [#]	20SA-10	2850	58	1710	3848	960
5 [#]	20SA-10	2850	58	1710	3848	960

Table 2: The model parameters of the pump.

$H = Hx - SQ^2$				
Serial number	H_{xi}	S_i		
1 [#]	$71.17\; k_1^2$	$7.488e - 6$		
2 [#]	$70.39\; k_2^2$	$1.780e - 6$		
3 [#]	70.39	$1.780e - 6$		
4 [#]	70.39	$1.780e - 6$		
5 [#]	70.39	$1.780e - 6$		
$P = P_0 + P_1Q + P_2Q^2 + P_3Q^3$				
Serial number	P_0	P_1	P_2	P_3
1 [#]	$146.4\; k_1^3$	$0.05\; k_1^2$	$4.4e - 5\; k_1$	$-1.44e - 8$
2 [#]	$230.5\; k_2^3$	$0.1025\; k_2^2$	$5.826e - 6\; k_2$	$-2.1e - 9$
3 [#]	230.5	0.1025	$5.826e - 6$	$-2.1e - 9$
4 [#]	230.5	0.1025	$5.826e - 6$	$-2.1e - 9$
5 [#]	230.5	0.1025	$5.826e - 6$	$-2.1e - 9$
$V = V_0 + V_1Q + V_2Q^2 + V_3Q^3$				
Serial number	V_0	V_1	V_2	V_3
1 [#]	$7.696e - 10k_1^2$	$-1.57e - 6\; k_1$	$3.84e - 4$	$1.483\; k_1^{-1}$
2 [#]	$9.62e - 11\; k_2^2$	$-3.936e - 7k_2$	$1.92e - 4$	$1.483\; k_2^{-1}$
3 [#]	$9.62e - 11$	$-3.936e - 7$	$1.92e - 4$	1.483
4 [#]	$9.62e - 11$	$-3.936e - 7$	$1.92e - 4$	1.483
5 [#]	$9.62e - 11$	$-3.936e - 7$	$1.92e - 4$	1.483

The whole flow date is shown in Figure 7, and these values are obtained from the water demand curve based on historical data.

5.2. Optimization Pump Scheduling and the Result

A one-day optimization result was made. Assumption was made that the shortest period for each combination of pumps is 1 h, that is, a pump can be switched off/on after having been active or inactive for at least 1 h.

Based on the flow demand in Figure 7 and the head demand in (4.10), the desired objective value is shown in Figure 8.

Genetic algorithm was selected as the optimization method because of its suitable characteristics for adaptability to complex optimization problem [18, 19].

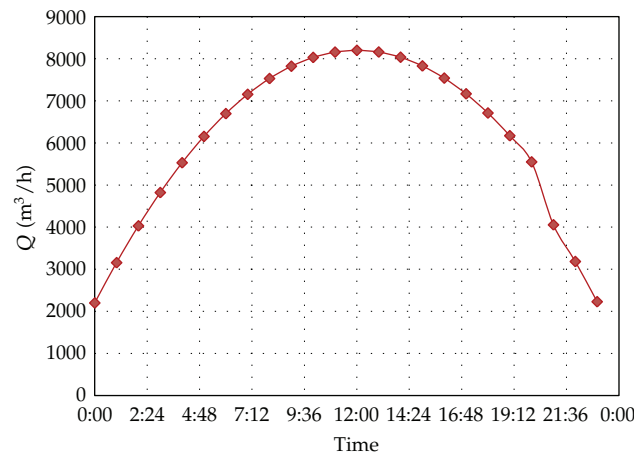


Figure 7: Daily variation curve of the entire flow demand.

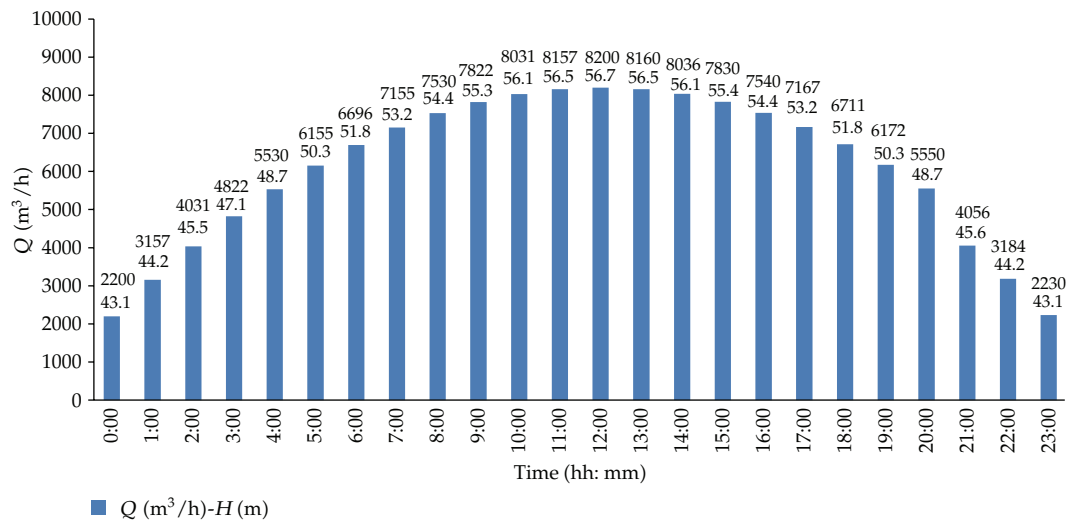


Figure 8: Daily supply index.

In this optimization problem, the single objective approach developed by Mackle was adopted because of its simplicity. The fitness function consists of the energy consumption cost and penalties for violation of the constraints of the system. All these factors were linearly weighted.

Then, the optimal result is shown in Figures 9, 10, and 11. A represents the result when only the minimum energy cost is considered, and B represents the result when both maintenance and energy costs are considered.

Figure 9 shows the result of the shaft power in the two different optimization objectives under the same water supply index. The result considered that the energy consumption was smaller, and the multiobjects are almost equal to the whole, whereas in some cases, they were smaller. This condition might be because the high efficiency point is the most reliable operation point in most situations; however, in some cases, the general shaft

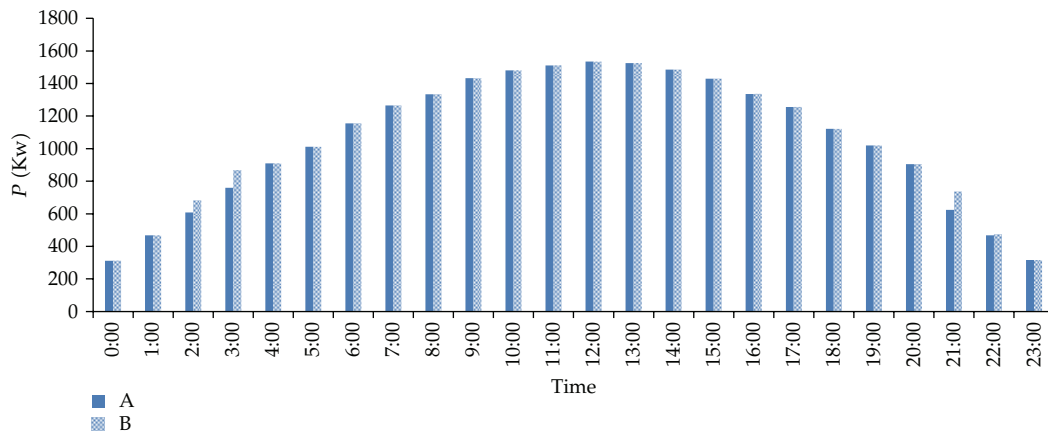


Figure 9: Input power of the different models.

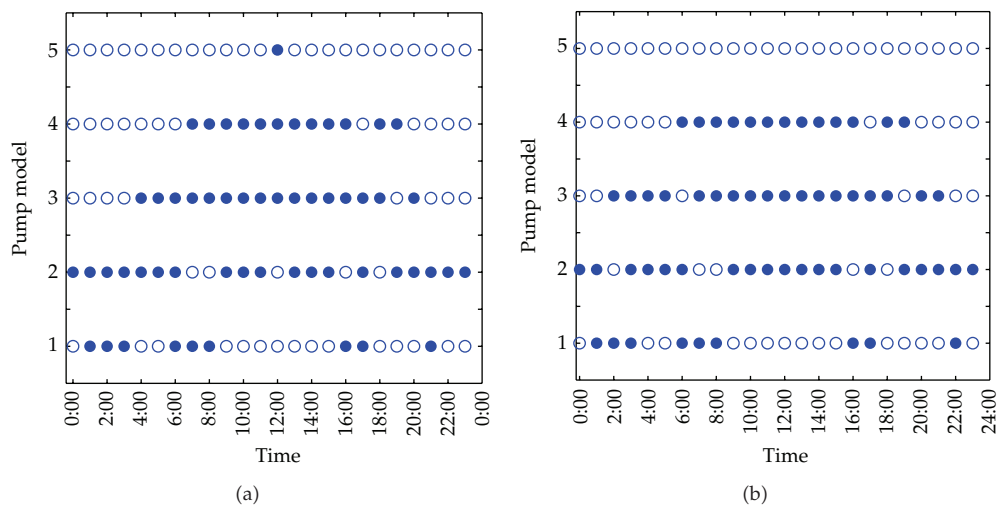


Figure 10: (a) Operation state of the form model. (b) Operation state of the new improved model.

power may be lower. For a single pump, operating in the high vibration point can cause the whole system to operate in unreliable conditions.

Thus, from energy-saving perspective, considering only energy cost as the main objective may be even better. However, when the switch number and reliability factor are considered in the optimization course, pump control are easy and simple using the least number of switches, and reliability will be enhanced when the operation point is close to the low vibration point as much as possible. Then, the pump system will be operating under high reliability condition, the maintenance cost will be reduced, and the service life will be extended. Subsequently, the whole LCC will be decreased by this kind of optimization scheduling model.

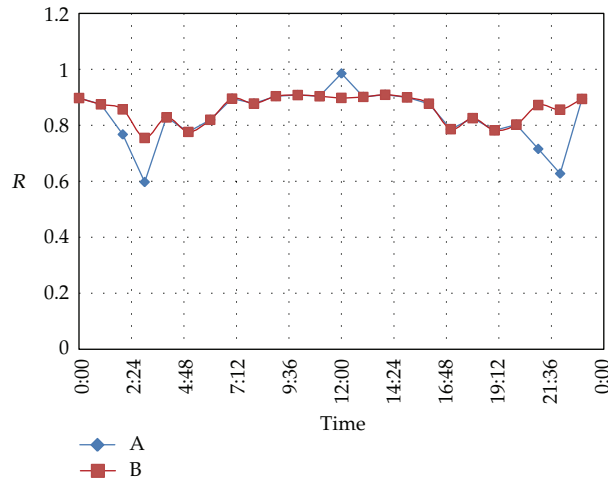


Figure 11: Reliability factor number of the different models.

6. Conclusions

In the present work, a new multiobjective approach, which takes into account operation reliability and maintenance cost incurred in operation, is presented.

The main conclusions from the current study are the following.

(1) Operation cost reduction and operation reliability enhancement are the focus of operation control for pump systems. Operation reliability for running pumps, which is a key to decrease unscheduled maintenance costs caused by the reliability incidents, and the wear cost during the operation should be considered.

(2) Vibration can affect the reliability and life of the equipment. An index reflecting the operation reliability and the wear degree in operation course could be considered as the normalization of vibration level.

(3) The idealized flow versus vibration plot should take a distinct bathtub shape. For this bathtub shape, a narrow sweet spot (80% to 100% BEP) exists that can be used to obtain low vibration levels, and the vibration also follows a similar law with the square of the rotation speed if it were not for resonance phenomena with a given pump.

(4) The maintenance cost which is not considered by traditional model concluding the unscheduled maintenance cost and the wear cost during the operation can be modeled as a function of the pump capacity and rotation speed based on the vibration characteristics. This function is then added to the traditional optimal scheduling model to create a new optimal scheduling model.

(5) Compared with the traditional method, the result of the new optimal model changes the result produced by the traditional one. It improves the operating conditions of the pump and enables the pump to run operate low vibration level. Therefore, maintenance cost can be reduced and the operation reliability can be enhanced to a certain degree.

Acknowledgments

This work is supported by National Outstanding Young Scientists Funds of China (Grant no. 50825902), Jiangsu Provincial Innovative Scholars "Climbing" Project of China (Grant no.

BK 2009006), the National Natural Science Foundation of China (Grant no. 50979034), and Priority Academic Program Development of Jiangsu Higher Education Institutions.

References

- [1] Hydraulic Institute, *Pump Life Cycle Costs: A Guide to LCC Analysis for Pumping Systems*, Parsippany, NJ, USA, 2001.
- [2] Z.-H. Wang, G.-H. Geng, and S.-K. Song, "Discussion on energy conservation of pump," *China Foreign Energy*, vol. 11, no. 5, pp. 73–76, 2006 (Chinese).
- [3] L. W. Mays, *Water Distribution Systems Handbook*, McGraw-Hill, New York, NY, USA, 2000.
- [4] L. E. Ormsbee and K. E. Lansey, "Optimal control of water-supply pumping systems," *Journal of Water Resources Planning & Management*, vol. 120, no. 2, pp. 237–252, 1994.
- [5] K. E. Lansey and K. Awumah, "Optimal pump operations considering pump switches," *Journal of Water Resources Planning and Management*, vol. 120, no. 1, pp. 17–35, 1994.
- [6] D. C. White, "Improve your project's prospects," *Chemical Processing*, vol. 67, no. 10, pp. 33–39, 2004.
- [7] H. P. Bloch and F. K. Geitner, *An Introduction Machinery Reliability Assessment*, Gulf Publishing, Houston, Tex, USA, 1994.
- [8] A. R. Budris, R. B. Erickson, F. H. Kludt, and C. Small, "Consider hydraulic factors to reduce pump downtime," *Chemical Engineering*, vol. 109, no. 1, pp. 54–60, 2002.
- [9] R. B. Erickson, E. P. Sabini, and A. E. Stavale, *Hydraulic Selection to Minimize the Unscheduled Maintenance Portion of Life Cycle Cost*, Pump Users International Forum, Karlsruhe, Germany, 2000.
- [10] A. E. Stavale, "Reducing reliability incidents and improving meantime between repair," in *Proceedings of the 24th International Pump Users Symposium*, pp. 1–10, College Station, Tex, USA, 2008.
- [11] X. Robert and P. E. Perez, "Operating Centrifugal Pumps Off-design- pumps & systems 20 suggestions for a new analysis method operating centrifugal pumps," April 2005, <http://www.pump-zone.com/articles/2.pdf>.
- [12] Y.-Y. Ni, *3-D unsteady numerical simulation and fluid-induced vibration for centrifugal pumps*, Ph.D. thesis, Jiangsu University, Zhenjiang, China, 2008.
- [13] Y.-Y. Ni, S.-Qi Yuan, Z.-Y. Pan et al., "Diagnosing the running condition of pump by its vibration character," *Drainage and Irrigation Machinery*, vol. 25, no. 2, pp. 49–52, 2007 (Chinese).
- [14] C. Vladimir, M. Heiliö, N. Krejić, and M. Nedeljkov, "Mathematical model for efficient water flow management," *Nonlinear Analysis*, vol. 11, no. 3, pp. 1600–1612, 2010.
- [15] API Standard 610, *Centrifugal Pumps for Petroleum Heavy Duty Chemical and Gas Industry Services*, Petrochemical and Natural Gas Industries, 9th edition, 2003.
- [16] C. Zhang, H. Li, M. Zhong, and J. Cheng, "The modelling and optimal scheduling for pressure and flow varying parallel-connected pump systems," *Dynamics of Continuous, Discrete and Impulsive Systems B*, vol. 11, no. 6, pp. 757–770, 2004.
- [17] T. C. Yu, T. Q. Zhang, and X. Li, "Optimal operation of water supply systems with tanks based on genetic algorithm," *Journal of Zhejiang University*, vol. 6, no. 8, pp. 886–893, 2005.
- [18] B. Barán, C. Von Lücken, and A. Sotelo, "Multi-objective pump scheduling optimisation using evolutionary strategies," *Advances in Engineering Software*, vol. 36, no. 1, pp. 39–47, 2005.
- [19] A. S. Dragan, A. W. Godfrey, and S. Martin, "Multiobjective genetic algorithms for pump scheduling in water supply," *Lecture Notes in Computer Science: Evolutionary Computing*, vol. 1305, pp. 227–223, 1997.